FRANTZ MARTINACHE

REPOUSSER LES LIMITES DE LA DIFFRACTION POUR L'ASTRONOMIE À HAUTE RÉSOLUTION ANGULAIRE

OBSERVATOIRE DE LA CÔTE D'AZUR

Copyright © 2018 Frantz Martinache

OBSERVATOIRE DE LA CÔTE D'AZUR

FRANTZMARTINACHE.EU

Première édition, Aout 2018

Table des matières

Introduction 7

Interférométrie à masque non redondant 11

Le projet SCExAO 63

Les noyaux de phase 183

Conclusion : vers un haut contraste robuste ? 273

Bibliographie 285

MUAD'DIB LEARNED RAPIDLY BECAUSE HIS FIRST TRAINING WAS IN HOW TO LEARN. AND THE FIRST LESSON OF ALL WAS THE BASIC TRUST THAT HE COULD LEARN. IT'S SHOCKING TO FIND HOW MANY PEOPLE DO NOT BELIEVE THEY CAN LEARN, AND HOW MANY MORE BELIEVE LEARNING TO BE DIFFICULT. MUAD'DIB KNEW THAT EVERY EXPERIENCE CARRIES ITS LESSON.

FRANK HERBERT, *DUNE*

Introduction

Ce mémoire est un résumé des activités de recherche que j'ai menées depuis la présentation de ma thèse de doctorat, soutenue le 5 juillet 2005 à l'Observatoire de Haute Provence, sur le thème de la *reconfiguration de pupille pour la haute résolution angulaire*. A peu près treize ans se seront donc écoulés au moment de la soutenance de cette thèse écrite dans le but d'obtenir une habilitation à diriger des recherches.



Figure 1: Le magnifique campus de l'Université de Cornell, dans la ville d'Ithaca, dans l'état de New York.

Ces treize années auront été passées en trois lieux différents : d'abord l'université de Cornell où j'ai, en collaboration avec le Professeur James Lloyd, entamé un programme observationnel exploitant la méthode de l'interférométrie à masque non-redondant qui transforme un télescope imageur classique en interféromètre... une véritable hérésie de ma part à l'époque, étant donné qu'une grande partie de ma thèse avait été dédiée au concept d'hypertélescope, dont la raison d'être est précisément d'accomplir la mission inverse ! Ce projet m'a pourtant mené à faire de nombreuses observations, principalement depuis l'Observatoire du Mont Palomar, avec le vénérable télescope de Hale qui, en combinant l'optique adaptative avec ce mode d'observation étrange dans l'infrarouge, s'est retrouvé à rivaliser avec le télescope spatial de Hubble en résolution angulaire.

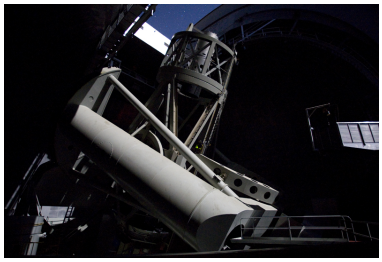


Figure 2: Le vénérable télescope de Hale, à l'observatoire du Mont Palomar.

Après trois ans à Cornell, je suis reparti travailler au télescope Subaru du National Astronomical Observatory of Japan (NAOJ) avec le Professeur Olivier Guyon. Le concept d'apodisation par reconfiguration de pupille appelé PIAA auquel j'avais contribué pendant ma thèse de doctorat s'est, durant les cinq ans passés au Subaru, progressivement transformé en instrument d'imagerie haut contraste, appelé SCExAO (Subaru Coronagraphic Extreme Adaptive Optics), qui est maintenant ouvert aux observations scientifiques. En complément de la conception et de l'intégration de l'instrument au télescope, j'ai profité de SCExAO pour développer des méthodes de contrôle de front d'onde plan focal de type speckle nulling.



Figure 3: Le télescope Subaru, à 4200 mètres d'altitude au sommet du Mauna Kea.

Alors que je me préparais à m'installer de façon permanente au Subaru pour continuer mon travail de développement et d'exploitation scientifique de SCExAO, j'ai été recruté en tant que maître de conférence à l'Observatoire de la Côte d'Azur que j'ai intégré à

l'automne 2013. Me voici donc aujourd'hui, présentant à l'OCA ma demande d'habilitation à diriger des recherches pour continuer mon travail à la frontière entre le développement et l'exploitation d'instrumentation en haute résolution angulaire. Etant donné que mon travail s'est fait dans trois endroits différents, il n'est pas étonnant de voir trois grands thèmes d'activités se dégager. Je vais donc organiser ce document en trois parties :



FIGURE 4: Entrée dans la Grande Coupole de l'Observatoire de la Côte d'Azur, emblématique de la ville de Nice.

— La première sera consacrée à l'illustration de résultats d'observation obtenus au Mont Palomar grâce à la technique de l'interférométrie à masque non-redondant utilisée avec optique adaptative. Nous verrons les raisons qui ont fait qu'une technique d'observation aussi exotique a progressivement réussi à coloniser tous les grands télescopes équipés d'optique adaptative qui se sont mis à utiliser des masques non-redondants... un succès qui culmine d'une certaine façon avec une place à bord du James Webb Space Telescope (JWST).

— La seconde partie sera consacrée au monde de l'optique adaptative extrême et de l'imagerie haut contraste, telle que j'ai contribué à la mettre en oeuvre dans le cadre du Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) instrument, maintenant ouvert aux observations scientifiques. Dans ce contexte, plutôt que de passivement étalonner les observables, nous irons activement moduler l'amplitude complexe des speckles dans des images pour créer, malgré un environnement instable, des régions à haut contraste, en particulier dans le régime des très faibles séparations angulaires.

— La troisième partie sera consacrée à une tentative réussie de conciliation des techniques d'étalonnage des observables de l'interférométrie à masque non-redondant avec les méthodes plus classiques d'observation, ie. utilisant la pleine pupille d'un télescope. Cette réconciliation a été permise par la mise au point d'une généralisation de la notion de clôture de phase que j'ai baptisée kernel (ou noyau) de phase et qui permet de s'affranchir de la stricte condition de non-redondance de la pupille.

CETTE SYNTHÈSE est particulièrement féconde en applications, allant du post-processing d'images d'archive à la définition de nouvelles méthodes d'analyse de surface d'onde. Cette idée de kernel et les applications qu'elle ouvre m'ont permis de décrocher un contrat ERC qui a démarré au mois d'octobre 2016 et se terminera en 2023. Mais avant d'en arriver là, il faut prendre cette histoire depuis le début et commencer par comprendre comment des astronomes en sont venus à vouloir masquer jusqu'à 90 % de la surface collectrice de grands télescopes, en réfléchissant au processus de formation des images en astronomie et à comment on peut repousser les limites de la diffraction.

A PROCESS CANNOT BE UNDERSTOOD BY STOPPING IT. UNDERSTANDING MUST MOVE WITH THE FLOW OF THE PROCESS, MUST JOIN IT AND FLOW WITH IT.

THE FIRST LAW OF MENTAT, FRANK HERBERT, *DUNE*

Interférométrie à masque non redondant

L'image au coeur de l'astronomie

L'image est à l'origine de la majorité des investigations scientifiques en astronomie. Extension directe de notre perception visuelle tellement intuitive, l'image est aujourd'hui la donnée de base à partir de laquelle un observateur va successivement pouvoir (1) identifier de nouvelles sources, (2) mesurer leur position et leur brillance par rapport à des sources dites de référence et (3) en suivre l'évolution en fonction du temps, de la longueur d'onde et de la polarisation. C'est à partir de ces mesures fondamentales qui forment un échantillon d'un espace des paramètres multi-dimensionnel (coordonnées, longueur d'onde, polarisation et temps), qu'un astronome va pouvoir construire un modèle amélioré qui raconte l'histoire d'un amas d'étoiles, d'un groupe de galaxies, de ce qui se trame au centre de notre Galaxie, ou d'un système planétaire en formation autour d'une étoile jeune du voisinage Solaire. La bonne interprétation des images est donc d'une importance capitale à toute forme de recherche en astronomie : les applications décrites dans ce mémoire sont destinées à aider à cette interprétation.



Figure 5: Exemple d'image astronomique particulièrement riche en information : le trapèze de la grande nébuleuse d'Orion, prise par l'instrument NICMOS équipant le télescope spatial de Hubble.

Que ce soit dans ce qu'on appelle véritablement un instrument imageur, un spectrographe ou alors un interféromètre, l'image est avant tout un lieu optique particulier, dans lequel les photons en provenance de multiples sources collectés par une surface optique où ils sont tous presque uniformément distribués, se retrouvent naturellement triés. Il est possible de décrire formellement le résultat de ce processus de ségrégation des photons, comme étant le résultat **I** du produit de convolution de deux fonctions : une qui est représentative de la distribution d'intensité qui décrit la source, ce qu'on appelle communément l'objet, et qu'on notera **O** ; l'autre qui décrit la réponse instrumentale, incluant l'atmosphère, le télescope et toutes les optiques rencontrées par la lumière avant d'arriver au détecteur. Cette réponse instrumentale, appelée fonction d'étalement du point (point spread function en anglais) est notée **PSF**, de telle sorte que :

$$\mathbf{I} = \mathbf{O} \otimes \mathbf{PSF}. \tag{1}$$

où $\otimes$ représente l'opération de convolution.



FIGURE 6: Image de Proxima Centauri prise par le Hubble Space Telescope. Le halo entourant l'étoile et les très évidentes aigrettes de diffraction présentes dans cette image nous rappellent que la contribution du télescope et de l'instrument (**PSF**) à l'image (**I**) ne peut pas toujours être négligée.

UNE GRANDE ÉNERGIE est dépensée dans la construction de grands télescopes, le design astucieux d'optiques de transfert et enfin, de l'instrumentation, maintenant *boostés* par l'utilisation d'optique adaptative (OA), dans le but de réduire la contribution de la fonction instrumentale et d'optimiser le processus de séparation des photons qui conduit à une meilleure caractérisation des sources observées. Pour une grande partie des observations en astrophysique, cette amélioration est satisfaisante à tel point qu'on identifie l'objet **O** directement à l'image **I** sans vraiment se préoccuper de la **PSF**. On pousse quelquefois les choses à leur paroxysme, par exemple sur les télescopes imageurs grand champ spatiaux comme Euclid pour lequel on ne cherche même pas à respecter le critère d'échantillonage minimal qui permettrait de caractériser cette **PSF**.

LA SÉGRÉGATION SPATIALE OPTIMALE DES PHOTONS dans le plan image d'un télescope est cependant fondamentalement limitée par le phénomène de diffraction (voir Fig. 6). L'ordre de grandeur qui règle cette limite est le rapport entre la longueur d'onde d'observation λ, et la dimension caractéristique (diamètre d'ouverture ou longueur de base interférométrique) de l'optique utilisée pour collecter la lumière. Pour rapidement estimer la résolution angulaire offerte par un télescope, il est utile de garder la formule suivante sous la main [1] :

$$\theta \approx 200 \times \frac{\lambda}{D}, \tag{2}$$

où θ est la résolution angulaire en millisecondes d'arc, λ la longueur d'onde en microns et D le diamètre du télescope en mètres. Une petite application numérique permet de vérifier qu'un télescope de 1 mètre comme C2PU observant dans le visible (λ = 0.5 μm) offre une résolution de 100 mas (0.1 seconde d'arc) et qu'un télescope de 8 mètres de diamètre comme un des quatre VLTs observant dans l'infra-rouge proche (λ = 1.6 μm) offre une résolution de 40 mas.

Même dans des conditions d'observation idéales, la ségrégation spatiale des photons par l'optique géométrique ne suffit pas à résoudre certains types de problèmes importants, tel que celui de l'iden-

1. Le facteur 200 dans la formule provient de la conversion de radians en millisecondes d'arc : 180 * 3600 / π / 1000 = 206.265 $\approx$ 200.

tification de sources faibles dans le voisinage direct d'un objet brillant, ou celui de la discrimination de sources si proches les unes des autres, qu'on les qualifie de non-résolues. Le plus en vogue de ces scénarios concerne le cas de l'imagerie haut contraste des planètes extrasolaires. Pour mettre les échelles de résolution angulaire précisées un peu plus haut en perspective dans ce contexte, on peut rappeler qu'une planète orbitant à une unité astronomique (UA) d'une étoile située dans le voisinage Solaire à 10 parsecs (pc) aura comme séparation angulaire maximale 0.1 seconde d'arc, ce qui est très proche du pouvoir de résolution limite des cas présentés.

La technique de l'interférométrie optique a pourtant depuis longtemps appris à composer avec cette limite de la diffraction et produit de façon routinière des résultats observationnels qui défient en apparence cette limite et ce, malgré (en comparaison d'un système produisant une image) un nombre très limité de quantités observables. Ceci est rendu possible par la bonne compréhension des sources d'incertitude et de biais affectant la mesure fondamentale de cohérence, qui a conduit à la formulation de quantités observables bien étalonnées. L'exemple le plus connu de ces quantités observables est la clôture de phase, qui permet, en post-processing de s'affranchir des erreurs induites par les défauts optiques et atmosphériques. La possibilité d'extraire et d'exploiter ces observables de qualité d'interférogrammes a justifié la mise en place d'une série de programmes d'observation utilisant notamment le télescope de Hale de l'Observatoire du Mont Palomar. Le succès de l'utilisation de cette technique depuis les télescopes de Keck et Palomar a conduit à une prolifération de ces masques sur l'essentiel des télescopes équipés d'optique adaptative qui ont trouvé là le moyen d'augmenter leur pouvoir de résolution pour un investissement matériel très minime.

Mesurer la cohérence des sources grâce à l'interférométrie

Dans sa forme la plus fondamentale[2], l'objet de l'interférométrie optique stellaire est la mesure de la cohérence mutuelle du champ électrique émis par une source, simultanément prélevé en deux stations d'observation formant ce que l'on appelle une ligne de base interférométrique. A partir d'une collection de mesures de cohérence pour plusieurs lignes de base, on peut remonter à la carte de luminosité de la source, grâce à un résultat connu sous le nom de théorème de Van Cittert - Zernike, qui est l'équivalent en interférométrie de la relation de convolution objet - image décrite par l'équation 1.

2. Le théorème de Van Cittert Zernike identifie le degré de cohérence mutuelle du champ électrique émis par une source astrophysique et mesuré par deux stations d'observations formant une ligne de base interférométrique de coordonnées (u, v) à la transformée de Fourier de la carte d'intensité normalisée décrivant la source.

$$\gamma(u,v) = \int I(\alpha,\beta) e^{-i2\pi(\alpha u+\beta v)}\, \mathrm{d}\alpha \mathrm{d}\beta$$

Cette mesure de cohérence, souvent notée γ, est aussi appelée visibilité complexe : c'est, comme son nom l'indique, un nombre complexe, séparable en deux composantes qui sont respectivement le module de visibilité $|\gamma|$ et sa phase Φ.

Le module de visibilité mesuré par une ligne de base interféro-

métrique est lié à la taille angulaire caractéristique de la source observée dans la direction formée par la paire d'ouvertures. Si ce module (une fois étalonné) vaut l'unité, l'objet observé peut être considéré comme un point source. Si ce module prend une valeur intermédiaire entre 0 et 1, l'objet observé est résolu : la mesure du module peut servir à en contraindre le diamètre angulaire, si l'on introduit un modèle *a priori* de l'objet, comme un modèle de disque uniforme, où un profil d'assombrissement centre-bord plus sophitsiqué.

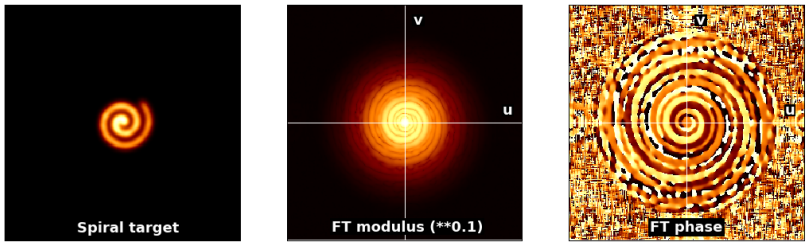


FIGURE 7: Exemple d'objet à géométrie spirale et la visibilité complexe qui lui est associée, séparée en module et phase pour tout le plan (u,v). Un disque uniforme de diamètre équivalent présenterait également une signature en amplitude, mais pas en phase.

LA PHASE EST RELIÉE AUX PROPRIÉTÉS DE SYMÉTRIE de l'objet observé. Pour une source de géométrie centro-symétrique observée dans des conditions idéales (un détail sur lequel nous allons bien entendu devoir revenir) la phase mesurée par une ligne de base interférométrique ne peut prendre que deux valeurs : 0 ou 180 degrés. Une valeur intermédiaire de la phase révèle que l'objet présente une asymmétrie, qui peut soit être induite par la présence d'un compagnon de plus faible luminosité, une tâche sombre ou une surbrillance sur la photosphère.

ETANT DONNÉ LA MOTIVATION SCIENTIFIQUE générale de détection de compagnons faibles autour des étoiles du voisinage Solaire et de la caractérisation des environnements proches de ces étoiles, c'est essentiellement cette information de phase que l'on va tenter d'extraire de nos observations interérométriques : c'est donc sur la phase que je vais concentrer ma description.

La phase interférométrique

LA PHASE MESURE L'ASYMÉTRIE d'une source. La figure 8 rappelle, de façon schématique, le principe d'acquisition de franges d'interférences avec un masque interférométrique à deux trous observant dans des conditions idéales une source non-résolue. La figure de diffraction classique d'Airy de largeur $2.44\lambda/D$ qui serait obtenue si on retirait le masque interférométrique est remplacée par un système de franges d'interférences de période λ/B, où B est la distance entre les trous du masque, et s'étalant dans une enveloppe de rayon λ/h, où h est le diamètre d'un trou.

Ces franges sont parfaitement contrastées : la source est en effet non-résolue par l'interféromètre et la cohérence entre les deux champs électriques prélevés par les deux ouvertures est totale : $|\gamma| = 1$. Ces franges sont également parfaitement symmétriques et centrées sur le détecteur : la source est en effet observée sur l'axe.

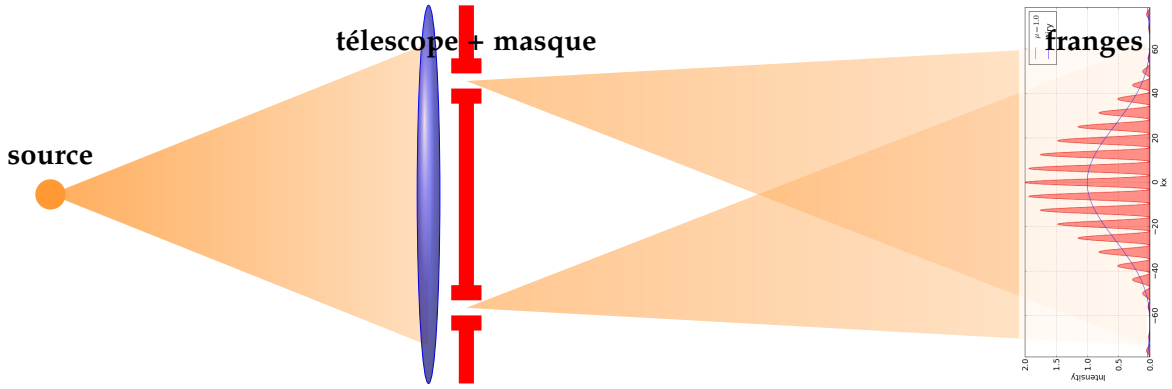


FIGURE 8: Acquisition de franges de Fizeau sur une source non-résolue par l'interféromètre : les franges résultantes sont parfaitement contrastées.

UNE SOURCE SUPPLÉMENTAIRE, non-confondue avec la première, produit également son propre système de franges d'interférences. La source étant située hors-axe, ces franges sont décalées par rapport aux précédentes. Cette lumière supplémentaire s'additionne de façon incohérente avec celle déjà présente dans le plan focal : les interfranges originalement parfaitement sombres du cas précédent (voir Figure 8) sont maintenant en partie comblées par la lumière de cette nouvelle source. Les franges, moins contrastées, sont caractérisées par un module de visibilité de valeur intermédiaire : $0 < |\gamma| < 1$. La figure 9 présente ce scénario.

En plus de la perte de visibilité, le centre de gravité du système de franges se déplace dans la direction de la position du compagnon, d'une mesure qui dépend du rapport de luminosité entre les deux objets : à ce déplacement du centre de gravité du système de franges, on associe une phase, décrivant (selon l'orientation du système de coordonnées retenu dans l'image) un "retard" ou une "avance" des franges enregistrées par le détecteur.

SI ON NÉGLIGE L'ENVELOPPE qui module les franges, ces dernières

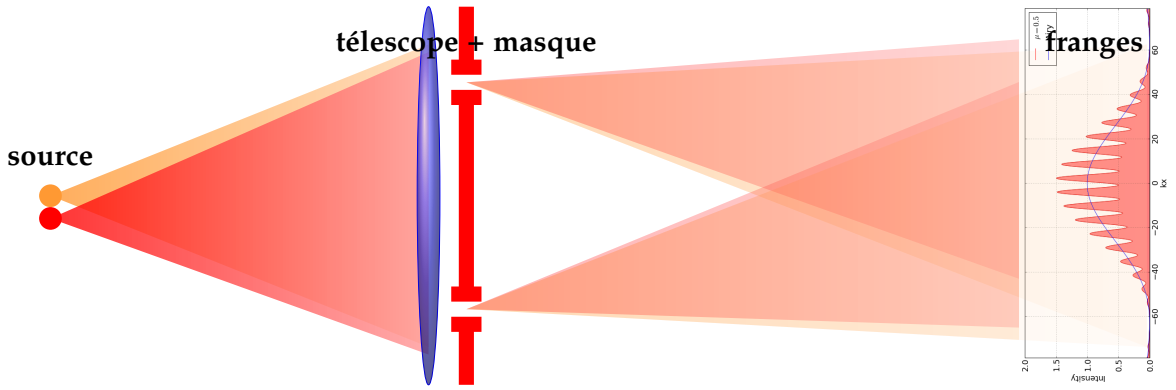


FIGURE 9: Acquisition de franges de Fizeau sur une source binaire partiellement résolue : les franges résultantes ont une visibilité intermédiaire et une phase non-nulle.

peuvent en effet être décrites par une équation simple qui inclut les informations composant la visibilité complexe :

$$I(x) = I_0 * \Big(1 + \mu \cos\big(kx - \phi\big)\Big), \tag{3}$$

avec k, le nombre des franges, imposée par le rapport λ/B où B est la distance entre les trous du masque, μ, le module de visibilité et ϕ la phase des franges.

ON DEVINE ASSEZ VITE que cette information de phase peut être très facilement corrompue, par exemple par une faible erreur de pointage du télescope, qui va résulter en un décalage du système de franges, absolument indiscernable de celui induit par la présence d'un compagnon. La turbulence atmosphérique, se traduisant par des fluctuations localse de l'indice de réfraction de l'atmosphère, induit un retard stochastique avec, aux longueurs d'onde du visible, une échelle de temps caractéristique de l'ordre de la milli-seconde. La figure 10 illustre cet effet. L'équation 3 doit être modifiée pour réfléter cette réalité :

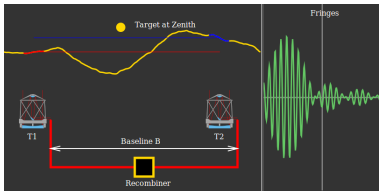


FIGURE 10: Illustration de l'effet du piston induit par l'atmosphère : l'information de phase, déduite de la position du packet de franges (sur la droite) sur le détecteur est décalée à cause du chemin optique supplémentaire introduit par les fluctuaitons de l'indice optique atmosphérique. Voir : http://frantzmartinache.eu/static/04_teaching/interfero_activity_018.html pour une démonstration dynamique de cet effet du piston atmosphérique.

$$I(x,t) = I_0 * \Big(1 + \mu \cos\big(kx - \phi + \Delta\varphi(t)\big)\Big), \tag{4}$$

et inclure un terme global de piston instrumental $\Delta\varphi(t)$, potentiellement variable dans le temps, qui est inséparable de la phase Φ de la source observée. L'information de phase telle qu'elle a été introduite semble tout simplement inutilisable. Pour la recouvrer, il faut avoir recours à une astuce, qui requiert un masque interférométrique à au moins trois ouvertures qui va permettre la construction d'une nouvelle grandeur : la clôture de phase.

La clôture de phase

UNE TROISIÈME OUVERTURE permet en effet de récupérer une fraction de l'information de phase. La figure 11 illustre l'apport d'une troisième ouverture. Chacune des trois lignes de base interférométrique : B_{12}, B_{23} et B_{32} donne accès à une mesure de visibilité com-

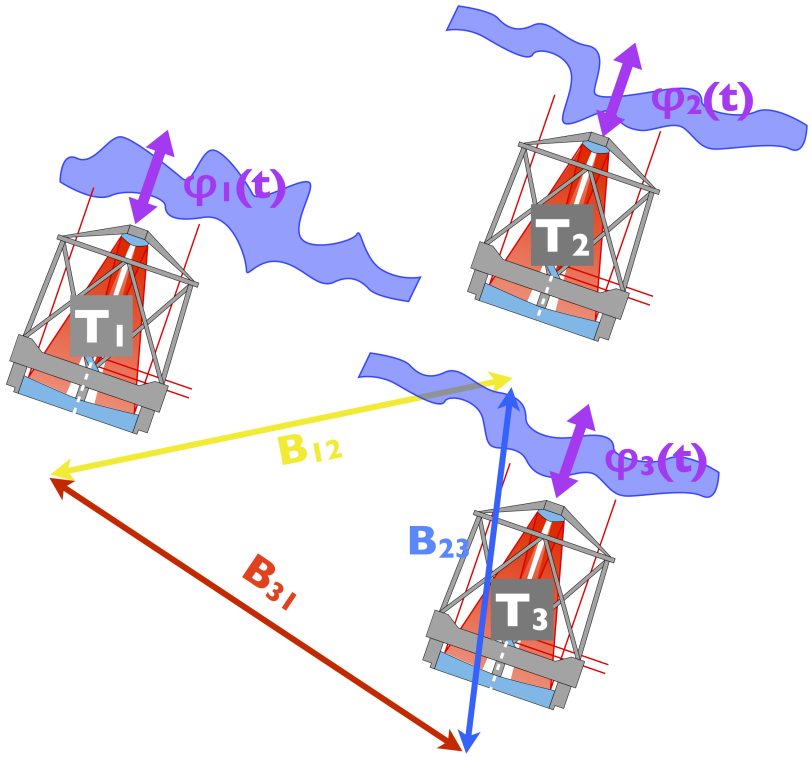


FIGURE 11: Trois ouvertures (ou des télescopes distincts comme en interférométrie longue base) recombinées interférométriquement faisant toutes l'expérience de piston atmosphérique.

plexe, comprenant notamment les phases : $\Phi(1-2)$, $\Phi(2-3)$ et $\Phi(3-1)$. Chacune de ces phases est bien affectée par le problème qui vient d'être mentionné : chaque phase mesurée peut s'écrire comme la somme de la phase de la source astrophysique $\Phi_O(i-j)$ (l'information à laquelle on souhaite avoir accès) et de la phase instrumentale $\Delta\varphi_{ij}$ le long de la base correspondante (i, j). C'est le retard de l'onde électromagnétique prélevée par les deux ouvertures i et j qui compose ce "piston" atmosphérique : $\Delta\varphi_{ij} = \varphi_j - \varphi_i$. On obtient le système d'équations suivant :

$$\begin{aligned}
\Phi(1-2) &= \Phi_O(1-2) + (\varphi_1 - \varphi_2) \\
\Phi(2-3) &= \Phi_O(2-3) + (\varphi_2 - \varphi_3) \\
\Phi(3-1) &= \Phi_O(3-1) + (\varphi_3 - \varphi_1).
\end{aligned}$$

CES TROIS PHASES INSTRUMENTALES ne sont pas indépendantes les unes des autres. Il suffit d'additionner ces trois équations pour s'apercevoir que les termes contribuant au piston s'annulent un à un : il est donc possible à partir de trois phases "polluées" par des effets instrumentaux de construire une nouvelle quantité observable, appelée clôture de phase[3] qui est indépendante de ces effets instrumentaux.

C'EST CETTE QUANTITÉ OBSERVABLE qui donne tout son intérêt à l'interférométrie à masque non-redondant : la possibilité de s'affranchir des effets atmosphériques ou de chemin optique non corrigés

3. R. C. Jennison. A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent. **MNRAS**, 118:276–+, 1958

par de l'optique adaptative située en amont (voir la section suivante) a permis à cette technique d'accomplir de véritables exploits observationnels, défiant les limites conventionnellement acceptées de résolution angulaire.

Interférométrie à masque non-redondant

Même s'il offre une grande robustesse à l'encontre des effets instrumentaux, un masque percé de trois petits trous, placé dans un plan conjugué avec la pupille du télescope rend les observations avec ce dernier particulièrement inefficaces. Quatre quantités observables (trois modules de visibilité + une clôture de phase) sont une maigre moisson en comparaison avec l'information qu'il est possible de collecter sur un détecteur typiquement constitué d'un milliard de pixels ! Heureusement, il est possible utiliser des masques plus riches, offrant simultanément accès à un plus grand nombre de lignes de bases interférométriques et de clôtures de phase.

LE MASQUE QUE J'AI LE PLUS UTILISÉ est un masque à neuf trous, tel que celui présenté dans la figure 12. L'ajout de trous augmente le nombre de bases interférométriques et permet de former plus de triangles auxquels on va pouvoir associer des clôtures. Un masque à N trous permet de former $N_B = \frac{N\times(N-1)}{2}$ lignes de base et $N_C = \frac{(N-1)\times(N-2)}{2}$ triangles indépendants[4]. Le masque à neuf trous permet donc la formation de 36 lignes de base et l'obtention de 28 clôtures de phase indépendantes.

4. J. D. Monnier. An Introduction to Closure Phases. In P. R. Lawson, editor, *Principles of Long Baseline Stellar Interferometry*, pages 203–+, 2000

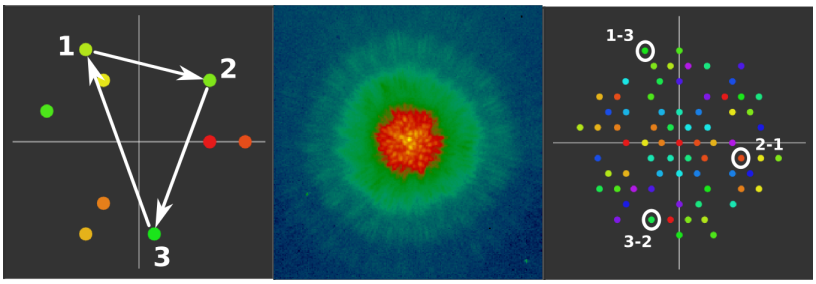


FIGURE 12: Exemple de masque interférométrique non-redondant à neuf trous. De gauche à droite : le masque, avec trois ouvertures (et trois bases) mises en évidence ; l'interférogramme produit par un tel masque en utilisant un filtre H ; la couverture (u, v) correspondante au masque, avec trois points mis en évidence, qui correspondent aux lignes de base du panel de gauche.

POURQUOI NE PAS UTILISER UN MASQUE RICHE comprenant un plus grand nombre de trous ? C'est cette question qui a motivé le développement de l'idée de kernel de phase, dont je parlerai dans la troisième partie de ce mémoire. Pour le moment, il me faut exposer l'argumentation classique qui justifie l'utilisation de masques comme celui présenté dans la figure 12. La géométrie qui guide la position des trous n'est en effet pas anodine et est guidée par une contrainte de non-redondance.

Un réquisitoire contre la redondance

Le théorème de Van Cittert Zernike dans la forme qui a été rappelée plus haut nous dit que pour une source astrophysique donnée, la mesure de la cohérence ne dépend pas de la position absolue des ouvertures contribuant à l'interférence mais seulement de leur position relative, explicitée par les traditionnelles coordonnées (u,v) : cela veut dire qu'une autre paire de télescopes identique à la première, simplement translatée n'apporte rien de nouveau et mesure la même visibilité complexe. Un réseau interférométrique qui mesure la même information plusieurs fois est qualifié de **redondant**. Un masque interférométrique tel que celui à neuf trous introduit plus haut est délibérément conçu de façon à être **non-redondant** : chaque paire de trous forme une ligne de base interférométrique qui ne se retrouve qu'une seule fois dans le réseau.

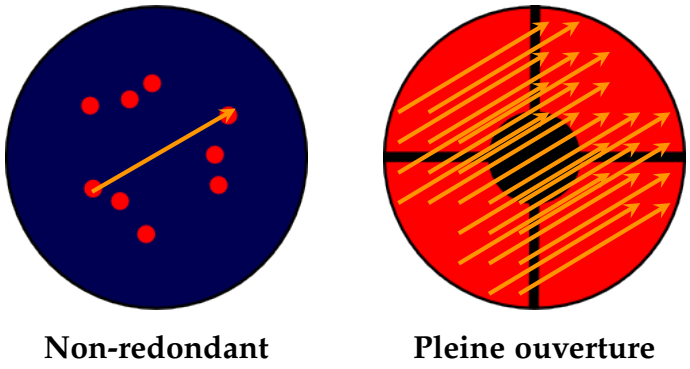


FIGURE 13: Propriété du masque non-redondant : la géométrie guidant la position des trous fait qu'une base interférométrique n'est jamais répétée. En l'absence de masque, cette même base interférométrique est échantillonnée un très grand nombre de fois.

LORSQU'ON COMPARE un masque interférométrique à neuf trous à la pleine ouverture d'un télescope (voir la figure 13), on voit qu'une ligne de base unique dans le cas non-redondant se trouve échantillonnée un très grand nombre de fois lorsque l'ouverture est pleine : en l'absence de masque, on pourra parler de **base interférométrique virtuelle** ou de **fréquence spatiale**. La présence d'une obstruction centrale et de l'ombre de supports portant le miroir secondaire du télescope peuvent altérer les détails de cette observation, mais la tendance est que plus la base virtuelle considérée est courte et plus elle est présente dans la pleine ouverture. Le problème associé à cette redondance est qu'elle est à l'origine d'une confusion supplémentaire qui conduit à une perte de l'information de phase et du module de visibilité.

Les conséquences de la redondance

Le mode de recombinaison interférométrique qui a été utilisé jusqu'ici est un mode dit **de Fizeau**, car il réplique le mode de recombinaison historique mis au point en 1870 par ce dernier. Dans ce mode, les interférences produites par toutes les bases interférométriques sont acquises dans un unique plan focal, produisant un in-

terférogramme tel que celui montré dans le panel central de la figure 12.

DANS LE CAS NON-REDONDANT, le traitement de cet interférogramme est simple : une transformée de Fourier de l'image révèle que l'information échantillonnée par le masque est directement accessible (voir la figure 14). Avec la pupille redondante, l'information est plus difficile à interpréter : les bases interférométriques virtuelles étant majoritairement redondantes, à une coordonnée (u,v) on peut associer la contribution de centaines de bases virtuelles.

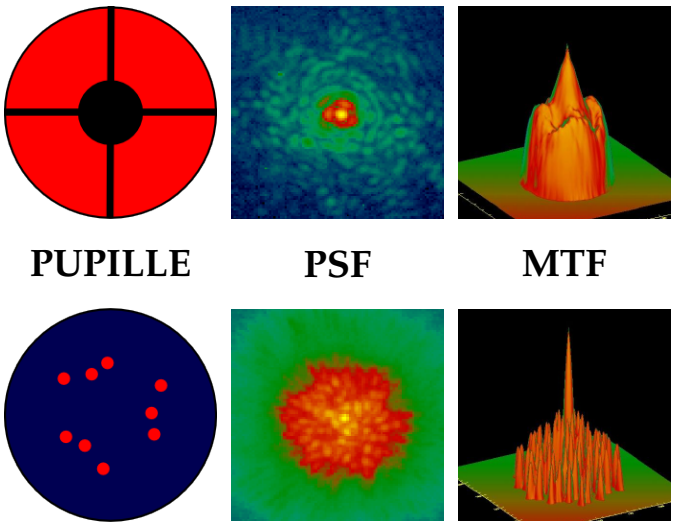


FIGURE 14: Comparaison par étape de l'imagerie classique "pleine ouverture" (rangée du haut) et de l'imagerie avec masque interférométrique non-redondant (rangée du bas). De gauche à droite, sont systématiquement présentés : (à gauche) la pupille utilisée, (au centre) un exemple d'image produite par cette pupille et (à droite), une vue 3D de la MTF (fonction de transfert en modulation, équivalente à un module de visibilité dans le contexte de l'interférométrie). A l'examen de cette MTF, on remarque, dans le cas "pleine ouverture" qu'au delà d'une certaine fréquence spatiale (dépendant des propriétés du miroir d'optique adaptative utilisé pour les observations), cette MTF présente des structures inattendues, entièrement attribuables aux résidus de correction adaptative. Dans le cas non-redondant, cette MTF est en comparaison, particulièrement propre. C'est cette propriété des masques non-redondants qui en fait l'intérêt principal.

Cette situation ne serait pas problématique en l'absence des perturbations instrumentales ou atmosphériques mentionnées plus haut, et qui sont responsables d'une confusion de la phase échantillonnée par chaque base virtuelle. Les amplitudes complexes échantillonnées par une base virtuelle redondante s'additionnent de façon cohérente : à chacune des bases mises en évidence en jaune sur le panel droit de la figure 13 est associé un phaseur, c'est à dire un vecteur dans le plan complexe dont le module est[5] le module de visibilité de la source astrophysique $|\gamma_0|$ et une phase aléatoire Φ affectée par les effets instrumentaux.

5. en l'absence de scintillation seulement

DANS LE MODE DE RECOMBINAISON FIZEAU, les contributions de toutes les bases virtuelles identiques sont regroupées : la visibilité complexe associée, enregistrée aux coordonnées (u,v) est le résultat de la somme des phaseurs associés à chaque base virtuelle contribuant à ce point (u,v). La figure 15 illustre ce cas de figure pour une base quatre fois redondante. C'est l'exemple classique du problème de marche aléatoire : on voit que non seulement, la phase de cette somme de phaseurs est aléatoire mais aussi que l'amplitude le devient également.

LA DESTRUCTION DE L'INFORMATION DE COHÉRENCE potentiellement collectable par un télescope observant de façon classique[6] est

6. c'est à dire, non masquée

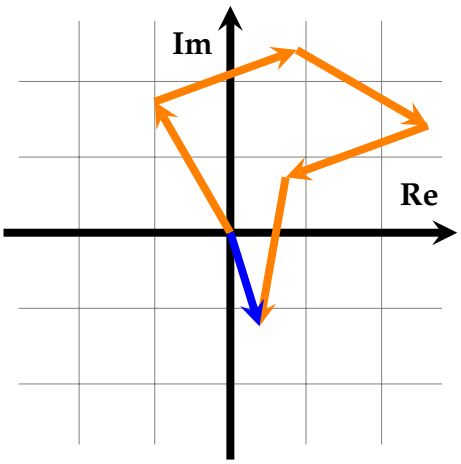


FIGURE 15: Somme de cinq phaseurs aléatoires dans le plan complexe : le résultat de cette somme est lui même aléatoire. Dans un tel cas, non seulement la phase mais également l'amplitude, associée au module de visibilité sont perdus.

le résultat d'un processus en deux étapes :

1. la turbulence atmosphérique et la phase instrumentale détruisent l'information de phase en introduisant un terme de piston aléatoire.
2. la redondance des bases interférométriques virtuelles présentes dans la pupille du télescope achève cette corruption de l'information en détruisant l'information du module de visibilité.

CETTE AMPLITUDE COMPLEXE ALÉATOIRE dans le plan focal se manifeste par la présence de tavelures ou speckles dans l'image. En combinant de nombreuses images speckles traitées dans l'espace de Fourier, il est possible de regagner accès au module de visibilité, tel que mis en oeuvre dans l'interférométrie des speckles originale [7]. Par comparaison, un masque non-redondant placé dans la pupille d'un télescope permet cependant à partir d'une seule image, de sauver une grande fraction de l'information incluant la phase. Le masque à neuf trous présenté plus haut donne instantanément accès à 36 modules de visibilité et 28 clôtures de phase, ce qui comparé à l'interférométrie optique longue base [8] est un mode d'observation finalement très efficace. La clôture de phase présente des caractéristiques qui la rendent très attractive : la possibilité de s'affranchir des effets atmosphériques lorsqu'on observe depuis le sol et d'éventuels défauts optiques lorsqu'on observe depuis l'espace ont rendu cette technique d'interférométrie à masque non-redondant particulièrement populaire. Des masques non-redondants sont en effet aujourd'hui disponibles dans la grande majorité des instruments d'imagerie haute résolution. Un masque non-redondant (voir la figure 16) fait même désormais partie des fonctionnalités affichées de l'instrument NIRISS qui ira dans l'espace à bord du JWST [9].

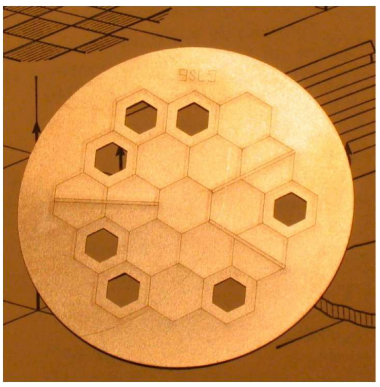


FIGURE 16: Prototype de masque non-redondant à sept trous qui fera partie des modes d'observation de l'instrument NIRISS, à bord du JWST.

7. A. Labeyrie. Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier Analysing Speckle Patterns in Star Images. **A&A**, 6:85, May 1970

8. Les instruments combinant simultanément quatre télescopes au VLTI sont en cours d'installation

9. A. Sivaramakrishnan, D. Lafrenière, K. E. S. Ford, B. McKernan, A. Cheetham, A. Z. Greenbaum, P. G. Tuthill, J. P. Lloyd, M. J. Ireland, R. Doyon, M. Beaulieu, A. Martel, A. Koekemoer, F. Martinache, and P. Teuben. Non-redundant Aperture Masking Interferometry (AMI) and segment phasing with JWST-NIRISS. In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442 of **Proc. SPIE**, page 84422S, Sept. 2012

Acquisition et analyse des données

La méthode d'acquisition de données interférométriques ressemble à celle des images conventionnelles acquises avec optique adaptative

dans un petit champ (de l'ordre de la seconde d'arc). Prenant en compte la faible transmission du masque, le temps d'intégration doit être optimisé de façon à ce que les franges de l'interférogramme (voir panneau central de la figure 14) soient correctement échantillonnées (trade-off avec le bruit de lecture du détecteur) sans aucune saturation. Pour que l'image puisse être interprétée dans l'espace de Fourier, il est en effet essentiel que la relation de convolution de l'équation 1 soit strictement vérifiée. Cela exclut donc l'utilisation de masques coronagraphiques. On privilégie autant que possible un temps d'intégration court, de façon à se placer le plus près possible du régime "frozen phase-screen" pour lequel la variance temporelle du front d'onde est minimale. Comme pour des acquisitions non-interférométriques, les interférogrammes sont ensuites corrigés du "dark" et du "flat field"

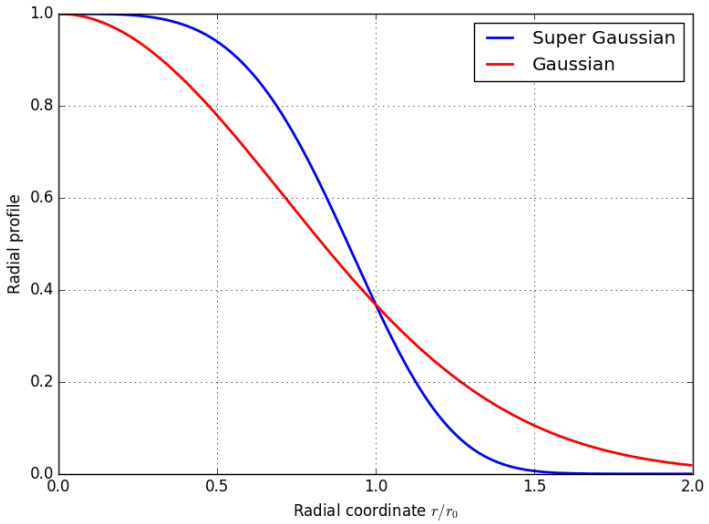


FIGURE 17: Profil du filtre super-Gaussien (exponentielle d'ordre 4), comparé à une fonction Gaussienne (d'ordre 2). Ce filtre est utilisé dans la réduction de données NRM.

LE RAPPORT ENTRE LA PLUS COURTE ET LA PLUS LONGUE DES BASES INTERFÉROMÉTRIQUES couvertes par le masque impose la taille du champ utile de ce masque : si la ligne de base la plus longue est en général comparable à un diamètre de la pupille original et donnant donc une résolution angulaire λ/D, la plus petite est pour les masques utilisés, typiquement quatre fois plus courte (voir par exemple le prototype du masque de NIRISS présenté en figure 16), ce qui va correspondre à un champ utile de taille $4\lambda/D$. Avant d'être interprété dans le domaine de Fourier, chaque interférogramme est recentré et multiplié par une fonction apodisante, de rayon caractéristique r_0 ajusté au champ utile suivant une loi :

$$f(r) = \exp\left(-(r/r_0)^4\right), \tag{5}$$

appelée super-Gaussienne (voir figure 17) et dont le but est de filtrer les parties de l'image contenant peu de signal mais pouvant introduire des biais dans la transformée de Fourier.

CES DONNÉES ACQUISES SUR LA CIBLE D'INTÉRÊT doivent être étalonnées par l'observation d'une ou préférentiellement de plusieurs étoiles non-résolues, n'appartenant pas à des systèmes multiples, de type spectral et de luminosité comparable à la cible d'intérêt (pour ne pas changer le régime de fonctionnement de l'optique adaptative) et situées dans une région du ciel suffisamment proche, en particulier en élévation (air-mass comparable) qui limite l'apparition d'effets systématiques non-contrôlés. L'acquisition de clôtures de phase sur plusieurs sources de calibration permet de construire un jeu de clôtures étalon, qui sont ensuite soustraites des mesures enregistrées sur l'étoile cible, comme on le fait en interférométrie longue base. Ces clôtures de phase étalonnées (voir l'exemple présenté dans la Figure 18) sont ensuite à interpréter.

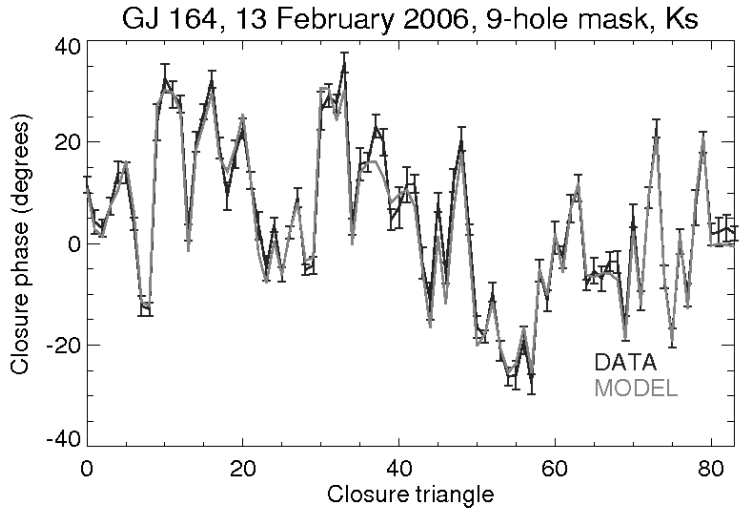


FIGURE 18: Exemple de clôtures de phase calibrées, acquises sur l'étoile GJ 164 dans un filtre K_S (autour de 2 μm) avec le masque à neuf trous installé dans une roue de l'instrument PHARO sur le télescope de Hale du Mont Palomar.

LES CLÔTURES DE PHASE SONT DES QUANTITÉS ABSTRAITES, qui encodent le degré d'asymétrie de la cible observée dans une direction dépendant du triangle de clôture considéré (voir la figure 12). Un objet parfaitement symétrique comme une étoile unique non-résolue ou une binaire de contraste de luminosité 1 : 1 présentera des clôtures de valeur alternant entre 0° et 180° mais pas de valeurs intermédiaires. L'exemple de la figure 18 montre que GJ 164 dévie significativement de ce genre d'objet. Pour aller au delà de ce simple constat, il va falloir tester des hypothèses de complexité croissante sur la nature de l'objet : compagnon, système triple, étoile résolue à tâches, rotateur rapide... et tester ces hypothèses.

NOTRE CONNAISSANCE A PRIORI DES OBJETS NOUS GUIDE : dans le cas de GJ 164, l'étoile principale est une naine M, située à environ 14 pc et qui n'est donc pas résolue par l'ouverture de 5 m du télescope de Hale, ce qui laisse les modèles multiples comme de plausibles interprétations des clôtures de phase. Le modèle binaire, le plus simple (trois paramètres : séparation angulaire ρ, angle de

position θ et contraste de luminosité c) et par conséquent le plus plausible permet, une fois ajusté, de très bien reproduire les clôtures effectivement enregistrées (voir encore une fois la figure 18).

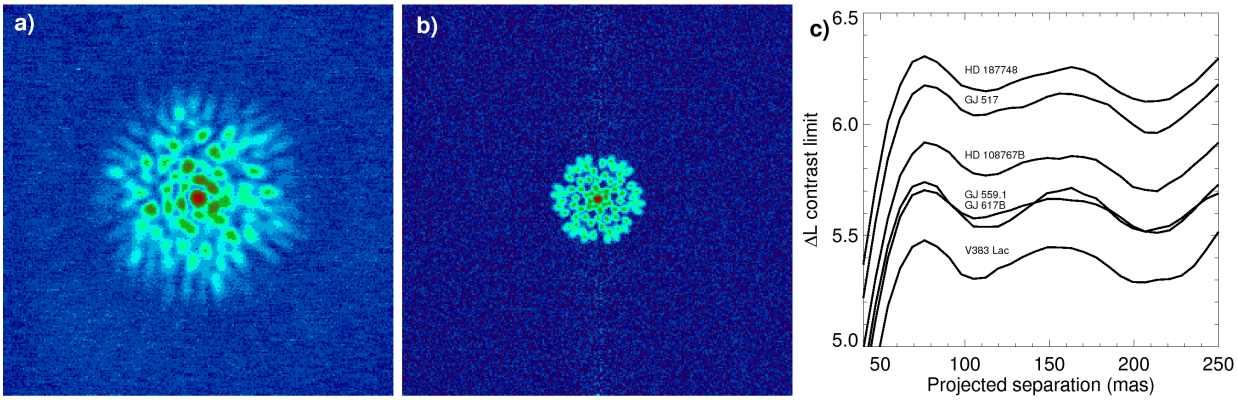


FIGURE 19: (a) Interférogramme produit par un masque à 9 trous, en bande L (λ = 3.6 μm), sur l'étoile HD 187748 observée depuis Keck avec NIRC2 le 26 juin 2008. (b) Le spectre de puissance de cette image, révélant les fréquences spatiales auxquelles le masque donne accès. (c) Limites de détection en contraste enregistrées pour les six objets observés lors de la même nuit, entre 5.5 et 6.5 magnitudes pour une séparation angulaire allant de 50 à 250 millisecondes d'arc.

Limites de détection et erreurs systématiques

Par construction, la clôture se présente comme une quantité observable non-biaisée : l'exemple présenté plus haut, de clôtures acquises au télescope et comparées à des clôtures correspondant à un modèle paramétrique est une illustration convaincante. Avec de tels observables, les erreurs statistiques associées aux mesures individuelles se moyennent en suivant le théorème central limite : l'incertitude suit une loi en $1/\sqrt{N}$ et il suffit d'augmenter le nombre N de mesures pour atteindre un niveau de précision, et donc de contraste, donné. La figure 19 résume le résultat d'observations en bande L avec l'instrument NIRC2 du télescope Keck. Pour aucun des six objets représentés, membres des associations jeunes à fort mouvement apparent [10] comme α Persei, AB Doradus et β-Pictoris, un compagnon n'a été détecté.

Les clôtures de phase étalonnées, et les incertitudes qui leur sont associées peuvent quand même servir dans une simulation de type Monté Carlo pour déterminer, à un niveau de confiance donné, les propriétés des compagnons qu'on aurait pû détecter dans l'espace séparation angulaire - contraste couvert par les observations. Ainsi, le panneau de droite de la figure 19 présente sur un même graphe, une série de courbes indiquant les limites de détection. On voit que les limites de détection en contraste (3-σ) tournent autour de six magnitudes, ce qui en échelle linéaire, correspond à un contraste de $\sim$ 1 : 250.

LA SENSIBILITÉ REPORTÉE est assez représentative des performances moyennes de la technique. Les relevés [11] [12] des associations Scorpius - Centaurus ou Taurus-Auriga, principalement conduits dans les bandes H et K, rapportent des limites de détection similaires.

10. T. M. Evans, M. J. Ireland, A. L. Kraus, F. Martinache, P. Stewart, P. G. Tuthill, S. Lacour, J. M. Carpenter, and L. A. Hillenbrand. Mapping the Shores of the Brown Dwarf Desert. III. Young Moving Groups. **ApJ**, 744:120, Jan. 2012

11. A. L. Kraus, M. J. Ireland, F. Martinache, and J. P. Lloyd. Mapping the Shores of the Brown Dwarf Desert. I. Upper Scorpius. **ApJ**, 679:762–782, May 2008

12. A. L. Kraus, M. J. Ireland, F. Martinache, and L. A. Hillenbrand. Mapping the Shores of the Brown Dwarf Desert. II. Multiple Star Formation in Taurus-Auriga. **ApJ**, 731:8, Apr. 2011

Moyenner plus de données pour réduire les incertitudes associées aux clôtures de phase atteint en pratique une limite, qui se manifeste par la présence d'une erreur systématique qui biaise les mesures de clôture.

Cette erreur systématique est invoquée dans l'interprétation des clôtures de phases $\hat{\phi}_i$ et des incertitudes σ_i qui leur sont associées. Le critère d'adéquation au modèle paramétrique binaire $f(a, \theta, c)$, est un critère de χ^2 défini par :

$$\chi^2 \equiv \sum_i \frac{(\hat{\phi}_i - f_i(a, \theta, c))^2}{\sigma_i^2}, \tag{6}$$

qui si toutes les incertitudes étaient contrôlées, devrait trouver un minimum égal à $\nu - 1$, ou ν est le nombre de degrés de liberté (nombre total de mesures de clôtures moins le nombre de paramètres) du modèle. En pratique, malgré une très bonne adéquation apparente entre les données et le modèle (voir la figure 18), le χ^2 réduit est bien supérieur à 1, suggérant que les incertitudes mesurées à partir de différentes réalisations du vecteur $\hat{\phi}$ sont sous-évaluées. La solution retenue jusqu'ici a été de rajouter, au dénominateur de l'équation 6, un terme d'erreur systématique σ_S, s'ajoutant en quadrature à l'incertitude, dont la magnitude est ajustée pour produire lors de son meilleur ajustement, un χ^2 réduit proche de 1 :

$$\chi^2 \equiv \sum_i \frac{(\hat{\phi}_i - f_i(a, \theta, c))^2}{\sigma_i^2 + \sigma_S^2}. \tag{7}$$

À L'HEURE DE LA RÉDACTION DE CE DOCUMENT, l'origine exacte de ces erreurs systématiques reste mystérieuse. La technique est utilisable et utilisée sur plusieurs instruments, qui présentent tous des comportements différents, les données NIRC2 étant par exemple beaucoup plus sensibles à l'étape de calibration que ne le sont celles de PHARO, ce comportement étant sans doute dû à la position exacte du masque non-redondant par rapport à la conjugaison du plan pupille de l'instrument. Je pense que la résolution définitive de cette question de l'origine des erreurs systématiques et la mise en oeuvre de procédures permettant de s'en affranchir requièrent un travail important, mêlant analyse de données réelles, simulation numérique et expérimentation en laboratoire. Cette question reste en partie en suspens. J'y reviendrai dans la troisième partie de ce document : ceci fait en effet partie des questions auxquelles, dans un cadre plus large, le projet KERNEL souhaite apporter des réponses.

LA QUESTION N'A CEPENDANT PAS ÉTÉ ÉVITÉE et a motivé la mise en oeuvre de plusieurs chantiers dont les conclusions ne sont que partielles. J'ai ainsi vers la fin de mon séjour à Cornell, commencer à implémenter une solution permettant de faire du filtrage spatial à Palomar. Une source possible des erreurs systématiques sur les clôtures de phase étant que les trous des masques non-redondants ne sont pas infiniment petits, ce qui fait que la condition de non-redondance

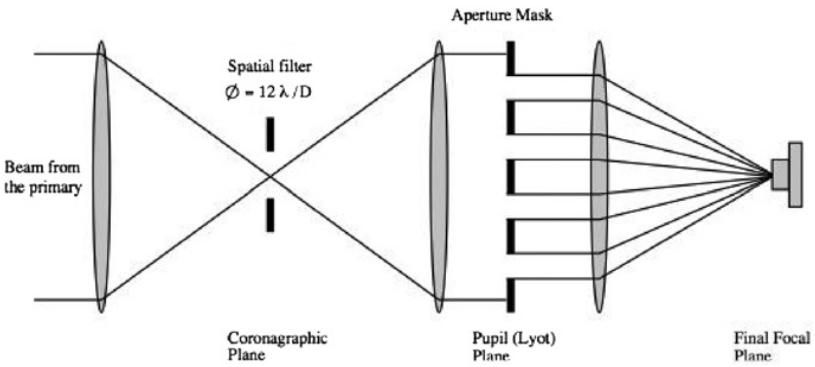


FIGURE 20: Schéma de principe de filtrage spatial simple, exploitant l'architecture de l'instrument PHARO qui donne accès à un plan focal intermédiaire avant le passage par le masque non-redondant. Ce filtre a été installé et testé au télescope entre 2007 et 2008.

n'est en réalité pas strictement vérifiée. Une solution définitive aurait été d'inclure un filtrage par trou, un peu à la manière de ce qui est désormais mis en oeuvre dans un instrument comme FIRST [13]. J'ai à cet effet commencé à me former aux techniques de nano-fabrication accessibles à Cornell. En parallèle, j'ai tenté d'exploiter l'architecture de l'instrument PHARO qui offre accès à un plan focal avant de passer par le masque non-redondant : un filtre spatial unique constitué (voir la figure 20) d'un trou de diamètre $12\lambda/D$ dans une plaque métallique, couplant un peu plus que le champ utile du masque mais lissant malgré tout le contenu à hautes fréquences spatiales du front d'onde à l'échelle d'un trou.

13. G. Perrin, S. Lacour, J. Woillez, and É. Thiébaut. High dynamic range imaging by pupil single-mode filtering and remapping. **MNRAS**, 373:747–751, Dec. 2006

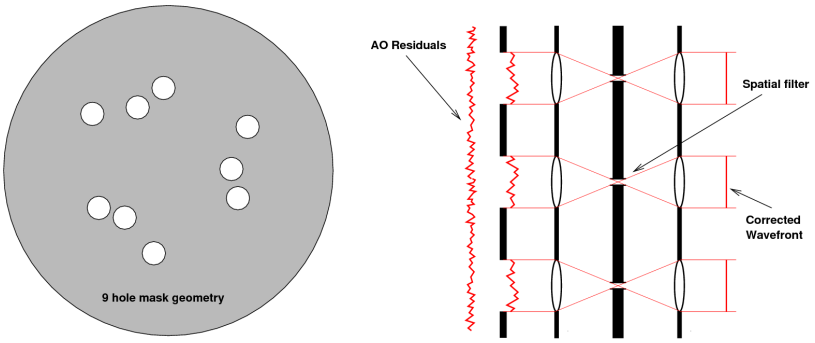


FIGURE 21: Schéma conceptuel pour un masque non-redondant incluant un filtrage spatial par trou, dont la fabrication était envisagée, en utilisant le laboratoire de nanofabrication de Cornell.

Ce masque a été installé dans l'instrument et a été testé dans des conditions d'observation non-idéales. Devenant plus ambitieux dans nos objectifs astrophysiques, le cadre était celui d'un programme de suivi de naines brunes L et M qui sont des objets faibles (magnitude en bande J comprises entre 13 et 15) et qui ont donc nécessité l'utilisation d'optique adaptative avec étoile guide artificielle laser (LGS), avec une correction adaptative médiocre (Strehl $\sim$ 5 - 10 %). En pratique, il a été difficile [14], de tirer avantage des propriétés de filtrage de ce masque : des variations de la position de la source à l'intérieur du filtre spatial lors du passage d'une source à une autre se sont traduits par des effets systématiques beaucoup plus difficiles à traiter. Etant donné que cette implémentation reproduit l'effet global de la fonction d'apodisation décrite plus haut qu'on applique en post-processing sans le problème de l'alignement relatif du masque phy-

14. D. Bernat, F. Martinache, M. Ireland, P. Tuthill, and J. Lloyd. The Use of Spatial Filtering with Aperture Masking Interferometry and Adaptive Optics. **ApJ**, 756:8, Sept. 2012

sique, cette approche n'a pas été poussée plus loin. Mon départ de Cornell a laissé de côté le chantier de mise en oeuvre d'un masque non-redondant intégrant un filtrage spatial beaucoup plus agressif par sous-ouverture (voir la figure 21) à la FIRST.

Après l'annonce de la découverte d'un compagnon planétaire en orbite autour de l'étoile β-Pictoris [15], une tentative d'observation de cet objet depuis Keck m'a sensibilisé à l'effet de la réfraction atmosphérique différentielle (β-Pic étant depuis Hawaii dans le meilleur des cas, seulement 20° au dessus de l'horizon) comme source plausible d'erreurs systématiques dans les clôtures. Pour référence, j'inclus donc dans ce mémoire, une note de recherche écrite vers 2009 et non-publiée. Cette note rassemble des résultats de simulation de l'effet de la réfraction atmosphérique et propose une procédure de calibration des données qui permet de s'en affranchir.

15. A.-M. Lagrange, M. Bonnefoy, G. Chauvin, D. Apai, D. Ehrenreich, A. Boccaletti, D. Gratadour, D. Rouan, D. Mouillet, S. Lacour, and M. Kasper. A Giant Planet Imaged in the Disk of the Young Star β Pictoris. *Science*, 329:57–, July 2010

Draft version April 3, 2009
Preprint typeset using LaTeX style emulateapj v. 03/07/07

RESEARCH NOTE: DIFFERENTIAL ATMOSPHERIC REFRACTION & APERTURE MASKING INTERFEROMETRY

Frantz Martinache[1]

Draft version April 3, 2009

ABSTRACT

The success of non redundant aperture masking inteferometry essentially relies of the self-calibrating properties of closure phases. For detections at moderate contrast ratio, our standard calibration procedure has proven quite satisfactory, as demonstrated by the repeted several discoveries of binary companions around field and cluster stars. If the calibration process was perfect, the contrast ratio limit of a dataset should theoretically increase with the total amount of data (square root, really). In practice, the limit however seems restricted to a typical ΔH ∼5-6, which implies that some systematic signal in the closure phase does not properly calibrate. The paper explores the effect of differential atmospheric refraction that can account for some part of the systematic systematic signal found in the closure phase.

Subject headings: techniques: AO, interferometric

1. INTRODUCTION

It seems most of our observations, no matter how great the seeing is, are limited to typical contrast ratio of 5 to 6 magnitudes, where calibration and increased amount of data should theoretically bring better contrasts. There must therefore exist some systematic signal in the closure phase that limits the quality of the calibration with the pipeline the way it works now.

Observations of β-Pictoris from Keck last winter showed that at least at low elevation, differential atmospheric refraction (there is no corrector in the AO system) could explain the otherwise somewhat disappointing limits.

This observation motivated the development of some sort of simulator, capable of generating data that can then be reduced with the existing data reduction pipeline. The simulator includes: photon noise, poisson sky noise (the easy stuff), some sort of AO-corrected kolmogorov phase screen (that needs some work), and differential atmospheric refraction (DAR). This research note gathers some results of this simulation, and shows that (if the model is correct), DAR can account for a significant fraction of the systematic errors even at fairly high elevation.

Plan: Presentation of the model; basic simulation and demonstration of the effects; study of the DAR closure phase signature and comparison with a partially resolved binary signal; discussion: how can we address the problem.

Filippenko (1982); Arribas et al. (1999).

2. DIFFERENTIAL ATMOSPHERIC REFRACTION

The refractive index of air is a function of wavelength. Except exactly at zenith, the atmosphere therefore behaves like a refraction prism that disperses the light. The effect becomes obvious when taking images of an object at low elevation with a broad band filter at a somewhat short wavelength: the PSF is not round anymore but shaped more like a water droplet. The model and formulas used in this note are extracted from pp. 262-263 of Allen's Astrophysical Quantities (4^{th} edition).

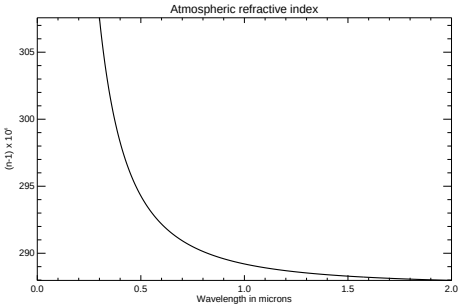


Fig. 1.— Refractive index of standard dry air versus wavelength.

The refractive index n_s of air at pressure $p_s = 1013.25$ hPa and temperature $T_s = 288.15$ K is given by:

$$n_s = 1 + (64.328 + \frac{29498.1}{146 - \sigma^2} + \frac{255.4}{41 - \sigma^2}) \times 10^{-6}, \quad (1)$$

where $\sigma = \lambda^{-1}$ and λ is the vacuum wavelength in μm. To account for the variations of atmospheric pressure p and temperature T, one applies the following correction:

$$n - 1 = (n_s - 1)(pT_s/p_sT). \quad (2)$$

Fig. 1 presents the curve of atmospheric refractive index as a function of wavelength for a standard atmosphere ($p = 1013.25$ hPa; $T = 273.15$ K). From one to two μm, the index varies by less than 2×10^{-6}. This very weak variation however affects the calibration of the masking data.

The atmospheric refraction is defined as the difference between the true (z_t) and apparent (z_a) zenith distances. For zenital distances less than 80°, the refraction R_{n_0} can be approximated as:

$$R_{n_0} = R_0 \tan z_t, \quad (3)$$

where the constant of refraction R_0 is:

[1] Subaru Telescope, Hilo HI

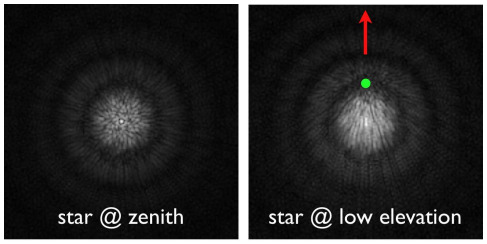


FIG. 2.— Illustration of the effect of differential atmospheric refraction on full H-band PHARO 9-hole mask data. The left panel shows the interferogram of a single star observed at zenith (therefore in the absence of DAR). The right panel shows the interferogram of another single star, this time observed at 20° elevation. The red arrow points toward the zenith: the actual orientation of the dispersion depends on the parallactic angle of the source. The green dot marks the point where the "normal" bandwitdh smearing observed on the left image compensates the dispersion.

$$R_0 = \frac{n_0^2 - 1}{2n_0^2}, \tag{4}$$

with n_0 evaluated for a standard atmosphere. For other temperature and pressure conditions, the following correction applies:

$$R = R_0(pT_0/p_0T). \tag{5}$$

This model can also be refined to include the effect of water vapor in the atmosphere that changes the refractive index. That won't be necessary here. At this point of the study, we're not really interested by an absolute model of the refraction but rather by a model relative to a given wavelength.

According to this model, at 1.6 μm, the refraction constant for dry air is $R = 52.42$ arcsec. The important number is however how much the refraction constant varies across the full bandpass of a filter: $\Delta R = 46.3$ mas for H, $\Delta R = 8.8$ mas for L_P, which tells how much the PSF spreads due to the diffraction for a zenith distance $z_t = 45°$.

On PHARO, the plate scale is approximately 25.2 mas/pixel. For $z_t = 80°$, $\Delta R \sim 250$ mas: the core of the PSF is elongated by a factor 10 along the direction of the parallactic angle. Fig. 2 compares two interferograms taken in H-band with PHARO (simulated data): one star at zenith and a comparable star at 80° zenith distance.

3. DAR & NRM

The interest of NRM essentially lies in the high quality self-calibration properties of closure phases. Systematic closure phases are substracted from a science target after acquiring a reference star in comparable observing conditions.

The effect of atmospheric refraction doesn't evade this general rule and if you're lucky enough to find a calibrator that's right next to your science target, then DAR or not, you're pretty safe as the DAR closure phase perfectly calibrates. However, depending on the brightness and spectral type of the science target, one sometimes need to slew the telescope by as much as 20 degrees. The differential refraction closure phase signature won't calibrate and one is left with a systematic term in the closure phase, which if strong enough, will limit the attainable contrast on a given dataset, no matter how much data one takes.

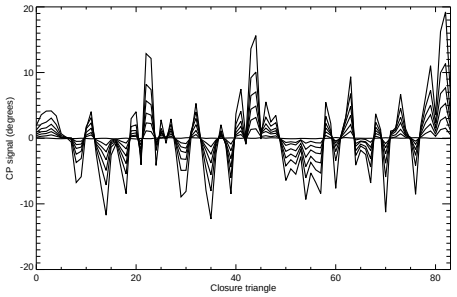


FIG. 3.— Closure phase signature of the differential atmospheric refraction for zenith distances going from 0 (horizontal line) to 60 degrees (signal of maximum amplitude) with a 10 degree increment. The amplitude of the DAR closure phase is proportional to $\tan z_t$.

elev.	90°	80°	70°	60°	50°	40°	30°	20°
90°	0.06	0.60	1.20	1.91	2.80	4.10	6.54	15.27
80°	0.57	0.05	0.63	1.35	2.26	3.58	5.97	14.78
70°	1.21	0.61	0.05	0.71	1.65	2.93	5.39	14.28
60°	1.95	1.34	0.72	0.05	0.90	2.24	4.66	13.71
50°	2.78	2.27	1.62	0.94	0.05	1.34	3.79	12.92
40°	4.07	3.55	2.96	2.24	1.31	0.05	2.47	11.85
30°	6.53	5.99	5.35	4.68	3.79	2.45	0.07	9.63
20°	15.28	4.80	4.26	3.75	2.97	1.88	9.62	0.09

TABLE 1
DIFFERENTIAL ATMOSPHERIC REFRACTION RMS CLOSURE PHASE (IN DEGREES) AS A FUNCTION OF ELEVATION ANGLES OF BOTH THE SCIENCE TARGET (COLUMN) AND THE CALIBRATOR (ROW).

3.1. *Targets at transit*

Figure 3 plots the calibrated closure phase of a star observed exactly at zenith, with a calibrator observed while it transits (parallactic angle is zero) at zenith distance going from 0 to 60 degrees. The figure shows that the closure phase signal is directly proportional to $\tan z_t$.

An ideal calibration procedure should leave (within statistical errors) zero closure phases. One convenient way to quantify the strength of the systematic DAR signal in the closure phase is therefore to use the closure phase RMS itself.

Table 1 summarizes the result of a series of numerical experiments: each dataset consists of ten frames in semi-perfect observing conditions on two reference stars observed while transiting at elevation going from 20 to 90 degrees.

Several comments on this table:

- Diagonal values are zero. If you can find a good calibrator right next to your science target, the DAR perfectly calibrates (well, at least in these simulations).
- The matrix is symetric.
- Even for small zenith distances, a 10°-difference of elevation has a 0.6° signature in the closure phase.

This last comment just shows how important this effect can be: for angular separations between 0.5 and 1 λ/D, the closure phase of a 100:1 contrast ratio binary also exhibits a typical $\sim 0.6°$ signature.

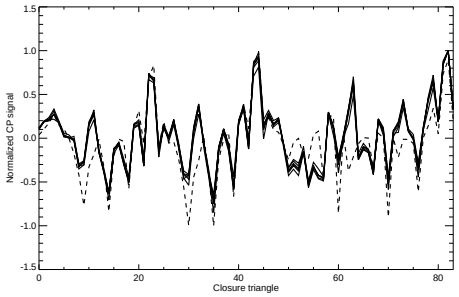


FIG. 4.— Similitude of DAR (solid lines) and low-sep low-contrast binary (dashed line) closure phase signal. All curves are normalized

3.2. *DAR and binary signal*

It's been shown that in first order approximation, differential refraction elongates the PSF in the direction of the parallactic angle, so one intuits that the DAR closure will mimic a partially resolved ($< \lambda/D$) low contrast binary, with a strong correlation between the actual contrast and angular separation.

From the simulations that produced the results of Table 1, one selects these for which the source is at zenith (table's first row) and the calibrator is between 90 and 30° elevation.

Figure 4 plots all the corresponding normalized closure phases, (solid lines) and compares them to the normalized closure phase of a low-separation low-contrast binary. These simulations confirm one's intuition and show that the DAR closure phases are very similar (yet not exactly) to the closure phase of a barely resolved binary.

3.3. *Parallactic angle*

The simulation has so far only been concerned with targets observed while they transit: the effect of the refraction is therefore always in the same direction on the detector and in the uv plan. In practice, one typically observes targets at different parallactic angles. For PHARO, which is at Cassegrain focus on an equtorial telescope, the direction of dispersion relative to the detector will therefore change with the parallactic angle. For NIRC2, which is normally used in vertical angle mode at the Nasmyth focus of an alt-azimutal telescope, the direction of dispersion relative to the detector remains a constant.

Unless the elevation h you're pointing a target that is low enough to be below the pole (i.e $h < \phi$, with ϕ the local latitude), the parallactic angle can take any value between 0 and 360°.

Table 2 presents the result of another set of simulations for which science target and calibrator share the same elevation but are observed at different parallactic angles.

This part needs more work, formulas for P.A. not complete yet.

4. TARGET SPECTRA

This is a little off-topic, since this note focuses on DAR, but still worth mentioning. For all the simulations made so far, the spectrum of the cal and science target is considered flat over the full H band. To get the feeling of how the actual spectra will affect these simulations, the

P.A.	**0°**	**15°**	**30°**	**45°**	**60°**
0°	0.05	0.44	0.86	1.38	1.84
30°	0.89	0.46	0.08	0.52	1.02
60°	1.88	1.46	1.00	0.55	0.04
90°	2.71	2.42	1.98	1.57	0.99
120°	3.41	3.12	2.78	2.36	1.90
150°	3.77	3.57	3.32	2.98	2.53
180°	3.83	3.69	3.54	3.25	2.94
210°	3.52	3.46	3.36	3.21	3.04
240°	2.94	3.00	3.05	3.04	3.06
270°	2.26	2.45	2.64	2.80	2.95
300°	1.56	1.84	2.14	2.47	2.80
330°	0.73	1.19	1.62	2.04	2.42

TABLE 2
DAR: SCIENCE TARGET AND CALIBRATOR HAVE SAME ELEVATION (30°), BUT ARE OBSERVED AT DIFFERENT PARALLACTIC ANGLES (P.A.). EACH COLUMN IS FOR A SCIENCE TARGET OF P.A. BETWEEN 0 AND 60°. EACH LINE IS FOR A CALIBRATOR OF P.A. GOING FROM 0 TO 360°.

elev.	**90°**	**80°**	**70°**	**60°**	**50°**	**40°**	**30°**	**20°**
90°	0.06	0.81	1.66	2.67	3.93	5.67	8.89	18.93
80°	0.82	0.08	0.90	1.86	3.10	4.88	8.07	18.20
70°	1.63	0.87	0.06	0.99	2.24	4.01	7.23	17.34
60°	2.63	1.83	1.00	0.08	1.24	3.03	6.25	16.45
50°	3.91	3.12	2.23	1.27	0.07	1.79	5.00	15.33
40°	5.66	4.87	4.05	3.09	1.79	0.04	3.20	13.71
30°	8.88	8.06	7.22	6.22	4.98	3.21	0.04	10.66
20°	18.97	18.17	17.43	16.48	15.35	13.71	10.71	0.06

TABLE 3
REVISED VERSION OF TABLE 1, NOW INCLUDING THE EFFECT OF A RAYLEIGH-JEANS SPECTRA.

star was modeled as a black-body in the Rayleigh-Jeans limit, for which the number of photons per unit of wavelength is proportional to λ^{-3}.

This ponderation gives more weight to the "blue-side" of the spectrum and therefore reinforces the level of asymmetry in the DAR-contaminated interferograms. The experience of Sec. 3.1 was repeated including this effect. Table 3 summarizes the results of the simulations and shows that a Rayleigh-Jeans spectra indeed strengthens the DAR closure phase signature.

Note that this doesn't change the structure of the closure phase signal itself but mostly its amplitude. The signal is still proportional to $\tan z_t$.

More generally, the impact of the differences in spectral type from science target to calibrator will also matter. This is likely to become the dominant source of systematic errors when observing from space with JWST (at least, no DAR there...). A library of low-resolution spectra and some simulations will soon become mandatory.

5. STRATEGIES TO ADDRESS THE PROBLEM

The strong similitude between the DAR and low-sep binary signals makes this whole DAR shebang a pretty annoying problem to deal with. At this point, two strategies are possible, and it's very likely that one is just too optimistic...

5.1. *A realistic simulation*

Although the model that's been used here is interesting, it certainly does not exactly reproduces the subtles yet important nuances that nature adds to our models.

Ideally, one would like be to be able to rely on a model that given the coordinates of your objects and the time of your observations, would tell you exactly how much DAR

systematics you should substract from your calibrated closure phase, before doing any sort of model fit to it. There are plenty of reasons for which you won't believe this approach is realistic.

There may be work to do for students looking for work on this topic: for an observing night, take only calibrators and try to pair them and see the residual closure phases you get, consider this as a DAR closure phase skymap for a given night. Use this to calibrate the model used in this research note (or something more elaborate) and then use the model to better calibrate the data on science targets.

5.2. *A new approach to calibration*

Currently, the calibration and the model fit of the closure phase are considered as two completely independent steps of the data reduction process:

- The current calibration script uses a accept/reject strategy. The user is invited to select the calibrators that produce closure phase histograms and visibilities that are consistent.
- The then calibrated closure phases are model-fitted, typically looking for a binary solution in a 3-parameter χ^2 space.

According to what we see in this research note, it seems that for high contrast companions, because of things like DAR, difference of spectral types, and other things we haven't thought of yet, the current calibration procedure is not adequate, and needs to be revised.

For DAR, one can always consider that for small differences of elevation, the DAR closure phase variations are linear. Instead of a accept/reject calibrator strategy, the calibration could consist in another χ^2 minimization problem:

$$\chi^2 = \sum_{triangles}^{N} (\sum_{i=1}^{n} \alpha_i (\text{src CP})_i - \sum_{j=1}^{p} \beta_j ((\text{cal CP})_j)^2, \quad (6)$$

where $(\alpha_i)_{i=1}^n$ and $(\beta_j)_{j=1}^p$ are the $n+p$ coefficients that will optimize the calibration of the data.

One should then run a similar procedure including the possibility of a binary in a $n+p+3$ parameter fit.

Verify the contrast limits by simulating fake binaries in the data.

6. CONCLUSION

This stuff obviously needs a fair bit of work, and requires to write and test a new IDL calibrate procedure.

There is one complication with NIRC2: the rotation of baselines in the uv plane that changes with the PA, which makes the proposed procedure impossible to apply as is. Somehow need to de-rotate the closure phases, which I am not sure is possible yet.

REFERENCES

Arribas, S., Mediavilla, E., García-Lorenzo, B., del Burgo, C., & Fuensalida, J. J. 1999, A&AS, 136, 189
Filippenko, A. V. 1982, PASP, 94, 715

Conclusion et publications

Même si son utilisation reste limitée à des contrastes modérés, l'interférométrie des masques non-redondants continue d'ouvrir un accès à une région de l'espace des paramètres serré sur les très faibles séparations angulaires (entre 0.5 et 4 λ/D), qui n'est encore que très partiellement couvert par les observations coronographiques. La seconde partie de ce mémoire présentera le travail que j'ai mis en place dans le cadre du projet SCExAO pour qu'un instrument imageur coronographique puisse commencer à sonder efficacement ce régime de séparation.

Les erreurs systématiques présentes dans l'analyse des données d'interférométrie à masque non-redondant conservent une part d'inconnue, malgré les éléments d'information apportés par les investigations évoquées plus haut. A l'heure actuelle, la contribution la plus importante à ces systématiques semble être la variance temporelle de la phase [16]. Un des objectifs du travail expérimental couvert dans le cadre du projet KERNEL qui généralise beaucoup des idées évoquées dans ce chapitre permettra, c'est mon ambition, d'affiner notre compréhension de ces systématiques et de proposer des moyens de s'en affranchir.

16. M. J. Ireland. Phase errors in diffraction-limited imaging: contrast limits for sparse aperture masking. **MNRAS**, 433:1718–1728, Aug. 2013

Pour conclure cette section du manuscrit, je présente deux articles dont je suis premier auteur, décrivant deux systèmes binaires d'étoiles naines de type spectral M : GJ 623 et GJ 164 [17]. Pour ces objets, on voit que les observations interférométriques faites depuis les télescopes de Palomar et de Keck qui résolvent le compagnon binaire, apportent des contraintes très fortes [18] sur les paramètres orbitaux et les propriétés physiques qui en sont déduites. Combinées à des observations astrométriques dans le cas de GJ 164 [19] ou de vélocimétrie radiale dans le cas de GJ 623 [20], ces mesures interférométriques ont conduit à des estimations à la fois précises et justes, des masses dynamiques des membres de ces associations.

17. Ces objets font partie du catalogue Gliese-Jahreiss qui répertorie les objets faisant partie du voisinage Solaire (distance inférieure à 25 pc).

18. G. Torres. Substellar Companion Masses from Minimal Radial Velocity or Astrometric Information: a Monte Carlo Approach. **PASP**, 111:169–176, Feb. 1999

19. F. Martinache, B. Rojas-Ayala, M. J. Ireland, J. P. Lloyd, and P. G. Tuthill. Visual Orbit of the Low-Mass Binary GJ 164 AB. **ApJ**, 695:1183–1190, Apr. 2009

20. F. Martinache, J. P. Lloyd, M. J. Ireland, R. S. Yamada, and P. G. Tuthill. Precision Masses of the Low-Mass Binary System GJ 623. **ApJ**, 661:496–501, May 2007

The Astrophysical Journal, 661:496–501, 2007 May 20
© 2007. The American Astronomical Society. All rights reserved. Printed in U.S.A.

PRECISION MASSES OF THE LOW-MASS BINARY SYSTEM GJ 623

Frantz Martinache,[1] James P. Lloyd,[1] Michael J. Ireland,[2] Ryan S. Yamada,[1] and Peter G. Tuthill[3]
Received 2006 December 4; accepted 2007 February 7

ABSTRACT

We have used aperture masking interferometry and adaptive optics (AO) at the Palomar 200 inch telescope to obtain precise mass measurements of the binary M dwarf GJ 623. AO observations spread over 3 yr combined with a decade of radial velocity measurements constrain all orbital parameters of the GJ 623 binary system accurately enough to critically challenge the models. The dynamical masses measured are $m_1 = 0.371 \pm 0.015\ M_\odot$ (4%) and $m_2 = 0.115 \pm 0.0023\ M_\odot$ (2%) for the primary and the secondary, respectively. Models are not consistent with color and mass, requiring very low metallicities.

Subject headings: binaries: general — stars: luminosity function, mass function — techniques: interferometric

Online material: color figures

1. INTRODUCTION

The mass of a star, along with its metallicity and age, is the fundamental parameter that determines its position along an evolutionary track. Even if binarity, rotation, magnetic fields, and other parameters also affect stellar interiors, what is known as the Vogt-Russel theorem remains an important rule, and the mass an essential parameter of stellar evolution. Binary systems offer the ideal test to infer dynamical masses independent of the use of a stellar model, such as those of Baraffe et al. (1998), or an empirical mass-luminosity (*M-L*) relation. Such relations, available for both visible (Henry et al. 1999) and near-infrared (Delfosse et al. 2000), are important astrophysical tools, fairly well constrained for intermediate-mass stars. However, the solar neighborhood is dominated by low-mass stars, in both number and total mass (Henry 1998), and as dust condenses in the atmospheres of these cool stars, the models meet new unknowns. There are also puzzles in the stellar structure of M dwarfs and the extrapolation of these models to substellar objects remains untested. Below 0.6 $M_\odot$, both *M-L* relations and models will benefit from model-independent determinations of high-precision stellar masses, which can be achieved by combining radial velocimetry measurements and high angular resolution imaging (Ségransan et al. 2000; Delfosse et al. 2004). The complementarity of the two techniques yields substantial benefits, even in the regime where the time baseline of the observations is shorter than the orbital period (Eisner & Kulkarni 2002).

We have observed known M dwarf binaries with precise radial velocity measurements published by Nidever et al. (2002). These late-type M dwarfs are ideal observing targets for imaging with the PALAO adaptive optics (AO) system and the PHARO infrared camera (Hayward et al. 2001), optimized for the near-infrared bands. At small angular separation (i.e., $<2\lambda/d$), the sensitivity of the detection of faint companions can be improved by combining AO with aperture-masking interferometry (Pravdo et al. 2006; Lloyd et al. 2006). This paper provides the astrometry of the binary system GJ 623, successfully observed with both AO and AO + aperture masking. Combined with the radial velocity measurement, the astrometry provides precise dynamical masses (better than 2%) of the GJ 623 binary system. Combined with the *J*, *H*, *K* photometry, this measurement adds new constraints to the models and *M-L* relations.

2. THE OBSERVATIONS

2.1. *The Primary*

GJ 623 (aka LHS 417, HIP 80346) is a high proper motion M2.5 dwarf 8 pc from the Sun. It is a long-known astrometric binary, first characterized by Lippincott & Borgman (1978). GJ 623 has proven to be an ideal test for different observational techniques, such as radial velocimetry (Marcy & Moore 1989; Nidever et al. 2002) and speckle interferometry (McCarthy & Henry 1987). This binary system has also been directly imaged in the visible with the COSTAR-corrected *HST* FOC (Faint Object Camera; Barbieri et al. 1996).

Our observations of GJ 623 were performed with the PHARO instrument on the Palomar 200 inch (5 m) telescope and with the NIRC2 instrument on the Keck II telescope. The companion of GJ 623 was detected in six observing runs: 2003 September, 2004 June, 2005 January, and 2006 February, at Palomar, using the *J*, *H* and K_s filters, and 2006 June and August at Keck, with the 1.58 μm $H_{\rm cont}$ filter. Tables 1 and 2, respectively, gather the astrometric and photometric measurements made at Palomar and Keck.

Four of our observations consist of conventional direct imaging with AO. These images were dark subtracted, flat-fielded, and analyzed with a custom IDL program, using the latest PHARO and NIRC2 plate scale and orientation characterization by Metchev & Hillenbrand (2004). The location of the companion (angular separation and position angle) and the contrast ratio are precisely determined with a cross-correlation of the images. Not surprisingly, doubling the diameter of the telescope (cf. the Keck data points) significantly improves the precision of the measurements and roughly reduces the error bars by a factor 3.

The data also include observations in aperture-masking interferometry with AO. This technique (Tuthill et al. 2000), recently described by Lloyd et al. (2006), is appropriate to the detection of faint companion at small angular separation (typically less than $2\lambda/D$), where direct AO imaging has so far proven difficult.

Indeed, present AO systems focus on achieving high contrast at moderate angular separation $>4\lambda/D$. Below this limit, the variance of the speckle background dominates the photon noise by

[1] Department of Astronomy, Cornell University, Ithaca, NY.
[2] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA.
[3] School of Physics, University of Sydney, Sydney, NSW, Australia.

TABLE 1
ASTROMETRIC MEASUREMENTS AT PALOMAR AND KECK: ANGULAR SEPARATION AND POSITION ANGLE OF GJ 623B

Julian Date (−2,450,000)	Band	Pupil (see text)	Separation (mas)	P.A. (deg)
2896.6	K_s	9 hole	240.4 ± 9.7	79.31 ± 2.0
3163.8	[Fe II]	18 hole	340.5 ± 20.4	49.16 ± 4.3
3402.1	K_s	Full aperture	350.7 ± 2.2	28.66 ± 0.5
3780.0	J, H, K_s	Full aperture	257.1 ± 3.1	−12.76 ± 0.6
3909.5	$H_{\rm cont}$	Nirc2	176.7 ± 1.1	318.2 ± 0.2
3962.3	$H_{\rm cont}$	Nirc2	138.7 ± 0.9	295.4 ± 0.2

several orders of magnitude (Racine et al. 1999). This prevents us from calibrating the point-spread function (PSF) with a precision sufficient to discriminate faint companions from the bright star's residual speckles.

This issue can be evaded by using interferometric techniques. By sampling a few spatial frequencies only, a mask located in the pupil plane permits us to completely decompose the PSF into a finite set of Fourier components. Non-redundancy of the baselines passed by the mask ensures that each frequency is sampled only once and the visibilities can be used to form closure phases (Baldwin et al. 1986; Readhead et al. 1988; Nakajima et al. 1989). This observable rejects both atmospheric noise and calibration errors of the wave front sensor. The only drawback is the transmission of such a mask: between 5% and 15% for the ones used at Palomar. However, in the speckle noise limited regime, light loss does not result in a loss of sensitivity.

Our analysis also uses a *HST* observation dating to 1994, and originally published by Barbieri et al. (1996): an angular distance of 330 ± 20 mas and a position angle of 7.0° ± 2.6° on 1994 June 11. Even though this is inferior to the precision achieved with Palomar and Keck, which are much larger telescopes, the fact that this observation was performed more than 10 yr ago adds an important constraint on the period of the binary.

From the 2MASS catalog, the apparent magnitudes of GJ 623AB are $J = 6.638 \pm 0.024$, $H = 6.141 \pm 0.021$, and $K = 5.915 \pm 0.023$.

The *Hipparcos* parallax $\pi = 124.34 \pm 1.16$ mas originally published by Perryman et al. (1997) was based on the radial velocimetry measurement of Marcy & Moore (1989). The parallax was recalculated by Jancart et al. (2005) after the publication of an improved radial velocity curve by Nidever et al. (2002). It is this revised parallax $\pi = 125.81 \pm 1.19$ mas (D. Pourbaix 2006, private communication) that we adopt here. The following absolute magnitudes can be deduced for GJ 623AB: $M_J = 7.137 \pm 0.052$, $M_H = 6.640 \pm 0.051$, and $M_K = 6.414 \pm 0.052$.

TABLE 2
PHOTOMETRIC MEASUREMENTS OF GJ 623B

Julian Date (−2,450,000)	Filter	Δmag
2896.6	K_s	3.051 ± 0.826
3163.8	Fe II	2.903 ± 0.476
3402.1	K_s	2.604 ± 0.047
3780.0	J	2.691 ± 0.038
	H	2.860 ± 0.039
	K_s	2.789 ± 0.014
3909.5	$H_{\rm cont}$	2.794 ± 0.033
3962.3	$H_{\rm cont}$	2.781 ± 0.016

TABLE 3
ORBITAL ELEMENTS

Parameter	Nidever	This Work
Π (mas)	...	125.81 ± 1.19
a (AU)	...	1.894 ± 0.019
α (mas)	...	237.28 ± 0.88
e	0.67 ± 0.01	0.631 ± 0.002
i (deg)	...	154.0 ± 0.1
Ω_1 (deg)	...	98.5 ± 0.47
ω_0 (deg)	251 ± 1	248.68 ±0.46
P (days)	1366.1 ± 0.4	1365.6 ± 0.3
T_P (reduced JD)	1298 ± 10	1313.3 ± 0.6
V_0 (km s^{-1})	−27.654 ± 0.3	−27.729 ± 0.005
A_1 (kms^{-1})	...	3.57 ± 0.01

3. METHOD: EXTRACTING ORBITAL PARAMETERS

The starting point in the determination of the orbital parameters of the GJ 623 system is the set of radial velocity measurements published by Nidever et al. (2002) and available in the online version of the Ninth Catalogue of Spectroscopic Binary Orbits[4] (Pourbaix et al. 2004). The parameters deduced from the Keplerian fit to the data are summarized in Table 3. They have been directly used by Jancart et al. (2005) to reprocess the *Hipparcos* intermediate astrometric Data of van Leeuwen & Evans (1998) for a complete characterization of the dynamical elements of the system.

The approach detailed in this paper consists of combining together, with no a priori assumptions, the original radial velocity data and our astrometric observations.

3.1. *Keplerian Orbits*

One uses the standard two-body solution to parameterize the location of the companion orbiting GJ 623A. To an observing date t, one associates an angle M called the mean anomaly:

$$M(t) = \frac{2\pi}{P}(t - T_P), \tag{1}$$

where P and T_P, respectively, represent the orbital period and the epoch at the periastron passage. In the orbital plane, the $(x,\ y)$ coordinates can simply be expressed as a function of another angle, the eccentric anomaly E. It is the angle between the direction of the periastron and the current position of the companion, projected onto the ellipse's circumscribing circle perpendicularly to the major axis, measured at the center of the ellipse (cf. Fig. 1):

$$M(E) = E - e\sin E \tag{2}$$

$$x(E) = a(\cos E - e) \tag{3}$$

$$y(E) = a\sqrt{1 - e^2}\sin E, \tag{4}$$

where a and e, respectively, represent the semimajor axis and the eccentricity of the orbit. Equation (2) is called Kepler's equation. One solves this equation, i.e., finds the value of E associated with a given M by using the following classical iterative algorithm:

$$E_0 = M$$
$$E_{n+1} = M + e\sin E_n. \tag{5}$$

[4] See http://sb9.astro.ulb.ac.be.

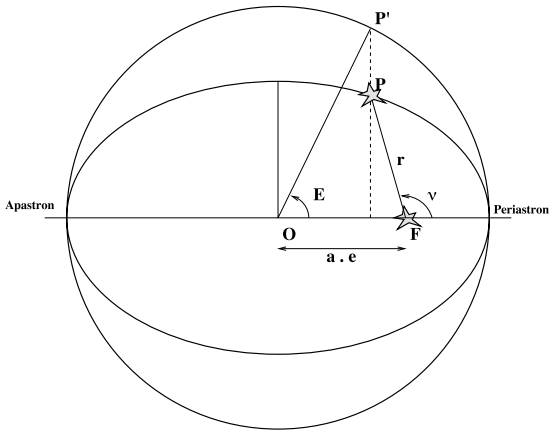


FIG. 1.—Parameterization of an elliptic orbit. Geometric relation between the eccentric anomaly E and the true anomaly ν. [*See the electronic edition of the Journal for a color version of this figure.*]

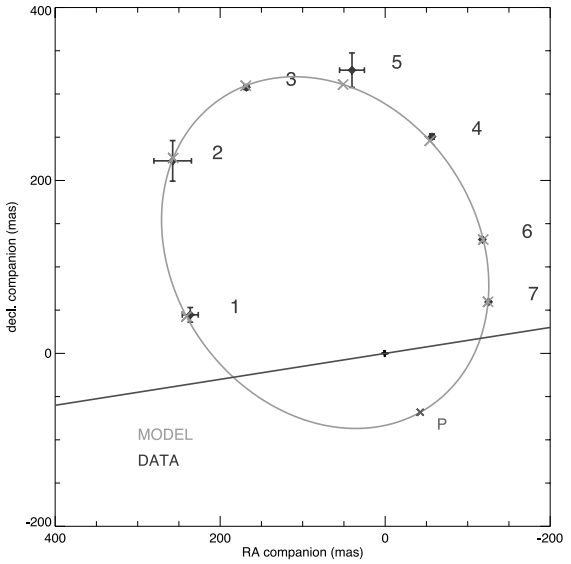


FIG. 2.—Orbit of GJ 623B. Measurements and associated uncertainties are represented by black points with error bars. *Points 1–4*: PHARO observations. *Point 5*: *HST* observations reported in Barbieri et al. (1996). *Points 6–7*: Nirc2 observations. *Gray lines*: Keplerian fit to the orbit. The straight line is the line of nodes, and the point P marks the position of the periastron. [*See the electronic edition of the Journal for a color version of this figure.*]

3.1.1. *Radial Velocity Orbital Models*

One can use the standard two-body solution (eqs. [3] and [4]) to compute the coordinates of the primary component in the orbital plane, relative to the center of mass of the system:

$$\begin{bmatrix} x_P(E) \\ y_p(E) \end{bmatrix} = \frac{m_2}{m_T} \begin{bmatrix} x(E) \\ y(E) \end{bmatrix}, \tag{6}$$

where m_T and m_2, respectively, stand for the total mass (primary+secondary) and the mass of the secondary. In the orbital plane, the velocity vector can be calculated with the following partial derivative:

$$V = \frac{\partial}{\partial t} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{\partial}{\partial E} \begin{pmatrix} x \\ y \end{pmatrix} \frac{\partial E}{\partial M} \frac{\partial M}{\partial t}. \tag{7}$$

This derivation leads to the following x and y components of the velocity:

$$\begin{bmatrix} v_x(E) \\ v_y(E) \end{bmatrix} = \frac{m_2}{m_T} \frac{2\pi a}{P(1 - e\cos E)} \begin{bmatrix} -\sin E \\ \sqrt{1 - e^2}\cos E \end{bmatrix}. \tag{8}$$

The radial velocity (RV) is the component of this velocity projected on the line of sight. Its expression therefore involves both the argument of the periastron ω_0 and the inclination of the system i. The RV is not sensitive to the orientation of the system on the sky, i.e., to the value of the argument of the ascending node Ω_1:

$$RV(E,\ \omega_0,\ i) = \left[v_x(E)\sin\omega_0 + v_y(E)\cos\omega_0\right]\sin i + V_0, \tag{9}$$

where V_0 is a constant term, the mean RV.

3.1.2. *Astrometric Orbital Models*

Unlike the case of RV data for which equations are expressed from the center of mass of the system, the astrometric measurements gathered in Table 1 provide the instant position of the companion relative to the primary. In this frame, the trajectory of the companion is an ellipse whose focus is the primary component. The parametric equation of this ellipse is

$$r(\nu) = \frac{a(1 - \epsilon^2)}{1 + \epsilon\cos\nu}, \tag{10}$$

where a is the semimajor axis, ϵ the eccentricity ($0 < \epsilon < 1$), and ν an angle called the true anomaly, which is the angle between the direction of the periastron and the current position of an object on its orbit (P in Fig. 1), measured at the focus of the ellipse.

Figure 1 illustrates the one-to-one correspondence between eccentric and true anomaly. The projection of the point P on the major axis provides the following relation:

$$r\cos\nu = a(\cos E - \epsilon). \tag{11}$$

Together, equations. (10) and (11) lead to the primary to secondary distance r and the true anomaly ν, as functions of the eccentric anomaly E:

$$r = a(1 - e\cos E), \tag{12}$$

$$\cos\nu = \frac{\cos E - \epsilon}{1 - \epsilon\cos E}. \tag{13}$$

At a given observing date t, one needs once more to solve Kepler's equation (cf. eq. [2]) to determine the corresponding eccentric anomaly E. The location of the companion along the orbit is provided by equations (12) and (13).

Contrary to the RV, which is the component of the velocity projected on the line of sight, one measures here the position of the companion projected on the celestial sphere. Right ascension, α, and declination, δ, of the secondary (relative to the primary) are given by the following relations:

$$\alpha = r[\cos(\nu + \omega_0)\sin\Omega_1 + \sin(\nu + \omega_0)\cos i\cos\Omega_1], \tag{14}$$

$$\delta = r[\cos(\nu + \omega_0)\cos\Omega_1 - \sin(\nu + \omega_0)\cos i\sin\Omega_1]. \tag{15}$$

3.2. χ^2 *Fitting*

We use the model presented in the previous section to fit a nine-parameter model to 26 observables. The observables are

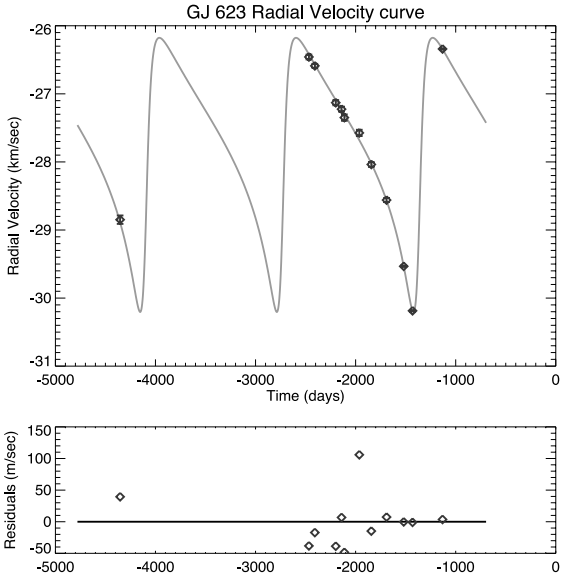


FIG. 3.—Radial velocity curve of GJ 623A. The measurements and associated uncertainties reported by Nidever et al. (2002) are represented by diamonds. The solid line represents the Keplerian orbit with the parameters derived by χ^2 fitting. [*See the electronic edition of the Journal for a color version of this figure.*]

two coordinates for seven astrometric data points and 12 radial velocities. The nine parameters are the six orbital elements: the semimajor axis a, the eccentricity e, the longitude of the ascending node ω_0, the inclination i, the argument of periastron Ω_1, and the orbital period P, plus the RV offset V_0 and the semiamplitude of the RV curve K_1.

In the case of a conventional analysis of RV data, the main observable is the semiamplitude K_1, which once combined with the period and the eccentricity of the orbit, provides the mass function

$$f(M_1,\ M_2,\ i) = \frac{M_2^3 \sin i^3}{(M_T)^2}. \tag{16}$$

This combined analysis RV+astrometry allows us to separate the geometrical effects from the semiamplitude. This produces another composite observable, somewhat simpler than K_1, a "pseudoamplitude," whose formal expression is

$$A_1 = \frac{M_2}{M_T}\frac{2\pi a}{P}. \tag{17}$$

This is the parameter that will be used to determine the dynamical masses of the two components of the binary.

The final result of this nine-parameter fit, with 17 degrees of freedom, is represented in Figure 2 for the astrometry and in Figure 3 for the radial velocimetry. Our solution exhibits a final reduced $\chi^2_\nu = 1.03$, which is, despite the heterogeneity of the data, close enough to unity to ensure confidence in our estimation of the error bars. It is dominated by the velocimetry, despite having fewer measurements than the astrometry.

The confidence interval of each of the parameters is determined by analysis of the likelihood function. If we assume that the noise associated to our measurements is Gaussian, and that our parameters are independent, the likelihood can be approximated by

$$L(\text{parameters}) \propto \exp\left(-\frac{\chi^2}{2}\right). \tag{18}$$

The computation of a nine-dimensional likelihood function requires a lot of CPU time. This difficulty may be circumvented by confining the search to a subset of the total space. This is achieved by fixing the values of certain parameters and calculating the joined likelihood for the remaining parameters. However, one needs to check a posteriori that the hypothesis of independent parameters is valid.

This analysis reveals that among the nine parameters, only the argument of the periastron ω_0 and the longitude of the ascending node Ω_1 exhibit significant correlation. These two parameters are constrained within 0.5° (cf. Table 3). The hypothesis of independent parameters is valid, as expected from our good coverage in both radial velocimetry and astrometry. The uncertainty associated with each parameter is taken equal to the standard deviation of its associated likelihood function. The Gaussian-like likelihoods of the parameters used in the determination of the dynamical masses are shown in Figure 4.

4. CHARACTERISTICS OF GJ 623AB

The orbital parameters and the confidence intervals derived from the likelihood analysis are summarized in Table 3. This

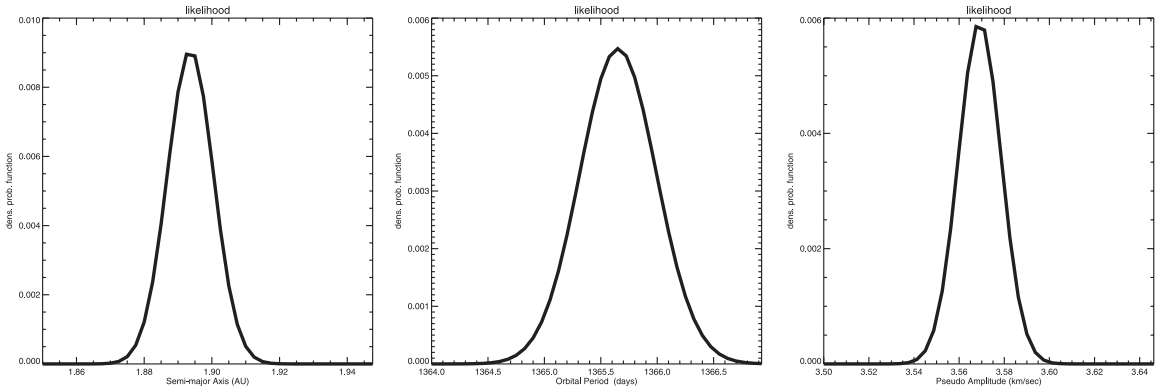


FIG. 4.—Likelihood functions of semimajor axis a, period P, and RV pseudoamplitude A_1, used to determine the dynamical masses of the GJ 623 system.

TABLE 4
DYNAMICAL MASSES

Quantity	Value ($M_\odot$)
Total mass	$M_T = 0.486 \pm 0.015$
Primary mass	$M_1 = 0.371 \pm 0.015$
Secondary mass	$M_2 = 0.115 \pm 0.0023$

combined analysis significantly improved the constraints on the epoch at periastron passage, from $T_P = 2451298 \pm 10$ (Nidever et al. 2002) to the new value $T_P = 2451313.3 \pm 0.6$. The rms of the RV fit is 56 m s^{-1} for a semiamplitude of $K_1 = 2.01 \pm 0.01$ km s^{-1}, which is comparable to the original fit (rms = 51 m s^{-1} for $K_1 = 2.08 \pm 0.04$ km s^{-1}).

4.1. *Dynamical Masses*

From the data summarized in Table 3, one can determine the dynamical masses of both components of the system. The revised parallax figuring in the table takes the binarity into account. The semimajor axis a, expressed in AU, once combined with the orbital period, gives the total mass M_T and the associated uncertainty σ_T:

$$M_T = a^3/P^2, \tag{19}$$

$$\sigma_T/M_T = \sqrt{(3\sigma_a/a)^2 + (2\sigma_P/P)^2}. \tag{20}$$

The likelihood analysis also determines the "pseudoamplitude" defined by equation (17) and its associated uncertainty.

This composite parameter, combined with the total mass, allows independent determination of the mass of the secondary M_2:

$$M_2 = M_T \frac{A_1 P}{2\pi a} = \frac{A_1 a^2}{2\pi P}, \tag{21}$$

$$\sigma_2/M_2 = \sqrt{(\sigma_{A_1}/A_1)^2 + (\sigma_P/P)^2 + (2\sigma_a/a)^2}, \tag{22}$$

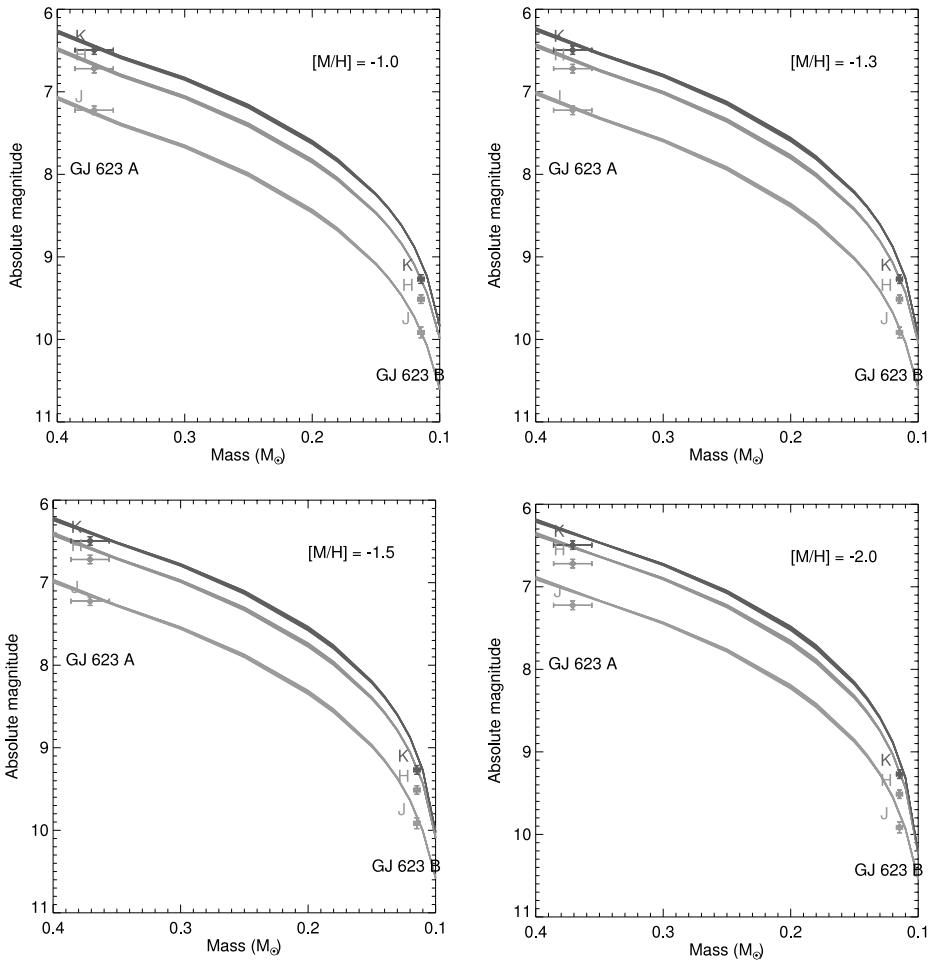


FIG. 5.—Mass-luminosity diagram for the GJ 623 system in *J*, *H*, and *K* bands. Comparison with the low-mass population II star models by Montalbán et al. (2000) in the 0.1–0.4 $M_\odot$ range, for different metallicities. [*See the electronic edition of the Journal for a color version of this figure.*]

and the mass ratio M_2/M_T:

$$M_R = \frac{m_2}{M_T} = \frac{A_1 P}{2\pi a}, \quad (23)$$

$$\sigma_R/M_R = \sqrt{(\sigma_{A_1}/A_1)^2 + (\sigma_P/P)^2 + (\sigma_a/a)^2}. \quad (24)$$

A similar analysis is possible for M_1. Errors on the masses are dominated by the uncertainty on the semimajor axis, itself dominated by the error on the *Hipparcos* parallax. As a consequence, with a fractional error of 1%, the mass ratio is better constrained than the mass of the secondary. The dynamical masses of the GJ 623 system are summarized in Table 4.

4.2. *Color, Metallicity, and Kinematics*

The multiwavelength contrast ratios found in Table 2 may be used to decompose the observed combined magnitude of the binary system into magnitudes for individual components. One determines the following absolute magnitudes: $M_J = 7.224 \pm 0.052$, $M_H = 6.719 \pm 0.051$, and $M_K = 6.495 \pm 0.052$ for GJ 623A and $M_J = 9.915 \pm 0.065$, $M_H = 9.512 \pm 0.052$, and $M_K = 9.269 \pm 0.053$ for GJ 623B.

With $J - K$ color indices of 0.729 ± 0.074 and 0.646 ± 0.084, respectively, for the primary and the secondary, the GJ 623AB system is bluer than the Delfosse et al. (2000) empirical *M-L* relation. Therefore, we suspect GJ 623 is of low metallicity. From the *Hipparcos* proper motion (1145.38, −452.37) mas yr^{-1} (Perryman et al. 1997) and the RV offset $V_0 = -27.4$ km s^{-1} determined from our model fit, one can calculate the Galactic space velocity $(U, V, W) = (-33, 14, -41)$ km s^{-1} after correction for standard solar motion.[5] This velocity is consistent with an old disk population, and therefore of subsolar metallicity (Chiba & Beers 2000), which is consistent with both components being bluer than the average field object.

[5] The sign convention is the one of the IDL astrolib `gal_uvw` procedures. Note that the literature is confusing on this matter. For instance, McCarthy & Henry (1987) provide numbers of comparable magnitude but with opposite sign for all (U, V, W) components.

Figure 5 compares the location of both components of the GJ 623 system in a mass-luminosity diagram to the low-mass population II models of Montalbán et al. (2000) for different metallicities. The trend we observe with these models (cf. Fig. 5) supports GJ 623AB being of subsolar metallicity. The model that best matches our measurements for the primary is for $[\mathrm{M/H}] = -1.0$. For the secondary, the best model predicts a slightly lower mass of $0.110 \pm 0.001\ M_\odot$ for $[\mathrm{M/H}] = -2.0$. This large discrepancy in metallicity is inconsistent with the assumption of a coeval binary. The very low metallicity for the secondary would be consistent with GJ 623B belonging to the Galactic halo, which is unlikely according to the kinematics. Therefore, we conclude that the models do not adequately fit the data.

5. CONCLUSION

The observation of binary systems is the only way to measure unbiased masses. As shown in this paper, combined with excellent radial velocity measurements, a few high–angular resolution images provide sufficient information to constrain the range of possible masses below the 5% precision that is required to seriously challenge the models at the low end of the main sequence.

The application of precision RV methods developed for planet searches combined with AO will provide a complete characterization of the stellar structure of the lower main sequence.

We thank the staff and telescope operators of Palomar Observatory and Keck Observatory for their support. F. M. thanks Terry Herter for his help. This work is partially funded by the National Science Foundation under grants AST 03-35695 and AST 05-06588. This publication makes use of the Simbad database, operated at CDS, Strasbourg, France and the data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. We wish to extend special thanks to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests. Without their generous hospitality, the observations presented herein would not have been possible.

REFERENCES

Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Nature, 320, 595
Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403
Barbieri, C., Demarchi, G., Nota, A., Corrain, G., Hack, W., Ragazzoni, R., & Macchetto, D. 1996, A&A, 315, 418
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Delfosse, X., Forveille, T., Ségransan, D., Beuzit, J.-L., Udry, S., Perrier, C., & Mayor, M. 2000, A&A, 364, 217
Delfosse, X., et al. 2004, in ASP Conf. Ser. 318, Spectroscopically and Spatially Resolving the Components of the Close Binary Stars, ed. R. W. Hilditch, H. Hensberge, & K. Pavlovski (San Francisco: ASP), 166
Eisner, J. A., & Kulkarni, S. R. 2002, ApJ, 574, 426
Hayward, T. L., Brandl, B., Pirger, B., Blacken, C., Gull, G. E., Schoenwald, J., & Houck, J. R. 2001, PASP, 113, 105
Henry, T. J. 1998, in ASP Conf. Ser. 134, Brown Dwarfs and Extrasolar Planets, ed. R. Rebolo, E. L. Martin, & M. R. Zapatero Osorio (San Francisco: ASP), 28
Henry, T. J., Franz, O. G., Wasserman, L. H., Benedict, G. F., Shelus, P. J., Ianna, P. A., Kirkpatrick, J. D., & McCarthy, D. W., Jr. 1999, ApJ, 512, 864
Jancart, S., Jorissen, A., Babusiaux, C., & Pourbaix, D. 2005, A&A, 442, 365
Lippincott, L. S., & Borgman, E. R. 1978, PASP, 90, 226
Lloyd, J. P., Martinache, F., Ireland, M. J., Monnier, J. D., Pravdo, S. H., Shaklan, S. B., & Tuthill, P. G. 2006, ApJ, 650, 131
Marcy, G. W., & Moore, D. 1989, ApJ, 341, 961
McCarthy, D. W., Jr., & Henry, T. J. 1987, ApJ, 319, L93
Metchev, S. A., & Hillenbrand, L. A. 2004, ApJ, 617, 1330
Montalbán, J., D'Antona, F., & Mazzitelli, I. 2000, A&A, 360, 935
Nakajima, T., Kulkarni, S. R., Gorham, P. W., Ghez, A. M., Neugebauer, G., Oke, J. B., Prince, T. A., & Readhead, A. C. S. 1989, AJ, 97, 1510
Nidever, D. L., Marcy, G. W., Butler, R. P., Fischer, D. A., & Vogt, S. S. 2002, ApJS, 141, 503
Perryman, M. A. C., et al. 1997, A&A, 323, L49
Pourbaix, D., et al. 2004, A&A, 424, 727
Pravdo, S. H., Shaklan, S. B., Wiktorowicz, S. J., Kulkarni, S., Lloyd, J. P., Martinache, F., Tuthill, P. G., & Ireland, M. J. 2006, ApJ, 649, 389
Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, PASP, 111, 587
Readhead, A. C. S., Nakajima, T. S., Pearson, T. J., Neugebauer, G., Oke, J. B., & Sargent, W. L. W. 1988, AJ, 95, 1278
Ségransan, D., Delfosse, X., Forveille, T., Beuzit, J.-L., Udry, S., Perrier, C., & Mayor, M. 2000, A&A, 364, 665
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555
van Leeuwen, F., & Evans, D. W. 1998, A&AS, 130, 157

The Astrophysical Journal, 695:1183–1190, 2009 April 20 doi:10.1088/0004-637X/695/2/1183
© 2009. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

VISUAL ORBIT OF THE LOW-MASS BINARY GJ 164 AB

Frantz Martinache[1,2], Barbara Rojas-Ayala[1], Michael J. Ireland[3], James P. Lloyd[1], and Peter G. Tuthill[3]
[1] Department of Astronomy, Cornell University, Ithaca, NY, USA
[2] Subaru Telescope, Hilo, HI, USA
[3] School of Physics, University of Sydney, Sydney, NSW, Australia
Received 2008 September 4; accepted 2009 January 30; published 2009 April 6

ABSTRACT

We report seven successful observations of the astrometric binary GJ 164 AB system with aperture masking interferometry. The companion, with a near-infrared contrast of 5:1 was detected beyond the formal diffraction limit. Combined with astrometric observations from the literature, these observations fix the parallax of the system, and allow a model-independent mass determination of both components. We find the mass of GJ 164 B to be $0.096 \pm 0.008\ M_{\odot}$. With data taken with the newly commissioned TripleSpec spectrograph, we examine GJ 164's near-infrared spectral features and compare them to a sample of M dwarfs of known metallicity. Models are not consistent with color and mass, requiring a very young age to accommodate a secondary too luminous, a scenario ruled out by the kinematics.

Key words: stars: low-mass, brown dwarfs – techniques: high angular resolution – techniques: spectroscopic

1. INTRODUCTION

The development of a self-consistent theory of the internal structure of stars and the construction of models which trace evolutionary behavior are major achievements of modern astrophysics. Both theory and models have however long been centered on intermediate- and high-mass stars, with little consideration of objects with masses below $0.6\ M_{\odot}$. The realization that the solar neighborhood is overwhelmingly dominated by low-mass stars (van de Kamp 1971; Henry 1998), and the discovery of brown dwarfs (Nakajima et al. 1995) have brought a lot more attention to the lower end of the main sequence and what lies below.

The complex physical nature of these very low mass objects makes their modeling intricate. Boundaries between stars, brown dwarfs, and planets, as well as theoretical relations that predict the luminosity and temperature as functions of age, mass, and metallicity, are still largely untested in the relevant range of parameters. Correct calibration of these models is of extreme importance, since the same self-consistent theory is used to estimate the mass of brown dwarfs (Burrows et al. 2006) and giant exoplanets (Baraffe et al. 2003) from their luminosity and age.

The observations required to challenge and improve the models are dynamical mass measurements of multiple star systems, combined with accurate photometry and distance determination. Besides gravitational microlensing (Paczynski 1986), the observation of binary systems and the use of Newtonian orbital dynamics provide the only method of directly measuring accurate stellar masses. With maximum sensitivity at separations greater than 1 arcsec, conventional Adaptive Optics (AO) observations at Palomar are limited to the most nearby objects or the consequent orbital periods become too long to lead to dynamical masses. We have therefore been using aperture masking interferometry (Tuthill et al. 2000) in conjunction with AO (Lloyd et al. 2006). The precision calibration of the data achieved with this observing mode indeed leads to reliable results up to and beyond the formal diffraction limit (super resolution), and very precise photometry. Combined with radial velocity (Martinache et al. 2007) or astrometry (Ireland et al. 2008), aperture masking interferometry has so far provided some of the most precise (dynamical) masses of objects below $0.1\ M_{\odot}$.

2. CHARACTERIZATION OF THE ORBIT

We report here aperture masking interferometry observations of the astrometric binary GJ 164 AB. GJ 164 (aka LHS 1642, G 175-19 or Ross 28) is a high proper-motion star, cataloged as an M4.5 dwarf (Reid et al. 1995), 11–13 pc distant. From the Two Micron All Sky Survey (2MASS) catalog, its apparent near-infrared magnitudes are $J = 8.773 \pm 0.032$, $H = 8.248 \pm 0.030$, and $K = 7.915 \pm 0.016$. The discovery of GJ 164 B was reported by Pravdo et al. (2004), as a part of the STEPS program of Pravdo & Shaklan (1996). The combination of the astrometry data from the discovery paper, combined with the aperture masking interferometry data reported in this paper provides a complete dynamical characterization of the system, resulting in masses that are independent of the use of a model or a mass–luminosity relation (MLR). GJ 164 is an especially interesting target for the mass of the B component lies very close to the substellar limit.

2.1. Aperture Masking Interferometry

Measuring the dynamical masses of a low-mass binary system requires some patience, for the observations have to cover at the very least a significant fraction of the orbital period. Because of its low mass (less than $0.6\ M_{\odot}$), a relatively short period (less than 10 yr) binary will have a semimajor axis smaller than 4 AU. Even for objects as close as 10 pc, the angular separation will not exceed 0.4 arcsec, a performance theoretically within the grasp of a 5–10 m class telescope equipped with AO. Yet, in practice, the residual quasi-static speckles inthe AO point-spread function (PSF) halo seriously limit the capability of detecting a companion at angular separation smaller than $2\text{–}4\lambda/D$ (Racine et al. 1999). Aperture masking interferometry with AO addresses this issue.

A nonredundant mask discards most of the pupil by sampling a few spatial frequencies only (Readhead et al. 1988; Nakajima et al. 1989). One admittedly loses most of the light (the 9-hole mask used for this works transmits approximately 15%) but rejects all atmospheric noise, as well as internal aberrations and noncommon path errors, which in turn, dramatically increases the signal-to-noise ratio (Tuthill et al. 2000). Used in conjunction, aperture masking interferometry and AO provide stable fringes, and enable long integration times, therefore making

Table 1
Aperture Masking Measurements at Palomar and Keck: Angular Separation and Position Angle of GJ 164 B

Julian Date (−2,450,000)	Band	Telescope	Sep. (mas)	PA (degrees)
3004.85	H	Palomar	82.8 ± 1.6	111.1 ± 0.9
3632.96	H	Palomar	53.3 ± 1.3	149.1 ± 1.2
3723.66	H	Palomar	80.1 ± 3.3	116.1 ± 2.1
3779.71	H, K_S	Palomar	86.1 ± 2.0	102.1 ± 0.9
3957.62	H	Keck	49.3 ± 0.5	47.0 ± 0.6
4018.98	K_S	Palomar	39.4 ± 1.3	344.8 ± 2.6
4078.79	K_S	Palomar	57.3 ± 2.2	294.1 ± 2.8

Table 2
GJ 164 Photometry

Julian Date (−2,450,000)	Filter	Contrast
3004.85	H	5.97 ± 0.32
3632.96	H	5.77 ± 0.46
3723.67	H	5.78 ± 0.63
3779.71	H	5.53 ± 0.60
	K_S	4.89 ± 0.19
3957.62	H	5.42 ± 0.03
4018.98	K_S	8.66 ± 4.81
	H	6.38 ± 1.65
4078.79	K_S	5.67 ± 1.17

faint target accessible. The only limit is whether the AO system can lock on the system. Thus, any target that can be observed with AO can also be observed with aperture masking interferometry. With a visible magnitude $V = 13.5$ (Weis 1996), GJ 164 can fairly easily be observed with aperture masking at Palomar and Keck.

The PSF of a nonobscurated telescope of diameter D is described by the radial Airy function, and a rule of thumb known as the Rayleigh criterion states that two sources are resolved if separated by at least $1.22\lambda/D$ (the first zero of the Airy function). One obvious advantage of aperture masking interferometry over standard direct imaging is that in masking the entire pupil except for two diametrically opposed holes (therefore separated by the distance D), the system produces fringes of period λ/D. Two sources are now resolved if separated by more than $0.5\lambda/D$, a condition also known as the Michelson criterion: masking the telescope offers a gain of 2.44 in resolution in a very straightforward manner.

Although useful for comparison, these criteria are however somewhat arbitrary and in practice, good calibration of the data in long baseline interferometry permits the resolution of structures smaller than $0.5\lambda/D$. See, for instance, Figure 1 of Kervella et al. (2008) that shows structures smaller than 2 mas being resolved by a ∼30 m baseline B for $\lambda = 2\mu$m, i.e., $\lesssim 0.15\lambda/B$.

The data presented in Tables 1 and 2 show that the companion to GJ 164 A is detected on multiple occasions beyond the formal diffraction limit, in what is sometimes called the super-resolution regime. Tuthill et al. (2006) provide a recent description of the general principles and performances of aperture masking interferometry, while Lloyd et al. (2006) detail the experiment undertaken at Palomar, and Kraus et al. (2008) the aperture masking mode of NIRC2.

The AO observations performed with PHARO (Hayward et al. 2001) at Palomar as well as with NIRC2 at Keck II span the range 2003 December to 2006 December. The companion was successfully detected all seven runs. The data are calibrated with observations of one or several stars of similar brightness for both the AO wave front sensor and the science camera to ensure comparable wave front correction and signal-to-noise ratio. The final data product of our custom software pipeline written in IDL, is a collectionof calibrated closure phases (Baldwin et al. 1986). This very robust observable rejects both atmospheric noise and calibration errors of the wave front sensor.

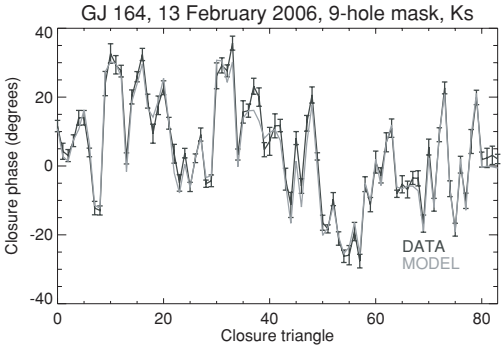


Figure 1. Calibrated closure phases observed on JD 2,453,779.71 with a K_S filter. A binary star model (light gray) with an 86 mas separation, PA of 102°, and a contrast of 5:1 satisfactorily matches the observations (dark gray). For an unresolved system, each closure phase would be either 0° or 180° with the same confidence intervals.

Each closure phase is a measure of the degree of asymmetry of the target for a given position angle (PA). Thus, a symmetric object like a point source or a near-equal binary exhibits 0° or 180° closure phase (Monnier 2000), while a nonequal binary exhibits closure phase of intermediate value. The 9-hole mask provides the simultaneous measurement of $\binom{9}{3} = 84$ possible closure triangles, $\binom{8}{2} = 28$ of which are theoretically independent. Figure 1 represents one such set of closure phases observed with the 9-hole mask in the K band.

The instantaneous geometry of a binary star is modeled by a set of three parameters: contrast ratio c, angular separation a, and PA θ. The example of one such model fitted to a set of closure phases is given in Figure 1. The aperture masking data processing software typically uses several hundreds of frames which are averaged to produce the closure phases. Due to the central limit theorem, the distribution of closure phases $\{\phi_i\}$ can be considered Gaussian, with uncertainties σ_i given by the standard deviation to the mean. The best fit minimizes the traditionally used goodness-of-fit parameter χ^2 defined as

$$\chi^2 \equiv \sum_i \frac{(\phi_i - f_i(a,\theta,c))^2}{\sigma_i^2}, \quad (1)$$

where f_i designates the model. A systematic error term is added to the closure phase dispersion σ_i to achieve a reduced $\chi^2_\nu = 1$. The likelihood of the parameters given the set of closure phases $\{\phi_j\}$ is related to the χ^2 by

$$L(a,\theta,c|\{\phi_j\}) \propto \exp(-\chi^2/2). \quad (2)$$

Normalized, L is the full joint probability density function for all three parameters. The confidence interval of one individual parameter is calculated by integrating out the two others. For example, the marginal probability density function for the

angular separation is calculated by integrating out the PA and the contrast ratio:

$$p(a) = \int d\theta\, dc\, L(a, \theta, c). \quad (3)$$

While the PA is naturally well orthogonal to the angular separation and the contrast ratio, the extraction of the other two parameters is a sensitive function of the angular separation. For large separation ($\sim\lambda/D$ and beyond), one sees (see Tables 1 and 2) that the contrast is better constrained than for the detections at smaller separations. Pravdo et al. (2006) show that for such detections at small angular separations, there is degeneracy between contrast and separation (see Figure 7 of the quoted paper). One can take advantage of the better-resolved observations for which the contrast is well constrained (JD 2,453,957.62 for the H band and JD 2,453,779.71 for the K_S band) and adopt these contrasts for the other observations, to get a better constraint on the angular separation. The results of this analysis: relative position and photometry are gathered in Tables 1 and 2. This and the more or less favorable seeing conditions explain the variable confidence interval sizes associated with each run. Overall, this analysis achieves a precision $\gtrsim$1 mas at separations as low as $\sim$40 mas in the K_S band ($0.5\lambda/D$) on good nights at Palomar.

2.2. Astrometry

GJ 164 was discovered as a binary in an astrometric survey by Pravdo & Shaklan (1996) and Pravdo et al. (2004). While the interferometry resolves the binary and provides its instant geometric configuration, the astrometry records the position of a star's photocenter, measured relative to several more distant stars. Repeated observations reveal the star's proper motion as well as its parallax. Once these two effects are subtracted from the astrometric signal, any residual wobble of the photocenter can be attributed to the presence of one or more companions, massive enough to cause an observable shift in position. The orbit of the companions can then be fully characterized.

Astrometric data were extracted from the figures of Pravdo et al. (2004) then added the motion due to the proper motion and the parallax given in the paper's tables, while the corresponding Julian dates were taken from the telescope observing schedule and therefore have a $\pm$2 day uncertainty. These extracted values are gathered in Table 3. We initially assume a 2.0 mas uncertainty for each individual measurement, as suggested by the residuals of the analysis performed by Pravdo et al. (2004).

The astrometric residual wobble however remains a function of the fractional light of the secondary, $\beta = L_2/(L_1 + L_2)$, with L_1 and L_2 respectively denoting the luminosity of the primary and the secondary, as well as the fractional mass of the secondary $f = m_2/(m_1 + m_2)$. In the limit at which the light of the secondary is negligible, the photocentric orbit is identical to the Keplerian orbit of the primary around the actual center of mass of the system. If the luminosity of the secondary is not so negligible, then the photocentric semimajor axis α is reduced proportionally to β. The ratio of the photocentric orbit α, and the Keplerian orbit a is (McCarthy et al. 1988)

$$\alpha/a = f - \beta. \quad (4)$$

A relatively small-mass, nonluminous secondary, and a relatively large-mass, luminous secondary will therefore have identical astrometric signatures. The data presented in Section 2.1 however rule out the latter possibility. Indeed, the aperture masking interferometry data gathered in Table 2 provide the H and K_S-band contrast ratios, which are turned into the following differences of magnitude:

$$\Delta H = 1.835 \pm 0.006$$
$$\Delta K_S = 1.721 \pm 0.097.$$

These contrast ratios can be used to decompose the 2MASS combined magnitudes of the binary into apparent magnitudes for individual components: $H = 8.432 \pm 0.030$ and $K = 8.117 \pm 0.017$ for GJ 164 A, $H = 10.267 \pm 0.030$ and $K = 9.834 \pm 0.038$ for GJ 164 B.

The corresponding $(H - K)$ color indices: $(H - K)_A = 0.315 \pm 0.034$ and $(H - K)_B = 0.433 \pm 0.048$ alone, provide a parallax-independent observable. While the $(H - K)$ color index of the primary is well compatible with an M4.5 spectral type (Henry et al. 1994; Koornneef 1983), the secondary is very unlikely to be any earlier than M8. In fact, according to Table 5 of Kirkpatrick et al. (2000), GJ 164 B qualifies as an M8.5 dwarf. Liebert & Gizis (2006) provide RI photometry for a sample of late M and L dwarfs. Out of their sample, we extract five M8.5 dwarfs, and can therefore estimate $\langle R - K_S \rangle = 6.41 \pm 0.26$ for this spectral type. Combined with the apparent K_S magnitude determined here, this provides an estimate of $R = 16.24 \pm 0.26$ for GJ 164 B, while $R = 12.4$ for GJ 164 A. This leads to an R-band contrast ratio $c = 33.5$, and to a fractional light of the secondary $\beta = 0.029 \pm 0.007$.

Table 3
GJ 164 Astrometry

Julian Date (−2,450,000)	ΔR.A. (mas)	ΔDecl. (mas)
0801.5	0.00	0.00
0801.5	0.00	−2.80
1088.3	−147.91	−646.96
1189.0	−328.82	−874.15
1189.0	−324.82	−874.05
1189.0	−325.32	−871.65
1436.5	−457.34	−1409.82
1494.0	−563.37	−1519.60
2657.5	−1639.04	−4132.08
2889.5	−1745.18	−4629.59
2889.5	−1745.18	−4627.29
2889.5	−1745.88	−4625.69
3030.5	−2014.31	−4953.05
3030.5	−2014.61	−4957.45
3051.5	−2044.87	−5016.85

3. CHARACTERISTICS OF GJ 164 AB

3.1. Orbital Parameters

The aperture masking observations gathered in Table 1 provide the instant position of the companion to the primary. These observations alone suffice to constrain all seven orbital elements of the binary, independently of the parallax, with the semimajor axis expressed in angular units.

We use a model identical to the one described in Section 3.1.2 of Martinache et al. (2007) to fit a 7-parameter model to 14 observables. The observables are two coordinates for seven aperture masking data points. The seven orbital parameters are: the angular semimajor axis α, the eccentricity e, the longitude of the ascending node ω_0, the inclination i, the argument of the periastron Ω_1, the period P, and the epoch at periastron passage T_P. Our solution exhibits a final reduced $\chi^2_\nu = 0.61$

Table 4
Orbital Elements

Parameter	AO Only	AO + STEPS
α (mas)	80.4 ± 1.3	80.5 ± 1.2
e	0.157 ± 0.012	0.161 ± 0.012
i (deg)	121.9 ± 0.9	121.9 ± 0.8
Ω_1 (deg)	272.4 ± 1.3	271.8 ± 1.2
ω_o (deg)	314.7 ± 5.0	311.7 ± 3.7
P (days)	734.3 ± 4.6	736.9 ± 1.7
T_P (reduced JD)	1868 ± 19	1856 ± 8

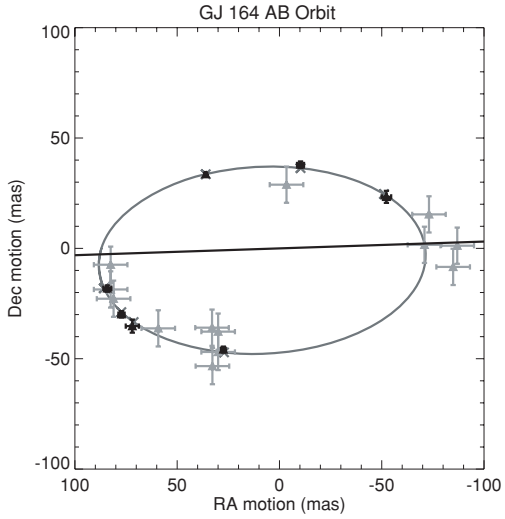


Figure 2. GJ 164 B ORBIT. The measurements and associated uncertainties of the aperture masking data are represented by dark gray rectangles. Light gray triangles represent the STEPS astrometry data. The relative apparent orbit (curve) is clockwise. Table 4 provides the corresponding orbital elements. The almost horizontal straight line represents the line of nodes.

for 7 degrees of freedom. Note that the temporal coverage of these observations definitely rules out the possibility of a ~4 yr orbital period, which although less likely, could not be excluded by Pravdo et al. (2004). The seven orbital parameters and the associated uncertainties are presented in Table 4.

We then proceed to a joined fit AO + astrometry, that is a 13-parameter fit to 40 observables. The extra parameters included in this new analysis are: the two components of the proper motion, the relative parallax Π as well as the semimajor axis of the motion of the primary around the center of mass of the system a_1. With a 2 mas uncertainty on the individual astrometric measurements, as assumed in Pravdo et al. (2004), the solution to this global fit exhibits a final reduced $\chi^2_\nu = 1.04$ for 27 degrees of freedom. Table 4 compares the seven orbital parameters and their uncertainties obtained from the fit of the AO data only and the joined fit: both approaches produce compatible solutions. Figure 2 represents the orbit resulting from this joined fit, onto which are superposed the AO and astrometry data. Table 5 summarizes the astrometric elements re-derived from the joined fit.

The probability that a set of parameters is contained within a volume of the parameter space is proportional to the likelihood function, itself proportional to $\exp(-\chi^2/2)$ (see Section 2.1). In an n-parameter fit, the marginal distribution for one parameter is calculated by integrating the likelihood over the $n-1$ other parameters. Tables 4 and 5 provide the 1σ confidence

Table 5
Astrometric Elements

Parameter	Value
Proper motion R.A. (mas)	-324.5 ± 0.3
Proper motion decl. (mas)	-808.3 ± 0.3
Relative parallax (mas)	70.0 ± 0.8
Photocentric semimajor axis (mas)	20.1 ± 0.8
Total mass ($M_\odot$)	0.343 ± 0.026

Table 6
Dynamical Masses: Obtained from the Joined Analysis AO+Steps

Parameter	Value
Total mass	$M_T = 0.343 \pm 0.026\ M_\odot$
Primary mass	$M_1 = 0.247 \pm 0.019\ M_\odot$
Secondary mass	$M_2 = 0.096 \pm 0.008\ M_\odot$
Mass ratio	$M_2/M_T = 0.279 \pm 0.012$

intervals for each parameter calculated this way. The method of Lampton et al. (1976) used by Pravdo et al. (2004) to calculate uncertainties however provides confidence intervals for the joined distribution, i.e., finds the hypercontour in the parameter space such that there is a 68.3% probability (i.e., 1σ for a normal distribution) of having simultaneously all parameters within that hypercontour. The approach used in this work provides a significantly smaller confidence interval for each parameter (compare Table 5 with Table 1 of Pravdo et al. 2004). In the determination of uncertainties on the dynamical masses (see Section 3.2), this proves essential since the masses are only a function of subset of all the 13 parameters the joined fit gives.

3.2. Dynamical Masses

Like Pravdo et al. (2004), we use a 2 ± 1 mas correction to convert the relative parallax $\Pi = 70.0 \pm 0.8$ mas, to the absolute parallax: $\Pi = 72.0 \pm 1.2$ mas. We can now determine the actual relative semimajor axis of the binary: $a = 1.1 \pm 0.2$ AU, as well as the total mass:

$$M_T = a^3/P^2, \quad (5)$$

that is, $M_T = 0.343 \pm 0.026\ M_\odot$. In the limiting case where the luminosity of the secondary is negligible, the fractional mass of the secondary (see Equation (4)) is given directly by the ratio of the photocentric and relative orbit semimajor axis, that is $\alpha/a = 0.250 \pm 0.010$. Section 2.2 however shows the secondary is not negligible. The finite R-band contrast ratio translates into a fractional light $\beta = 0.029 \pm 0.007$, which leads to the fractional mass $f = 0.279 \pm 0.012$. Note that a remarkable feature of the fractional mass is that it is absolutely parallax-independent.

From the total mass and the fractional mass, one infers the masses of both components: $M_1 = 0.247 \pm 0.019$ and $M_2 = 0.096 \pm 0.008$. All dynamical characteristics of the GJ 164 system are gathered in Table 6.

To check a posteriori the validity of the uncertainties on the masses, we have conducted a series of Monte Carlo simulations. For each simulation, a new set of data (aperture masking and astrometry) was generated using a normal distribution centered on the original data points with standard deviation the associated uncertainties. The model-fit procedure highlighted in Section 3.1 was then used to determine the corresponding orbital and astrometric elements, which lead to the dynamical masses M_1 and M_2. Figure 3 presents the histograms of M_1

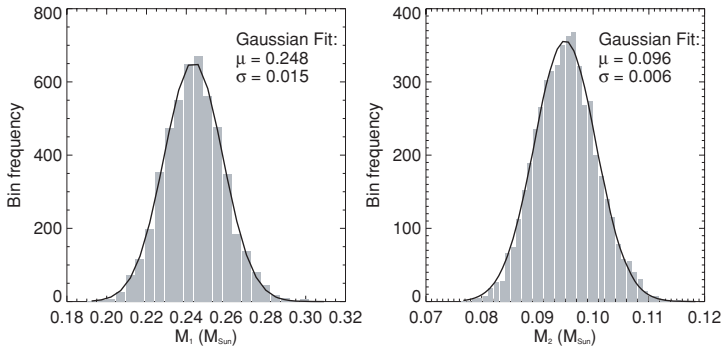


Figure 3. Histograms for M_1 and M_2 (gray bars) were calculated using a series of 5000 Monte Carlo experiments as described in Section 3.2. Each distribution is fitted with a Gaussian (solid black line) whose parameters (mean μ and standard deviation σ) are also printed.

and M_2 for 5000 such Monte Carlo simulations, as well as the mean and standard deviation of the corresponding distribution. With this analysis, we find that $M_1 = 0.248 \pm 0.015$ and $M_2 = 0.096 \pm 0.006$, which is in excellent agreement with the results and associated uncertainties quoted in Table 6.

4. SPECTROSCOPY

In this section, we present the *J*- and *K*-band spectra of GJ 164 AB. A careful characterization of the atmosphere and determination of abundances is necessary to understand the location of GJ 164 AB in a mass–luminosity (M/L) diagram, relative to an observational M/L relation or to numerical simulations of M-dwarf atmospheres. The spectrum and absorption features of GJ 164 AB were compared to the ones of other five M4–M5 dwarfs of known metallicity, aiming to determinate GJ 164 AB overall abundance.

4.1. Observations and Data Reduction

Our spectroscopic sample consists of five M4–M5 dwarfs (in complement to GJ 164 AB) associated with solar-type stars, whose metallicity was measured by Valenti & Fischer (2005). Assuming that both components formed from the same original molecular cloud, they should share the same metallicity. Their designation, spectral type, $(J-K)$ colors, as well as metallicities are listed in Table 7.

Near-infrared spectra of these objects were obtained with the recently commissioned TripleSpec spectrograph on the Palomar Hale Telescope (Wilson et al. 2004; Herter et al. 2008). TripleSpec at Palomar is a $\lambda/\Delta\lambda \approx 2500$–$2700$ cross-dispersed near-infrared spectrograph with a broad wavelength coverage across five simultaneous orders (1.0–2.4 μm) in echelle format. Its entrance slit is 1×30 arcsec and a notable feature is that it has no moving parts.

The data were reduced with an IDL-based data reduction pipeline developed by P. Muirhead for TripleSpec at Palomar. The data were sky-subtracted using a sky-frame made by median combination of the five exposures on different positions of the object along the slit. To correct for telluric absorption features and flux-calibrate their spectra, an A0 V star was observed as close to the science object star airmass as possible. Each sky-subtracted exposure was then divided by a normalized flat-field, wavelength calibrated and fully extracted. The spectra are flux-calibrated and telluric corrected using the IDL-based code `xtellcor_general` described in the paper by Vacca et al. (2003).

Table 7
The Spectroscopic Sample

Name	Sp. Typ.	$J-K$ (mag)	$[M/H]$ (dex)
GJ 324 B	M4	0.894	$+0.25 \pm 0.03$
GJ 544 B	M4[1]	0.894	-0.15 ± 0.03
GJ 611 B	M4	0.744	-0.49 ± 0.03
GJ 166 C	M4.5	0.785	-0.08 ± 0.03
NLTT 25869	M5	0.872	$+0.27 \pm 0.03$

GJ 544 B is listed in SIMBAD as an M6.

4.2. The Near-Infrared Spectrum of GJ 164

The final *J*- and *K*-band spectra of GJ 164 AB and the five M4–M5 dwarfs of known metallicity are shown in Figures 4 and 5, respectively. The equivalent width (EW) of prominent features were estimated using the IDL-based function `measure_ew` by N. Konidaris and J. Harker. Each spectrum was normalized and the EWs were calculated by computing the ratio of the area of a feature to a pseudocontinuum. This pseudocontinuum is a linear interpolation of the "clean" regions on either side of the feature. Uncertainties in the EW were obtained from the errors in the estimated pseudocontinuum and the errors in the measured intensities of each feature, as described in Sembach & Savage (1992).

The EW of seven prominent features in *J* band, shown in Figure 4, as well as Na I (2.206 μm and 2.209 μm), Ca I (2.261 μm, 2.263 μm, and 2.265 μm), and ^{12}CO(2,0) (2.293 μm) in *K* band are shown in Figure 6. Open circles depict the stars with $[M/H] < [M/H]_\odot$ (GJ 544 B, GJ 611 B, and GJ 166 C) and filled circles depict the stars with $[M/H] > [M/H]_\odot$ (GJ 324 B and NLTT 25869). A star represents GJ 164 AB.

The EW of the features in the *K* band of GJ 164 AB (see Figure 6) seems to be compatible with the values of the stars with solar or above-solar metallicity in our sample. Its *K*-band spectrum differs appreciably from the one of GJ 611 (the lowest metallicity object in the sample), where all absorption features appear weaker compared with the rest of the sample, an expected manifestation of its low abundance nature. However, most of the features present in the spectra of these stars, especially in the *J* band, are extremely temperature-dependent (Ali et al. 1995; Jones et al. 1996; except from the Fe I lines that do not exhibit any correlation with spectral type; Cushing et al. 2005), and prone to non-LTE effects, which make it difficult to assert unambiguously whether or not GJ 164 AB is a metal-rich star.

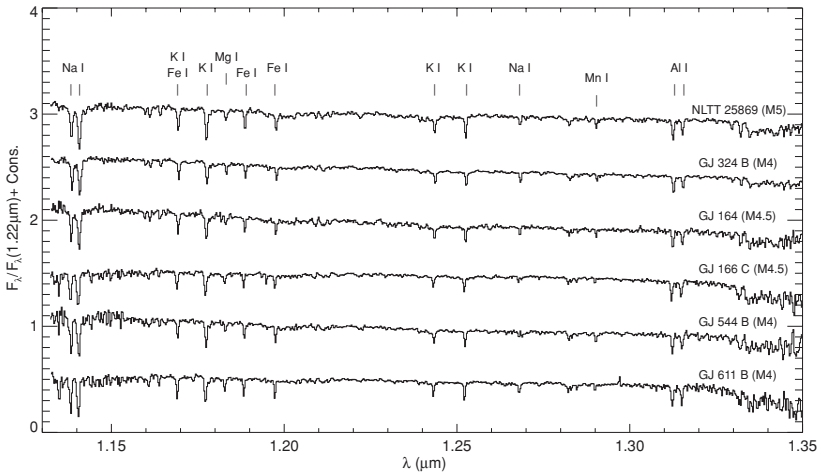


Figure 4. *J*-band spectra of the stars ordered from top to bottom by decreasing metallicity (listed in Table 7). The most prominent spectral features are highlighted.

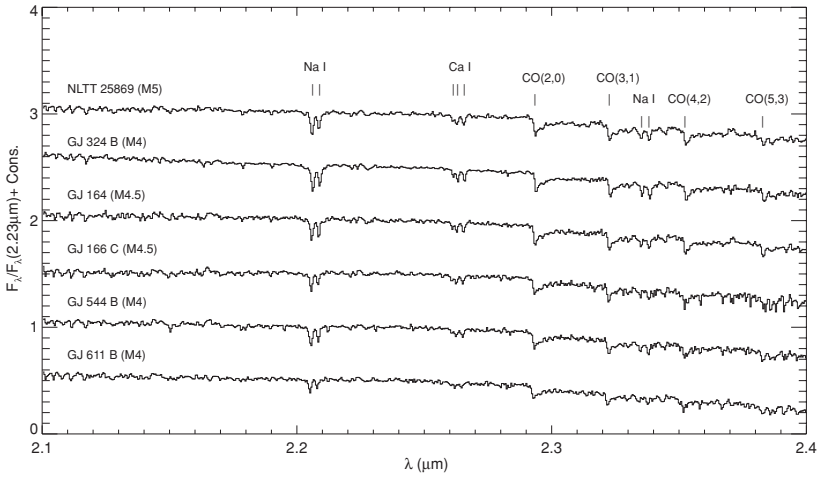


Figure 5. *K*-band spectra of the stars ordered from top to bottom by decreasing metallicity (listed in Table 7). The most prominent spectral features are highlighted.

5. MASS–LUMINOSITY RELATIONS

The individual apparent magnitudes calculated in Section 2.2, combined with the absolute parallax deduced from the dynamical analysis performed in Section 3.2, determine the following absolute H and K magnitudes: $M_H = 7.719 \pm 0.047$ and $M_K = 7.404 \pm 0.040$ for GJ 164 A, $M_H = 9.554 \pm 0.047$ and $M_K = 9.121 \pm 0.052$ for GJ 164 B.

The $(H - K)$ colors, and to a lesser extent the infrared spectroscopy data presented in Section 4, indicate that GJ 164 is unlikely to be metal-poor. Figure 7 compares the location of GJ 164 A and B in a mass–luminosity diagram, relative to the low-mass solar metallicity models of Baraffe et al. (1998, hereafter referred to as BCAH98) for the K band. For masses below 0.4 $M_\odot$ (but above the substellar limit), models predict very little evolution between 0.5 and 5 Gyr. Figure 7 therefore only plots two isochrones, for 0.1 and 1 Gyr.

The aperture masking data places the primary less than 1σ away from the 1 Gyr isochrone, and 2σ away from the 0.1 Gyr isochrone. The hypothesis of an evolved age is supported by the kinematics of the system. Indeed, with the proper motion $(-324.5, -808.4)$ mas yr^{-1}, the absolute parallax $\Pi = 72 \pm 1.2$ mas derived in Section 3.2, as well as the radial velocity $V = -29.9$ km s^{-1} measured by Pravdo et al. (2004), one can calculate the Galactic space velocity $(U, V, W) = (-24, -20, -47)$ km s^{-1} after correction for standard solar motion.[4] When compared to the properties of the nearest young moving groups (see Table 1 of López-Santiago et al. 2006), GJ 164 exhibits a velocity component perpendicular to the Galactic plane too large to be associated with any of these groups. From the kinematic properties of the M-Dwarf Survey of Reid et al. (2002), it is however impossible to conclude whether GJ 164 belongs to the thin or the thick disk. Figure 7 also plots the empirical K-band MLR established by Delfosse et al. (2000) for low-mass field stars, valid for $4.5 < M_K < 9.5$. Here again, the location of GJ 164 A seems in fairly good agreement with the MLR.

[4] The sign convention is the one of the IDL astrolib gal_uvw procedure, with U positive toward the Galactic anti-center, V positive in the direction of Galactic rotation, and W positive toward the North Galactic Pole.

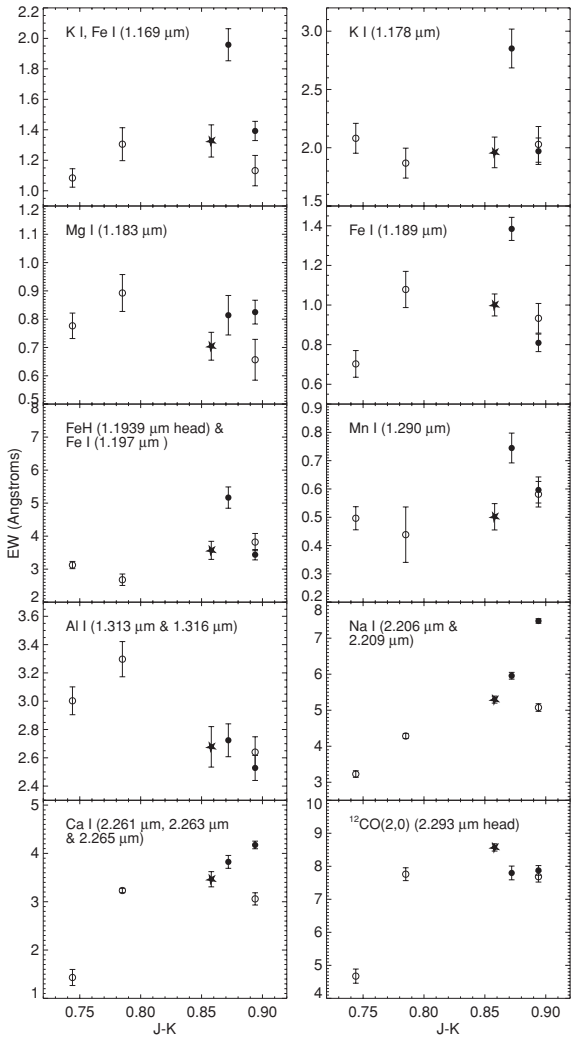


Figure 6. EWs of some of the most prominent lines in the sample spectra with 1σ errors vs. $(J - K)$ colors. GJ 164 is represented by the five branch star, while metal-rich objects ($[M/H] > 0.0$) are represented by filled circles and metal-poor objects ($[M/H] < 0.0$) by open circles. The high value of GJ 164's ^{12}CO(2,0) (2.293 μm) EW may be attributed to its secondary. The 5:1 K-band contrast ratio of the binary and the increasing strength of the CO bands with later types in the M-dwarf sequence, explain how GJ 164 B (unresolved with TripleSpec) may contribute to the strength of these spectral features.

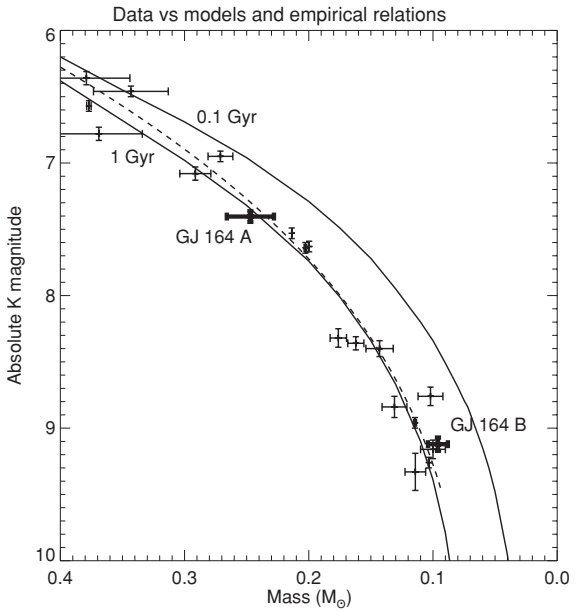


Figure 7. GJ 164 A and B in a mass–luminosity diagram (thick data points). The other data points were used by Delfosse et al. (2000) to establish an empirical K-band MLR for field stars (dashed line). The absolute K magnitudes are compared to the BCAH98 isochrones for 0.1 and 1.0 Gyr (two solid lines).

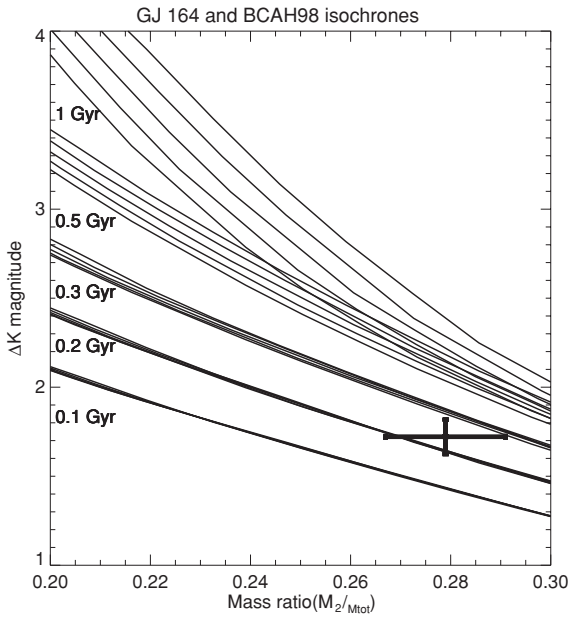


Figure 8. Mass ratio and K-band contrast ratio of GJ 164 (data point of coordinates (0.25, 1.8)) are compared to BCAH98 isochrones for ages between 0.1 and 1 Gyr. Each age generates a family of models taking into account the uncertainty on the total mass of the binary.

The case of GJ 164 B is a little bit more puzzling. Indeed, with a mass $\lesssim$0.1 $M_{\odot}$, GJ 164 B lies at the very end of the Delfosse et al. (2000) MLR, which makes comparisons with the MLR somewhat futile. The models however go well beyond. The reader will observe that as the mass decreases the luminosity they predict becomes a very steep function of the mass. Although GJ 164 B only lies 1σ away from the 1 Gyr isochrone, based on the dynamical mass and luminosity measurements reported in this work, the models seem to underestimate the luminosity of a <0.1 $M_{\odot}$ star by half a magnitude. A way to accommodate for this discrepancy would be for the GJ 164 A to be a tight binary, like for example the case of GJ 802 exposed in Ireland et al. (2008). Spectroscopic observations by Pravdo et al. (2004) however showed no evidence that GJ 164 A is a spectroscopic binary.

To provide a stronger case, Figure 8 plots more of the BCAH98 isochrones (0.1, 0.2, 0.3, 0.5, and 1 Gyr), this time in the plane mass ratio f (see Equation (4)), against the K-band magnitude difference. ΔK is a direct measurement made with

the aperture masking interferometry data (see Table 2), while f is a derived product of the joined fit, obtained in Section 3.2. Both observables are parallax-independent, which makes them more robust than the individual masses and luminosities.

Although taking into account all the uncertainties of our analysis by generating families of models within ± 1 standard deviation of the total mass of the binary, the low significance discrepancy between the data and the models presented in Figure 7 now turns out to be $\sim 2\sigma$ away from the 1 Gyr isochrone: the models presented in Figure 8 clearly favor a very young age (250 Myr), which accommodates for the apparent excessive luminosity of GJ 164 B, but is incompatible with the kinematics of the system. Note that the same conclusion holds when performing a similar analysis using H-band MLRs.

6. CONCLUSION

GJ 164 is an astrometric binary, whose orbit was resolved by aperture masking interferometry observations at Palomar and Keck. The consistency of the data is excellent, and exhibits a 1–2 mas precision in average for each measurement, made below what is usually accepted as the resolution limit of a telescope. These data, combined with earlier astrometric measurements, provide the following dynamical masses: $0.247 \pm 0.019\, M_{\odot}$ for the primary and $0.096 \pm 0.008\, M_{\odot}$ for the secondary.

Analysis of its infrared spectrum do not allow to unambiguously determine whether or not GJ 164 AB is metal-rich. When attempting to use theoretical models for solar metallicity very low mass stars to derive an age from the mass and luminosity of both components, we find that the models do not adequately fit the data, requiring very young ages to accommodate for an overluminous secondary. There are only a handful number of dynamical masses for stars below $0.1\, M_{\odot}$. The models are still waiting to be challenged in this critical region of the main sequence, right above the substellar limit.

The observations made here suggest that the models underestimate the luminosity of very low mass stars that have settled on the main sequence, but the precision of the dynamical data we present here, of the order of 10%, however does not provide a strong constraint.

A precise measurement of the contrast ratio in the R band, where the astrometry data was taken would already help to better constrain the mass ratio and ultimately improve the constraint on the individual masses, and make our statement about the models stronger. The very small separation (at most, 90 mas) makes this measurement difficult, even for *Hubble Space Telescope*. But because it lies so close from the substellar limit, GJ 164 B is a target that deserves most attention. In the meantime, a radial velocity curve would provide an independent measurement of the mass ratio. With an expected radial velocity amplitude of 2.8 km s^{-1}, and a period of two years, observations of GJ 164 with TripleSpec and TEDI will settle this uncertainty.

We thank the staff and telescope operators of Palomar Observatory and Keck Observatory for their support. This work is partially funded by the National Science Foundation under grants AST-0504874 and AST-0705085. This publication makes use of the Simbad database, operated at CDS, Strasbourg, France and the data products from the 2MASS, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. We extend special thanks to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests. Without their generous hospitality, the observations presented herein would not have been possible.

REFERENCES

Ali, B., Carr, J. S., Depoy, D. L., Frogel, J. A., & Sellgren, K. 1995, AJ, 110, 2415
Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Nature, 320, 595
Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, A&A, 337, 403
Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701
Burrows, A., Sudarsky, D., & Hubeny, I. 2006, ApJ, 640, 1063
Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, ApJ, 623, 1115
Delfosse, X., Forveille, T., Ségransan, D., Beuzit, J.-L., Udry, S., Perrier, C., & Mayor, M. 2000, A&A, 364, 217
Hayward, T. L., Brandl, B., Pirger, B., Blacken, C., Gull, G. E., Schoenwald, J., & Houck, J. R. 2001, PASP, 113, 105
Henry, T. J. 1998, in ASP Conf. Ser. 134, Brown Dwarfs and Extrasolar Planets, ed. R. Rebolo, E. L. Martin, & M. R. Zapatero Osorio (San Francisco, CA: ASP) 28
Henry, T. J., Kirkpatrick, J. D., & Simons, D. A. 1994, AJ, 108, 1437
Herter, T. L., et al. 2008, in Proc. SPIE 7014, Ground-based and Airborne Instrumentation for Astronomy II, ed. I. S. McLean & M. M. Casali (Bellingham, WA: SPIE), 70140X
Ireland, M. J., Kraus, A., Martinache, F., Lloyd, J. P., & Tuthill, P. G. 2008, ApJ, 678, 463
Jones, H. R. A., Longmore, A. J., Allard, F., & Hauschildt, P. H. 1996, MNRAS, 280, 77
Kervella, P., et al. 2008, A&A, 488, 667
Kirkpatrick, J. D., et al. 2000, AJ, 120, 447
Koornneef, J. 1983, A&A, 128, 84
Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762
Lampton, M., Margon, B., & Bowyer, S. 1976, ApJ, 208, 177
Liebert, J., & Gizis, J. E. 2006, PASP, 118, 659
Lloyd, J. P., Martinache, F., Ireland, M. J., Monnier, J. D., Pravdo, S. H., Shaklan, S. B., & Tuthill, P. G. 2006, ApJ, 650, L131
López-Santiago, J., Montes, D., Crespo-Chacón, I., & Fernández-Figueroa, M. J. 2006, ApJ, 643, 1160
Martinache, F., Lloyd, J. P., Ireland, M. J., Yamada, R. S., & Tuthill, P. G. 2007, ApJ, 661, 496
McCarthy, Jr, D. W., Henry, T. J., Fleming, T. A., Saffer, R. A., Liebert, J., & Christou, J. C. 1988, ApJ, 333, 943
Monnier, J. D. 2000, in Principles of Long Baseline Stellar Interferometry, ed. P. R. Lawson (Pasadena, CA: NASA), 203
Nakajima, T., Kulkarni, S. R., Gorham, P. W., Ghez, A. M., Neugebauer, G., Oke, J. B., Prince, T. A., & Readhead, A. C. S. 1989, AJ, 97, 1510
Nakajima, T., Oppenheimer, B. R., Kulkarni, S. R., Golimowski, D. A., Matthews, K., & Durrance, S. T. 1995, Nature, 378, 463
Paczynski, B. 1986, ApJ, 304, 1
Pravdo, S. H., & Shaklan, S. B. 1996, ApJ, 465, 264
Pravdo, S. H., Shaklan, S. B., Henry, T., & Benedict, G. F. 2004, ApJ, 617, 1323
Pravdo, S. H., Shaklan, S. B., Wiktorowicz, S. J., Kulkarni, S., Lloyd, J. P., Martinache, F., Tuthill, P. G., & Ireland, M. J. 2006, ApJ, 649, 389
Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, PASP, 111, 587
Readhead, A. C. S., Nakajima, T. S., Pearson, T. J., Neugebauer, G., Oke, J. B., & Sargent, W. L. W. 1988, AJ, 95, 1278
Reid, I. N., Gizis, J. E., & Hawley, S. L. 2002, AJ, 124, 2721
Reid, I. N., Hawley, S. L., & Gizis, J. E. 1995, AJ, 110, 1838
Sembach, K. R., & Savage, B. D. 1992, ApJS, 83, 147
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555
Tuthill, P., et al. 2006, in Proc. SPIE 6272, Advances in Adaptive Optics II, ed. B. L. Ellerbroek & D. Bonaccini Calia (Bellingham, WA: SPIE), 62723A
Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
van de Kamp, P. 1971, ARA&A, 9, 103
Weis, E. W. 1996, AJ, 112, 2300
Wilson, J. C., et al. 2004, in SPIE Conf. 5492, Ground-based Instrumentation for Astronomy, ed. A. F. M. Moorwood & M. Iye (Bellingham, WA: SPIE), 1295

The Astrophysical Journal, 756:8 (15pp), 2012 September 1 doi:10.1088/0004-637X/756/1/8
© 2012. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

THE USE OF SPATIAL FILTERING WITH APERTURE MASKING INTERFEROMETRY AND ADAPTIVE OPTICS

David Bernat[1], Frantz Martinache[2], Michael Ireland[3,4], Peter Tuthill[5], and James Lloyd[1]
[1] Department of Astronomy, Cornell University, Ithaca, NY 14853, USA
[2] National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720, USA
[3] Department of Physics & Astronomy, Macquarie University, NSW 2109, Australia
[4] Australian Astronomical Observatory, P.O. Box 296, Epping, NSW 1710, Australia
[5] Sydney Institute for Astronomy, School of Physics, University of Sydney, Sydney, Australia
Received 2011 March 1; accepted 2012 April 9; published 2012 August 9

ABSTRACT

Non-redundant aperture masking interferometry with adaptive optics (AO) is a powerful technique for high contrast at the diffraction limit with high-precision astrometry and photometry. A limitation to the achievable contrast can be attributed to spatial fluctuations of the wavefront—those within a sub-aperture and across sub-apertures—and temporal fluctuations within a single exposure. Spatial filtering addresses spatial fluctuations within a sub-aperture. An optimized pinhole in the focal place preceding the aperture mask is one approach for reducing the variation of the wavefront within a sub-aperture. Similarly, a weak spatial filtering effect is shown to be provided by post-processing the images with an apodized window function, typically used to minimize detector read noise and contamination from wide-separated sources. We explore the effects of spatial filtering through calculation, simulation, and observational tests conducted with a pinhole and aperture mask in the PHARO instrument at the Hale 200″ Telescope at Palomar Observatory. We find that a pinhole decreases stochastic closure phase errors and calibration errors, but that tight restrictions are placed onto the alignment of binary targets within the pinhole. We propose an observation strategy to relax these restrictions. If implemented, the pinhole could potentially yield an increase in achievable contrast by up to 10%–25% in H and K_s bands, and more at very high Strehl ($\gtrsim$80%). We also conclude that correcting low-order wavefront modes within the sub-apertures will be key for reaching high contrasts with extreme-AO systems such as the Gemini Planet Imager and PALM3K to search for planets.

Key words: binaries: close – brown dwarfs – instrumentation: adaptive optics – instrumentation: high angular resolution – techniques: high angular resolution – techniques: interferometric

1. INTRODUCTION

Current planetary searches using a coronagraph (e.g., Hinkley et al. 2011) excel at obtaining very high contrast (10^5:1) but are unable to probe at close separations ($\lesssim$300–500 mas) where the field of view is blocked by the occulting spot. This separation rules out the observation at physical separations $\lesssim$10 AU for many host stars. The detection of planetary companions with techniques providing both high contrast and high resolution will play a key role in identifying the full distribution of Jupiter-class planets and for investigating the mechanisms of planetary formation (Kraus et al. 2009), migration, and stability (Veras et al. 2009; Raymond et al. 2009 and, e.g., HR 8799, Fabrycky & Murray-Clay 2010).

Non-Redundant Aperture Masking Interferometry (NRM or *aperture masking*) in conjunction with adaptive optics (AO) is well established for yielding much more precise astrometry and photometry than AO alone at close separations (e.g., Kraus et al. 2008, 2011) and for the detection of high-contrast companions (Lloyd et al. 2006). The application of aperture masking with AO on 5–10 m class telescopes achieves contrasts of 10^2–10^3:1 at and outward of λ/D and nicely complements companion searches with a coronagraph. The increased contrast and resolution of NRM has been used with great effect for stellar multiplicity studies (Kraus et al. 2008, 2011), the detection of short-period brown dwarfs for dynamical mass measurements (Lloyd et al. 2006; Bernat et al. 2010), and a high-contrast search for inner planetary companions to HR 8799 (Hinkley et al. 2011). With the Gemini Planet Imager (Macintosh et al. 2008) and Project 1640 IFS (Hinkley et al. 2009) equipped with non-redundant masks, aperture masking interferometry from the ground will play a key role in the detection of exoplanets at close separations.

NRM observations use a mask to transform the full telescope aperture into a sparsely populated set of sub-apertures, constructed so that no two pairs of sub-apertures share the same baseline direction and length (i.e., are non-redundant). The strength of this technique draws on its measurement of the closure phase quantity (e.g., Jennison 1958; Readhead et al. 1988; Cornwell 1989), an observable that naturally mitigates the effect of wavefront errors on scales larger than a sub-aperture. These same wavefront errors produce the speckles of direct imaging, which dominate the image noise by orders of magnitude whether arising from atmospheric variation (Racine et al. 1999) or quasi-static instrumental effects (Hinkley et al. 2007). Their mitigation by closure phases enables higher contrast despite the blocking of 90%–95% of the flux by the aperture mask.

The achievable contrast of the technique is limited by its own calibration challenges, one of which arises from quasi-static spatial variation of the wavefront within the sub-aperture. Such wavefront errors erode the effectiveness of using closure phases (resulting in *redundancy noise* in the language of Readhead et al. 1988 and others). Similarly, decreasing the wavefront variation within the sub-aperture can increase the contrast of NRM observations further.

Optical and infrared long-baseline interferometers have implemented single-mode fibers and pinholes which spatially filter (i.e., smooth) aperture wavefront errors to improve measurements of complex visibility (Shaklan & Roddier 1987; du Foresto et al. 1997). Poyneer & Macintosh (2004) have studied

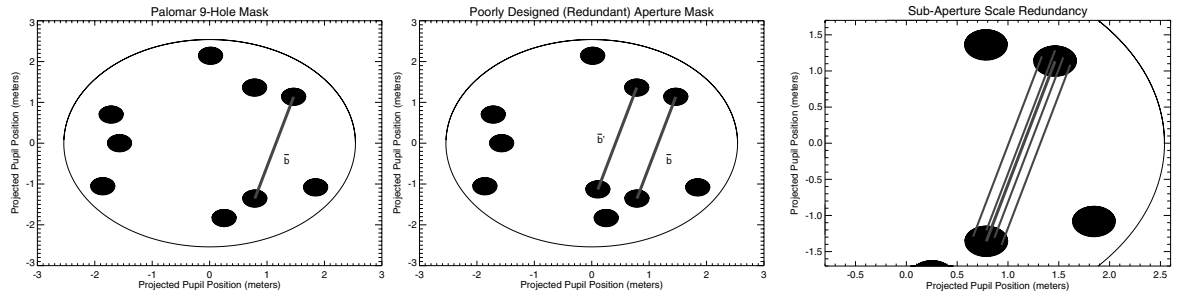


Figure 1. Aperture masks are designed to be non-redundant, but some redundancy persists because of the finite sub-aperture size. Left: the Palomar nine-hole mask. Each pair of sub-apertures acts as an interferometer. Center: a redundant mask. Two pairs of sub-apertures transmit the same baseline. As a result, the baseline carries redundancy noise into its closure phase. Right: because of the finite hole size, every baseline is redundant on sub-aperture scales. Spatially filtering the wavefront smoothes the wavefront phase, reducing noise from the sub-aperture redundancy.

the use of a pinhole to develop an AO wavefront sensor which more accurately measures the wavefront above a sub-aperture. Likewise, spatially filtering the wavefront errors with a pinhole placed in the image (focal) plane before the aperture mask may provide a means for substantially increasing the achievable contrast of NRM observations.

This paper provides a comprehensive analysis of NRM with pinhole spatial filtering. To establish its theoretical foundation, we present an analytic description of the combined technique (Section 2). Using an accurate simulation of an aperture masking equipped telescope, we derive the optimal pinhole size and estimate its expected performance, limitations, and restrictions. We also derive several results applicable to general aperture masking observations. (Sections 3 and 4). We have also installed a pinhole in the PHARO instrument of the Hale 200″ Telescope at Palomar Observatory. To study how spatial filtering improves the sensitivity of NRM observations and influences the astrometric and photometric characterization of discovered binary systems, we observed 26 single stars and 4 known binaries with and without the pinhole (Section 5). This paper further develops the understanding of NRM as a technique for detecting high-contrast companions and provides several new findings to guide the design of NRM experiments for next generation AO systems and instruments dedicated to exoplanet detection.

2. APERTURE MASKING WITH SPATIAL FILTERING

2.1. Aperture Masking: Current Technique

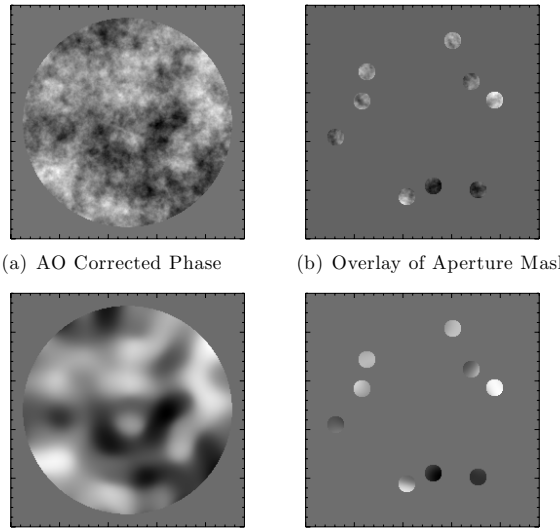
An aperture mask is positioned in the pupil plane behind the AO system and transforms the full aperture into a sparsely populated set of sub-apertures (Figure 1(a)). The resulting image of the target is a set of over-lapping fringes called the *interferogram*. The amplitude and phase of each fringe correspond to the measurement of one particular component of the complex visibility with spatial frequency $\boldsymbol{b}/\lambda$, where $\boldsymbol{b}$ is the baseline vector and λ is the observing wavelength. Multiplying the complex visibility of specific baseline triplets formed by three sub-apertures creates bispectra (Lohmann et al. 1983), the argument of which is the closure phase (Jennison 1958; Cornwell 1989). Closure phases are robust against pupil-plane phase errors, which are a source of direct imaging point-spread function (PSF) calibration errors and speckle noise, and provide the mechanism for obtaining more precise astrometry and photometry with the aperture masking technique.

Typical observations (including those in this paper) are conducted by taking one or more sets of target images interspersed with sets of one or more calibrators (unresolved stars near in the sky and of similar magnitude and color.) After the basic processing of the raw images (see, for example, Lloyd et al. 2006; Martinache et al. 2007; Bernat et al. 2010), the phase and log-amplitude of each fringe are extracted from the images and used to construct bispectra and closure phases. Mean values are obtained by averaging the quantities over a single set, and error estimates are derived from the scatter. Multiple sets are combined by weighted average. Calibration is performed by subtracting the closure phase and log-amplitude of the reference stars. Amplitudes are generally not used in companion searches because calibration is subject to stable atmospheric seeing and often is measurable to only 10%–50% (Tuthill et al. 2006). A binary model is fit to the closure phase data to minimize χ^2; a positive detection results in measurement of the binary parameters (separation, orientation, and wavelength-dependent contrast ratio). Errors in binary parameters are often taken from the curvature of the χ^2 space at minimum. Many examples of this implementation for the detection of stellar companions can be found in the literature (Kraus et al. 2008; Bernat et al. 2010; Kraus et al. 2011).

For targets in which the AO system provides stable and mostly coherent ($\sigma^2_{\rm rms} \lesssim 0.1$ rad^2) correction of the wavefront on sub-aperture scales, this observing mode typically measured closure phases with an error scatter of 1°–2°in the H band (1.6 μm) after a few minutes on a bright target, equivalent to a contrast of detection of about 100–200:1 at λ/D. This technique is calibration limited by a systematic closure phase component of one to several degrees which likely arises from quasi-static instrumental wavefront errors. Additional calibrators usually decrease, but do not fully eliminate, this component. To account for this, an additional error term is added in quadrature to the closure phase errors until the best-fitting model yields a χ^2 of unity.

2.2. Aperture Masking: Why Spatial Filter? Calibration Errors

A critical requirement of the aperture masking design is that each pair of sub-aperture creates a unique interferometric baseline. The lack of baseline redundancy ensures that any spatial frequency measured can be traced back to the interference of a unique pair of sub-apertures (Haniff et al. 1987; Roddier 1986). Closure phases constructed by non-redundant baselines will be less affected by pupil-plane phase which would otherwise



(a) AO Corrected Phase (b) Overlay of Aperture Mask

(c) AO + Spatial Filter Phase (d) Overlay of Aperture Mask

Figure 2. Effect of the pinhole filter on sub-aperture scale phase variation. (a) The AO corrected wavefront phase. Small-scale spatial inhomogeneities are apparent. (b) The AO corrected wavefront with an overlay of the aperture mask. Note that the wavefront phase is inhomogeneous within the sub-aperture. (c) The AO corrected wavefront after spatial filtering. The small-scale features are smoothed out; the wavefront exhibits structure with a characteristic scale close to that of the sub-apertures. (d) Within each sub-aperture, the spatially filtered phase is much more uniform.

distort measurements of the spatial frequency phase. When two or more baselines contribute to the same spatial frequency, the power adds partially incoherently depending on the phase difference of each contributing baselines (e.g., Figure 1(b)). A random component will be introduced into the resulting spatial frequency phase which cannot be removed by closure phases (yielding a so-called non-zero closure phase). This component, termed *redundancy noise*, is largest when the redundant baselines are incoherent and zero when the they are coherent. Readhead et al. (1988) provide an extensive treatment of redundancy noise for seeing-limited imaging.

The mask cannot be entirely non-redundant. The finite sub-aperture size means that baselines are redundant at least within a sub-aperture (Figure 1(c)). With sub-aperture scale correction provided by the AO system, this *sub-aperture redundancy* noise is largely removed (Tuthill et al. 2006), as compared to the uncorrected case, but still gives rise to closure phase measurement errors due to the remaining spatial incoherence within the sub-aperture. In completely analogous fashion, temporal variations to the baseline phase during a single exposure create *temporal redundancy* which also give rise to closure phase errors (again, see Readhead et al. 1988).

It may be illuminating to contrast the uncorrected and corrected cases. With uncorrected observing, redundancy restricts sub-aperture sizes to smaller than the characteristic size of atmospheric turbulence and exposure times shorter than the atmospheric coherence time. AO removes both constraints since good correction supplies a stable, mostly coherent wavefront across the full aperture.

Current NRM masks are designed with sub-apertures on the order of the AO actuator spacing; this need not be the case and will likely not be so in upcoming extreme-AO aperture masking

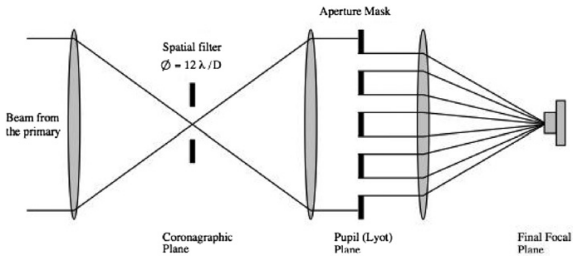


Figure 3. Optical setup for pinhole-filtered aperture masking interferometry at Palomar. One takes advantage of the coronagraphic capabilities of PHARO by inserting the aperture mask in the Lyot wheel and the spatial filter in the Slit wheel.

experiments. Atmospheric and AO residuals decorrelate on timescales much shorter than the exposure length and thus likely only contribute to the stochastic variation of closure phases from one image to the next. Changes in seeing between target and calibrator observations change the magnitude of the stochastic variability of closure phases (the signal to noise of the measurement), but do not introduce calibration offsets (Roddier 1986), and can thus be minimized by additional exposures. This is precisely why closure phases provide a robust measurable unlike visibility amplitudes, which are poorly calibrated if seeing changes (Tuthill et al. 2006).

Quasi-static instrumental wavefront errors contribute to all spatial scales and vary on timescales of tens of minutes (e.g., Bloemhof et al. 2001; Hinkley et al. 2007), or with movement of the telescope (Marois et al. 2005). As long as these errors remain static over a single exposure, large-scale wavefront changes (those larger than a sub-aperture) are removed by closure phases. One of the great advantages of using non-redundant masking to mitigate the quasi-static imaging problem is that closure phases require only a single image (they "self-calibrate"; Cornwell 1989). Instrumental errors change between observations of the target and calibrator, but only those smaller scale than a sub-aperture contribute to aperture masking calibration errors.

We propose that spatial filtering the wavefront before its propagation through the aperture mask can be used to reduce the inhomogeneity of the wavefront phase within the sub-aperture and, by virtue of being more coherent, reduce sub-aperture redundancy noise. Figure 2 illustrates the effect of an optimized pinhole spatial filter (discussed in the next subsection) to smooth the phase of a simulated AO corrected wavefront (Strehl ratio ∼50% in K_s). The portion of the wavefront sampled by each sub-aperture is much more uniform after spatial filtering.

We anticipate that spatial filtering can lead to several measurable improvements of the closure phases, including decreases in systematic calibration error and stochastic variation from one image to the next. Measured visibility amplitudes, by virtue of their dependence on sub-aperture inhomogeneity, can show less degradation and be more robust to changes in seeing.

2.3. Pinhole Filtering

One possible implementation of the pinhole spatial filter, reminiscent of the four planes of a coronagraph (see Ferrari 2007 for a review), is shown in Figure 3. The pinhole, positioned at the center of the field and in what is usually referred to as the coronagraphic plane, blocks light beyond its inner transmission region. This acts to spatially filter the wavefront before entering the non-redundant aperture mask, located in a re-imaged pupil

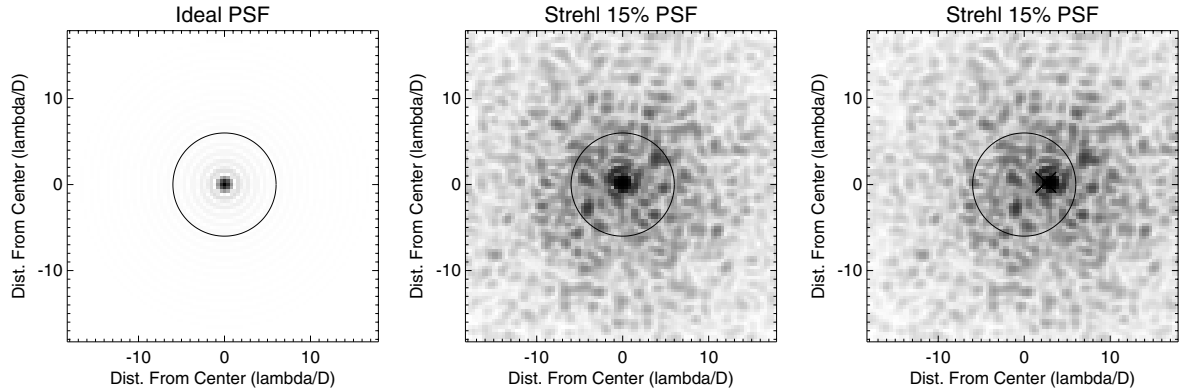


Figure 4. Imaging an unresolved target through a pinhole. The PSF of three targets is shown with a pinhole of size $6\lambda/D$ overlaid. Square root contrast scaling is used to highlight the truncated flux. Left: the pinhole, located in an image plane, truncated the portion of electric field which forms the outer rings of the PSF. In perfect seeing, the total flux blocked is very low. Center: wavefront errors dispel flux outward creating a diffuse halo around the target. The blocked flux increases, and more power aliases back into sub-aperture scales, resulting in closure phase errors. Right: when asymmetrically truncated, the center of light shifts toward the pinhole center (black ×). Each component of a binary is truncated differently, leading to errors in astrometry or contrast.

plane. (The mask location is equivalent to the Lyot-plane of coronagraphy.)

The transmission profile of the pinhole is a top-hat function, with diameter $d_{\rm spf}$ which will be expressed for convenience in units of λ/D, with D the diameter of the telescope. At the location of the aperture mask, this is equivalent to the wavefront convolved by a Jinc function[6] of characteristic diameter $\lambda/d_{\rm spf}$.

Efficient spatial filtering aims to make the wavefront uniform within one sub-aperture of the non-redundant mask, and so one wants the characteristic diameter to be matched to the diameter of a sub-aperture, suggesting

$$d_{\rm spf} \approx \frac{D}{d_{\rm sub}} \text{ (in units of } \lambda/D). \quad (1)$$

A larger pinhole decreases the impact of the spatial filtering; a smaller pinhole restricts the field of view and begins to impinge on the signal we wish to measure.

This technique differs from the form of spatial filtering implemented in optical and infrared long baseline interferometers, where the signal from each telescope is injected into a single-mode fiber or pinhole before beam combination and extraction of the interferometric signals (du Foresto et al. 1997). The use of single-mode fibers on long baseline interferometers has shown increases in phase estimations by a factor of two, or more in bad seeing, despite the loss of flux (Tatulli et al. 2010). The implementation of this approach with non-redundant aperture masking would require the injection of light from each sub-aperture into a single-mode fiber or pinhole before recording the interferogram. Given the simplification of the single pinhole implementation, it offers, a priori, an appealing alternative (Keen et al. 2001).

2.3.1. Aperture Masking through a Pinhole

Given that the aperture mask is located in a pupil plane, the PSF of masking interferometry is invariant to target position over a wide field of view (several arcseconds). With the pinhole filter in place (in an image plane), the PSF and system response vary with target position relative to the pinhole, a smaller effective field of view.

The Appendix shows that the use of the pinhole filter alters the measured complex visibility of a point source:

$$V_{\rm spf}(\boldsymbol{b}) = T(\boldsymbol{b}) \int d\boldsymbol{\theta} \Pi[\boldsymbol{\theta}/d_{\rm spf}] \times e^{2\pi i \boldsymbol{\theta}\cdot\boldsymbol{b}/D} \times \mathrm{PSF}[\theta - \boldsymbol{\alpha}], \quad (2)$$

where $\boldsymbol{\alpha}$ is the location of the point source on the sky (measured in angular units of λ/D), and $\boldsymbol{\alpha} = 0$ corresponds to perfect alignment of the source in the pinhole. The baseline being measured is $\boldsymbol{b}$, corresponding to a spatial frequency $\boldsymbol{u} = \boldsymbol{b}/\lambda$. The complex visibility is usually expressed as a function of the spatial frequency; here we present it as a function of baseline and wavelength for clarity. The telescope transfer function, $T(\boldsymbol{b})$, gives the spatial frequency response of the aperture or aperture mask in terms of the baseline. It is defined explicitly in the Appendix. The top-hat function, $\Pi[|\boldsymbol{\theta}|/d_{\rm spf}]$, is equal to one for $(|\boldsymbol{\theta}|/d_{\rm spf}) < (1/2)$ and zero otherwise.

From three baselines vectors, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $-\boldsymbol{b}_1 - \boldsymbol{b}_2$, a closure phase is the argument of the product of the complex visibilities:

$$\phi_{\rm cp} = \arg\left[V(\boldsymbol{b}_1)V(\boldsymbol{b}_2)V(-\boldsymbol{b}_1 - \boldsymbol{b}_2)\right]. \quad (3)$$

The observed visibility amplitudes and closure phase change by an amount that depends on the position of the target with respect to the pinhole (its alignment) and the PSF. (e.g., Figure 4). The transmission of the pinhole blocks high spatial frequency content of the wavefront, even in the absence of wavefront errors. In this way, removal of the higher spatial frequencies can alias into changes in lower, baseline frequencies, similar to the effect observed by Poyneer & Macintosh (2004). Those on sub-aperture scales compete with closure phase measurements.

As the PSF varies, so too will the aliased phase errors. This introduces a stochastic component to the closure phase errors which appears to undercut the effectiveness of the pinhole. A portion of the phase errors will also remain fixed and be strictly a function of the asymmetrical truncation of the target and its mean PSF. The latter we term misalignment error to distinguish it from other calibration components. Generally speaking, both will increase as the target is positioned further from the pinhole center.

[6] The Jinc function is defined as $\mathrm{Jinc}(x) \equiv (J_1(x)/x)$. It is the two-dimensional Fourier transform of a circular aperture and is related to the Airy function by $\mathrm{Airy}(x) = 4\,\mathrm{Jinc}(x)^2$.

Qualitatively, we can illustrate that the asymmetrical truncation of the target shifts the visibility and closure phase. Consider Figure 4(c). The visibility phase is influenced by the position of the target on the sky; the asymmetrical truncation of the target shifts its center of light, and so too the visibility phase. While closure phases are normally invariant to the position of the target, the pinhole breaks this invariance. The overall isotropy of the PSF and circular pinhole shape suggests that the misalignment error is symmetric about the origin. That is, just as aligning the target at two diametrically opposed positions will induce equal and opposite shifts in the center of light, the overall symmetry suggests equal and opposite phase errors as well.

The aliased phase errors are small as long as the majority of the flux resides within the pinhole and only competes with the closure phase measurements if the power aliases back into sub-aperture scale wavefront deviations. For this reason, shorter wavelengths will be preferentially advantaged due to their small PSF, assuming that the Strehl ratio is the same at each wavelength. Typically, AO residuals will also be larger for shorter wavelengths (producing a more substantial halo) and this partially counteracts the shorter wavelength advantage. More generally, this suggests that to minimize aliasing, the pinhole could be optimized to filter only frequencies that are not corrected by the AO system; in that case the pinhole size would be matched the AO actuator size instead of the masking sub-aperture size.

Each target of a binary will be truncated differently. The misalignment error can be calculated directly by replacing the single-star PSF in Equation (2) or Figure 4 with the binary image. This will be approximately the same as the error of each component individually. In other words, the misalignment error of the primary added to the misalignment error of the secondary multiplied by the secondary–primary contrast ratio. Assuming that the misalignment error is point symmetric about the origin, aligning the binary center of light near the pinhole center appears to be a viable strategy for minimizing the misalignment error.

If the misalignment component is large enough as to become the limiting component in closure phase noise, there are three approaches which can be employed for removing the component from science data. Empirically, the component can be calibrated by observing a single star aligned at the same location as the primary under similar AO performance. In practice, this may be limited by the accuracy with which one can repeatedly (visually) align a target within the pinhole, and the component arising from the (unknown) companion remains. Computationally, if the location of the primary is known along with an estimate of the PSF, then the misalignment component can be included in the mathematical closure phase model that is fit to data. This approach requires a direct image of the primary within the pinhole to determine its location and estimate its PSF. Both may detract from the overall efficiency of observing. Finally, a third option is to take several sets of aperture masking data with the target (center of light) at various alignments (i.e., dithering) close to the pinhole center. Specifically, if diametrically opposed alignments have misalignment errors opposite in sign, then the component could be averaged out.

2.4. Post-processing with a Window Function

In this course of our investigation of the pinhole filter, we also characterize the use of a tapered window function on masking images to reduce the impact of various noise sources and small-scale wavefront errors on closure phases. For example, we use a super-Gaussian function, $\exp(-kx^4)$, with our experiments for its flat-top and quickly tapering Fourier transform. For concreteness, we will use this same function as an illustrative example here (Figure 5). The window function is described by the size of its half-width at half-maximum (HWHM) measured in pixels on the detector or in units of λ/D.

Multiplying the images by a window function that retains the intensity of the interferogram core but tapers at its edges greatly reduces the contamination of detector read noise and dark current, or scattered light from nearby targets not of scientific interest. Per-pixel Gaussian-distributed read noise, when Fourier transformed, results in a per-baseline Gaussian-distributed uncertainty. The per-baseline noise has a magnitude that is directly related to the total transmission of the window function. Therefore, a tighter window function removes more read noise.

However, complex visibilities are extracted from the Fourier transform of the image, so the application of a window function is identical to the convolution of the complex visibilities with the Fourier transform of the window function. This mixes the complex visibilities of the baselines and, ultimately, restricts the minimum window function size.

The aperture mask is designed to be non-redundant for baselines extending between the centers of sub-apertures; the finite size of the sub-apertures give rise to islands of transmitted power in the power spectrum, which we call *splodges* (see Figure 5, bottom left). The splodges of a completely non-redundant mask do not overlap, and hence the separation between neighboring splodge peaks is about $2\,d_{\rm sub}$. (In practice, some overlap at the splodge edges may be exchanged for better coverage or larger sub-apertures.) Furthermore, neighboring splodges may arise from completely separate pairs of sub-apertures (i.e., separate baselines), and one would not assume any coherence between the baseline phases. For this reason, window functions narrower than about $\lambda/2d_{\rm sub}$, or those without quickly tapered Fourier transforms, will mix the incoherent baselines of neighboring splodges and rapidly increase the error of closure phases, similar to redundancy noise.

The optimally sized window function will balance the reduction of read noise with the increase of redundancy noise.

Post-processing with a window function also reduces the impact of small-scale wavefront errors. This is most readily recognized by noting that the outer rings of the PSF encapsulate the high spatial frequency content of the wavefront; removing this flux also removes the high spatial frequency content of the wavefront.

This method of applying a post-processing window function to spatially filter differs in two major ways from implementing an on-telescope pinhole. Applying filtering in post-processing allows easy tailoring of the shape of the window function (including suppression of the tails of its Fourier transform), a task considerably harder with an apodized pinhole. The pinhole, however, filters the wavefront phase noise at each instant; the window function only filters some form of the time-accumulated phase noise.

3. SIMULATED OBSERVATIONS

We have developed an accurate simulation of the Palomar aperture masking interferometry experiment in order to explore the effect of spatial filtering on wide band data and to optimize the pinhole for implementation on the telescope.

The Palomar nine-hole aperture mask has been described previously in Lloyd et al. (2006) and is shown in Figure 1(a).

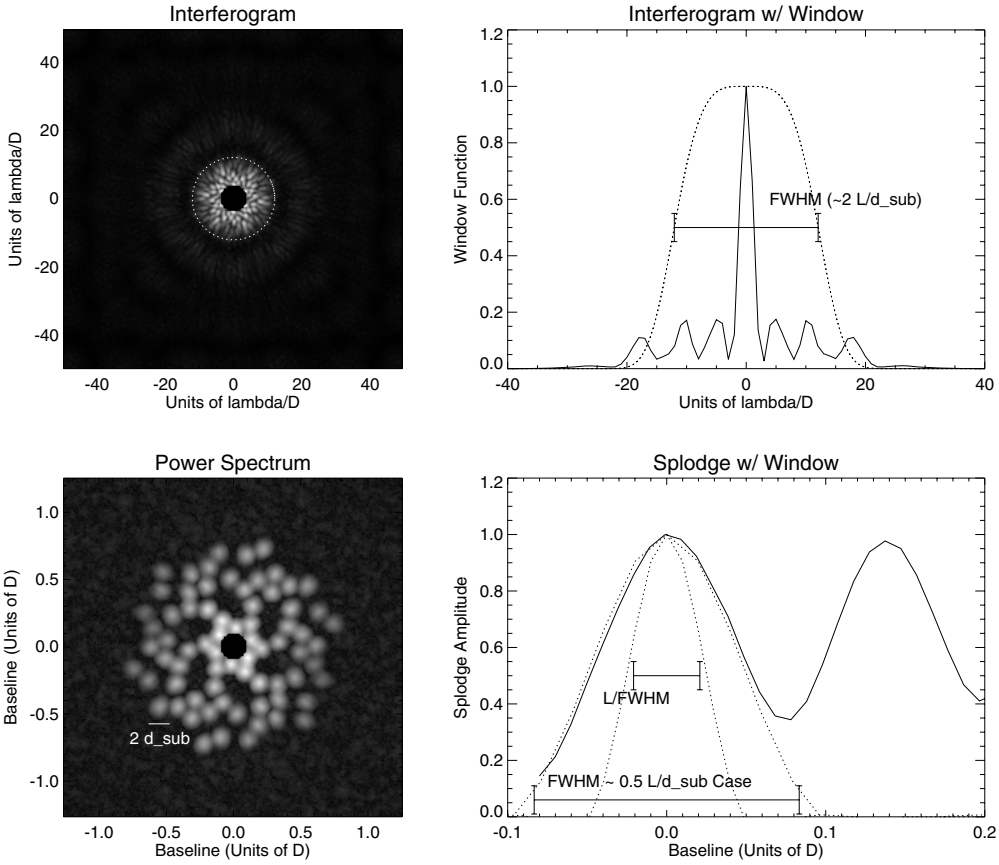


Figure 5. Effect of a window function. Top left: the aperture mask produces a set of interference fringes beneath an envelope of size $\lambda/d_{\rm sub}$, as seen in this Palomar nine-hole masking image. The central peak has been zeroed to highlight the envelope and outer rings. The radius of the white ring is $\lambda/d_{\rm sub}$. Bottom left: the power spectrum of the same interferogram, presented in units of baseline/D rather than spatial frequency. Each island of transmitted power (or *splodge*) is of size $2d_{\rm sub}$. This is expected, as the transmission function is related to the autocorrelation of the mask. Top right: using a window function of characteristic HWHM $\lambda/d_{\rm sub}$ (here, a super-Gaussian) removes the interferogram wings and its associated read and wavefront noise. Bottom right: the window function produces a convolution kernel of size λ/2HWHM. Note that a window function larger than 0.5 $\lambda/d_{\rm sub}$ creates a kernel larger than $2d_{\rm sub}$ and mixes neighboring splodges, adding redundancy noise (see the text).

Placed in the pupil plane, the mask has baselines ranging from 0.71 m to 3.94 m and sub-apertures that are 0.42 m in diameter.

As discussed in Pravdo et al. (2006) and Bernat et al. (2010), aperture masking and the implementation of closure phases work most effectively using the short exposure times, when large-scale wavefront errors (be they atmospheric or instrumental) can be regarded as approximately static. The typical aperture masking operation at Palomar uses exposure lengths of 431 ms. Over this timescale, the evolution of the atmosphere produces a highly dynamic wavefront on sub-aperture scales.

3.1. Characterization and Simulation of Palomar's Atmosphere

Studies by Ziad et al. (2004) and Linfield et al. (2001) conducted with the Palomar Testbed Interferometer (PTI) and Hale 200″ Telescope confirm that the atmospheric turbulence power spectrum is approximately Kolmogorov with an outer scale most often within the range 10–50 m. The median-seeing Fried parameter (r_0) has been measured to be 9.0 cm at 550 nm, dropping to 3.8 cm during bad, but regularly observed seeing (Dekany et al. 2007).

Advection (wind) speeds drive the evolution of atmospheric structure within the sub-aperture on intervals between 0.1 and 1.0 s. Using measurements of the Palomar atmospheric temporal structure function over four nights, Linfield et al. (2001) derive wind speeds typically less than 4 m s^{-1}. For shorter timescales, the temporal structure function decays exponentially. The characteristic decay time, which scales with $\lambda^{6/5}$, has been measured in two surveys to vary between 15–80 ms (Ziad et al. 2004) and 60–150 ms (Linfield et al. 2001) in K_s.

Our simulations assume an outer scale of 50 m and Fried parameter value 9.0 cm, which correspond to 48 cm in K_s and 32 cm in H. We use wind speeds of 5 m s^{-1} and temporal decay time of 60 ms.

3.2. Numerical Simulation

We generate time-evolved Kolmogorov phase screens following the procedures of Lane et al. (1992) and Glindemann et al. (1993). These phase screens are characterized by three parameters: the Fried parameter of an instantaneous phase screen, the wind velocity which blows the static phase screen across the aperture, and an additional parameter driving the

decorrelation of high spatial frequencies between time steps. From this last parameter emerges the exponentially decreasing temporal structure function for short timescales.

A single 431 ms exposure image is constructed by adding 24 sub-images, each of which is an instantaneous snapshot separated in time by 18 ms. We chose this time step to sufficiently sample the atmospheric evolution over relevant timescales: the wind sub-aperture crossing time is 84 ms; the coherence decay time is 60 ms.

The generated images are 512 × 512 pixels, designed to preserve the pixel scale of the PHARO detector. In frequency space, the 5.08 m aperture spans slightly less than 256 × 256 pixels in the K_s band. Phase screens of this size are generated; this corresponds to an inner scale of approximately 2 cm and sub-apertures approximately 450 pixels in area. We repeated several simulations with four times as many pixels and arrived at similar results; from these results we conclude that our simulation is well sampled.

The simulation generates one sub-image for each phase screen using an incoming monochromatic wavefront perturbed by the screen. After passing through the telescope aperture, the wavefront is corrected by AO, then propagates through the spatial filter, the aperture mask, and, finally, forms an image on the detector. We ignore read noise and photon noise. This simulation is similar to that used by Sivaramakrishnan et al. (2001).

We model the AO system as an instantaneous high-pass filter of the form $A(k) = 1/(1 + (k_c/k)^2)$, with a cutoff imposed by the Nyquist frequency of the actuator spacing, $k_c = N_{\rm act}/2D$ (about 1.4 cycles m^{-1} at Palomar). This filter is applied to the Fourier transform of the phase screens and accurately reproduces the wavefront residuals observed under optimal operation at Palomar (Dekany et al. 2007). This overestimates the typical AO performance on faint targets, particularly the performance of tip-tilt suppression. To account for this, we apply the following modified AO filter function, which degrades the low spatial frequency AO response:

$$A(k) = \min\left[\frac{1}{1 + (k_c/k)^2}, 0.05\right],$$
$$k_c = N_{\rm act}/2D. \quad (4)$$

To include the polychromatic effects, each wide band transmission filter is divided into a number of subintervals (generally four). A monochromatic image is generated for each subinterval (taking into account wavelength-dependent effects), and these images are added to form a single polychromatic sub-image. The sub-images are co-added to create a single 431 ms exposure.

This produces direct imaging Strehl ratios of 40%–50% in the H band and 60%–70% in K_s.

4. THE PALOMAR PINHOLE EXPERIMENT

4.1. Pinhole Implementation on PHARO

The PHARO instrument (Hayward et al. 2001) is especially suited to a pinhole implementation. PHARO has been designed with coronagraphic capabilities, giving access to both a focal plane and a pupil plane before the final focal plane (Figure 3).

The pinhole is placed in the focal plane before the aperture mask. Based on our simulations from the next subsection, a pinhole of angular diameter 0.779 arcsec was installed at the Lyot Stop in PHARO in 2008 June. This pinhole was chosen for optimal use at 1.6 μm (H band), corresponding to an angular size of 12 λ/D. The pinhole is of size 8.7 λ/D at 2.2 μm (K_s band).

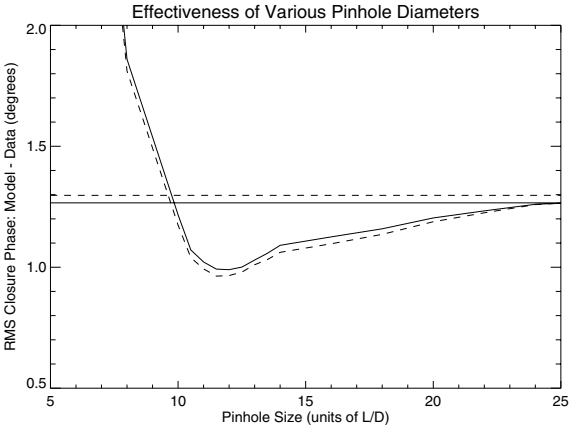


Figure 6. The rms fit residuals of simulated data (AO corrected H-band observing, no read noise) with pinholes of various size, analyzed with (solid line) and without (dashed line) a window function. The model used is an analytic model of closure phase for a single star without noise or wavefront errors. The horizontal line is the measurement level without any pinhole in place. The pinhole filter is most effective within the range 11–14 λ/D.

4.2. Pinhole Size Optimization

With sub-aperture sizes of 0.42 m, the estimate from Section 2.3 suggests the optimal pinhole size to be $D/d_{\rm sub} \approx 12\lambda/D$. Using the simulation of the Palomar aperture masking experiment described in Section 3.2, we can determine the optimal pinhole size under typical observing conditions.

For various pinhole sizes, we simulated 100 H-band images of a 10:1 binary with companion separation 150 mas under typical Palomar observing conditions, without read noise, and with the primary aligned at the center of the pinhole. Images of a single star were simulated to be used as a calibrator. We analyzed these images both with and without a window function, and calculated the root-mean-squared (rms) residuals of the simulated closure phases fit to a model binary. Figure 6 shows these results for each pinhole size (with window function, solid line; without window function, dashed line). For reference, the rms obtained with no pinhole in place is included as a horizontal line.

The spatial filter performance can be broken into following three classes:

1. A small pinhole ($\lesssim$10 λ/D) impinges on the field of view, rejecting enough of the light coming from the off-axis companion so that the closure phases will not match the model. (See also Section 2.3.1.)
2. A very large pinhole ($\gtrsim$20 λ/D) provides very little filtering and the data are statistically similar to the unfiltered case.
3. The operational pinhole range (10–20 λ/D) reduces the image to image variation of closure phases and produces data that better fit its model. The range 11–14 λ/D is most effective, reducing the fit residuals by roughly 25%.

These results agree with the estimate at the top of this subsection.

Finally, we note that because the window function itself provides some spatial filtering, the inclusion of the pinhole provides slightly less improvement if compared to the case in which no window function is used.

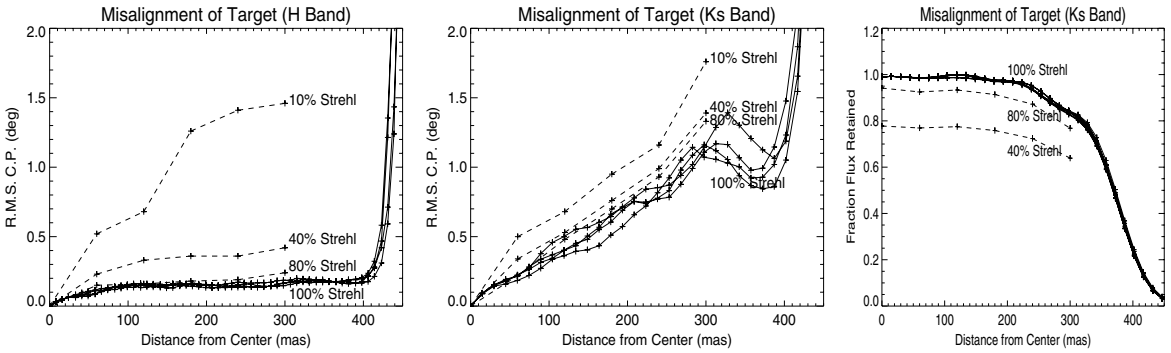


Figure 7. Misalignment of a single star within the pinhole introduces closure phase errors. (a) The Palomar nine-hole mask, overlaid with three closure triangles for which the misalignment errors are calculated. (b) Error due to misalignment at 1.6 μm (H band) as a function of target distance from the pinhole center. (Several azimuthal orientations are plotted for each separation.) (c) Closure phase errors at 2.2 μm (K_s band) in which the pinhole is smaller. In both cases, errors in visibility amplitude are below 0.005 for the same ranges.

4.3. How Important Is Target Placement?

In Section 2.3.1, we showed that the asymmetrical truncation of the target by the pinhole alters the measured closure phases, even in the absence of wavefront errors, an effect we called misalignment error.

Equation (2) can be directly integrated to determine the misalignment error when a target is observed through the pinhole. Alternatively, we chose to simulate the effect to more accurately reflect the details of our analysis pipeline. For clarity, we present the rms closure phase deviation between the pinhole and non-pinhole values with the Palomar nine-hole mask (averaged over the set of 84 closure phases) in Figure 7.

By inspection, we see that in the H band and at high levels of correction (Strehl $\gtrsim$40%), the pinhole introduces less than 0.°3 rms closure phase error at nearly any alignment within the 440 mas diameter pinhole. At a diameter of $12\lambda/D$, less than 2% of the total flux resides outside the pinhole and very little aliases back into the baseline frequencies. The misalignment deviation is also insensitive to the orientation of the misalignment; the aperture mask is three-fold symmetric in physical space, but provides a uniform coverage of spatial frequencies. At moderate Strehl ratios, the misalignment deviation increases, but observationally, this increase is balanced by a larger benefit of spatial filtering.

The alignment requirements are tighter in K_s. Because the pinhole is designed for the H band, its size is only $8.9\lambda/D$ in K_s. Good positioning is even more important given that AO residuals and closure phase errors are usually smaller. Alignment can introduce up to 1.°0 of closure phase errors if misaligned by 200–300 mas, or roughly 2–3λ/D.

The presence of a companion adds an additional misalignment error, although this error will be attenuated by the secondary–primary contrast ratio. Therefore, particularly for high-contrast binaries, the misalignment errors provided here give good rule-of-thumb indications of how well a target must be positioned within the pinhole. Given that aperture masking observations at Palomar typically yield 1°–2° of closure phase, targets must be positioned close to the pinhole center during K_s-band observations.

Finally, we note that the percentage of blocked flux is a weak function of the target alignment, and a strong function of the AO performance and size of the direct imaging halo (Figure 7, right panel). In the observations of Section 5 that compare pinhole and non-pinhole observations, the percent loss of flux is a measurable quantity. From the results of this plot, it is clear that large flux drops can be attributed to momentary drops in AO performance or very large shifts in alignment. Astrometric jitter on the scale of hundreds of mas is never seen. Practically, if a series of images contains one or a few with large flux drops, and the target was initially well aligned, then identifying large flux drops is a useful proxy for excluding frames taken with poor AO performance.

4.4. Window Function: Optimal Size and the Palomar Nine-hole Mask

For our experiment, we use a super-Gaussian function, $\exp(-kx^4)$, for its flat-top and quickly tapering Fourier transform. We describe the window function by the size of its HWHM, $w_{\rm pix}$, measured in pixels on the detector.

The optimal window function size finds the balance between eliminating read noise and increasing redundancy noise. We present a measure of the effectiveness of various window function sizes in the presence of various levels of read noise in Figure 8.

We simulated 100 images (256 × 256 in size) of a 10:1 binary with companion separation 150 mas and of a calibrator in the same fashion as described in Section 4.2 but without a pinhole in place. To each set of images we added Gaussian read noise with a per-pixel noise level ranging from 0.2% to 5.0% of the peak intensity. (This corresponds to targets of 7th to 10th magnitude when taking 6 s exposures on the PHARO detector.) The images were processed with window functions of various sizes and we calculated the rms residuals of fits to a model binary. In Figure 8, we plot the ratio of the measured rms with a window function to the same set of images without the window function. For all levels of read noise, an optimally sized window function improves the model fits. We caution that, because the effect of read noise on closure phases scales with image size and the choice of default image size is arbitrary, the results can only be evaluated qualitatively.

Several features are apparent. First, the optimal window function is approximately $\lambda/d_{\rm sub}$, or $\sim 12\lambda/D$ for the Palomar aperture mask, with tighter windows for the high-noise cases; one expects the optimal window size to be a function of the aperture mask. This makes qualitative sense: the interferogram image is a set of interference fringes under an Airy function envelope of characteristic size $\lambda/d_{\rm sub}$. One would expect the optimal window function to crop out those pixels with signal

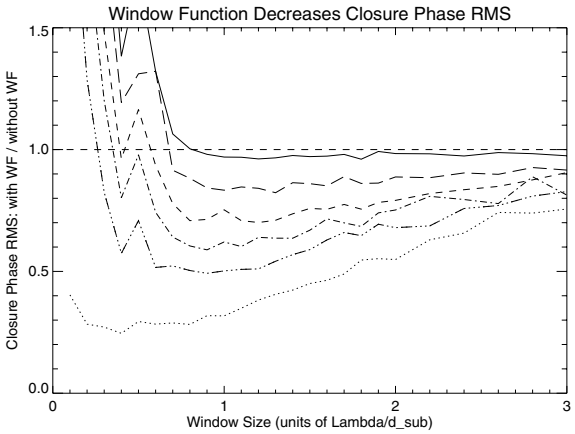


Figure 8. Window functions reduce closure phase error from read noise. These curves display the reduction of rms closure phase errors by using the *ratio* of the windowed and non-windowed cases (i.e., the *y*-axis is a ratio, with 1.0 corresponding to no improvement). These curves, from top to bottom, display the ratio of rms closure phase error when read noise is 0% (top, solid), 0.2%, 0.4%, 0.6%, 1.0%, and 5.0% (bottom, dotted) of the peak image intensity. The optimal window function is typically of size $\sim\lambda/d_{\rm sub}$, or $\sim 12\lambda/D$ for the Palomar aperture mask, with higher read noise favoring tighter window functions. Smaller window functions quickly add large amounts of redundancy noise (See the text). Note: even with no read noise (solid curve), a window function reduces closure phase errors, indicating that the window function provides an effect similar to spatially filtering the wavefront.

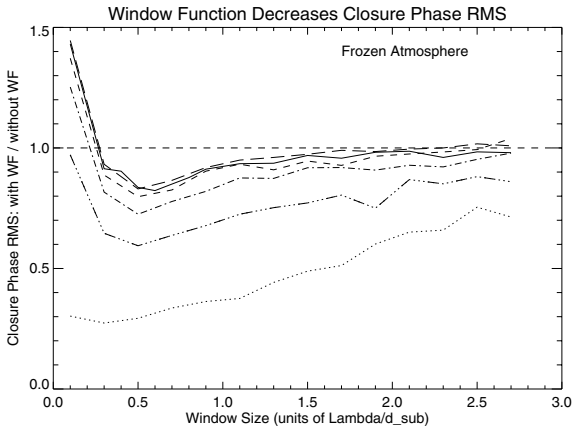


Figure 9. Curves showing the effectiveness of the window function as in Figure 8, except that the wavefront is static over an exposure. Again, these curves display the reduction of rms closure phase errors by using the *ratio* of the windowed and non-windowed cases (i.e., the *y*-axis is a ratio, with 1.0 corresponding to no improvement). Most notably, a window function provides better spatial filtering when the wavefront is static (solid line).

to noise less than one. Beyond $\lambda/d_{\rm sub}$, the intensity of the interferogram drops below a few percent of its peak value, comparable to the read noise.

Second, too small a window function quickly adds redundancy noise into the measurements. This turning point is near $0.5\ \lambda/d_{\rm sub}$.

Third, even in the absence of read noise, a window function decreases closure phase noise (solid curve). This demonstrates the capacity of the window function to spatial filter the wavefront phase noise, even though the improvement is only about 3%–4%.

To emphasize the utility of the window function as a spatial filter of wavefront errors, we simulated another set of images in which the wavefront is static over the exposure. With only spatial variation of the wavefront, the window function reduces closure phase errors by nearly 20% (Figure 9).

By comparing these results to those obtained with an optimized pinhole (25%), it is clear that the pinhole and window functions only provide comparable spatial filtering when the wavefront phase errors is static during an exposure. When accounting for the more realistic, dynamic wavefront, the pinhole provides superior spatial filtering.

Misaligning the peaks of the window function and interferogram introduces a tiny amount of closure phase error. Even an unrealistic misplacement of $10\lambda/D$ (about 26 pixels on PHARO at $1.6\,\mu$m) with a tight window of size $0.7\ \lambda/d_{\rm sub}$ produces an error below $0\overset{\circ}{.}06$. This will likely be relevant for observations taken by space telescopes.

5. OBSERVATIONS

Between 2008 June and 2009 September, we observed 26 single-star targets and 4 known binaries spanning infrared magnitudes between 6.0 and 9.0 using the PHARO instrument on the Hale 200″ Telescope at Palomar Observatory.

Each aperture masking observation was conducted with and without the pinhole in place to compare the effectiveness and practicality of using the pinhole filter during ground observing. An observing sequence consisted of sets of 20 images (6 second exposures) with the Palomar 9-hole aperture mask in several standard infrared bands. Typically, images were taken in a particular band with the pinhole in place, then again with the pinhole removed, until the sequence of bands had been taken. Similar observations of a calibrator were then taken. This was done to minimize changes in seeing between comparison observations and to minimize the effect of instrumental wavefront changes from slewing the telescope on calibration. Care was taken to use identical detector configurations (position, etc.) and to align the target center of light at the center of the pinhole. This alignment procedure added approximately two to five minutes additional overhead to each observation set.

The aperture mask, data taking procedure, and custom IDL pipeline to analyze aperture masking images have been previously described by Lloyd et al. (2006), Kraus et al. (2008), and Bernat et al. (2010). The mean and variance of the closure phases and amplitudes were calculated (across the set of images) and calibrated. The calibrated closure phases were found by subtracting the calibrator closure phases from those of the target; their errors were estimated by adding in quadrature the errors of the targets and calibrator. The closure phase signal of a single star is zero; deviations from zero may indicate the presence of a companion or result from wavefront errors. The log-amplitudes are calibrated identically and are included to determine if spatial filtering improves measurement of the amplitudes, but are not used for the analysis of the binary candidates. Companions are located by fitting the closure phase data set with a three-parameter binary closure phase model (separation distance, position angle, and contrast ratio) to minimize χ^2.

The targets are bright enough that the per-pixel read noise is at or below a few percent. For the analysis that follows, we use the results of Section 4.4 and a window function with HWHM of $1.0\ \lambda/d_{\rm sub}$ (30.9 pixels in the *H* band, 41.5 in the K_s band).

5.1. Pinhole Stability and Target Alignment

Images of the pinhole were taken at several periods throughout the run from which its location on the detector can be measured. The location of the pinhole remained accurate to less than two pixels for the entire run, with the exception of two instances in which the Slit wheel appeared to lodge the pinhole at an alternative location. This was easily repaired by putting the Slit wheel into its "HOME" position.

Explicit (direct) images of the target within the pinhole were taken only rarely. However, the locations of the interferogram centers provided an accurate location of the binary targets on the detector. Drift of the target was typically less than one-half pixel during a set of 20 images (about 2 minutes). Comparison of the target and pinhole locations indicate that our alignment was accurate to within four pixels of the measured pinhole center. We conclude that the largest potential source of misalignment error arises from the initial pointing accuracy within the pinhole, and not jitter of the target or pinhole during data taking.

The effect of the pinhole was seen to alter several observables. Most directly, when observing reference stars, the pinhole induced a drop of target flux by roughly 35% in H and 15% in K_s. The pinhole is located after the AO system in the focal plane, hence these values reflect the percentage of the direct imaging PSF which falls outside the pinhole. These percentages are consistent with the pinhole size and the level of AO performance (Strehl ratios of approximately 10% in H and 30% in K_s). Binary targets show a similar drop in flux despite their larger size, which we can take as an indication that their center of light is well aligned.

Simulations and calculations using a sample of direct images show that the percentage of blocked flux is a weak function of the target alignment, and so large flux variations do not indicate a high level of astrometric jitter. Instead, this percentage is a strong function of the AO performance (and size of direct imaging halo). Low transmission (below 40%) well identified individual images blurred by temporary drops in AO correction. High flux transmission well predicted good fits to models (e.g., Figure 10).

5.2. Calibrators: Pinhole Filtering Produces Lower Closure Phase Variance and Higher Amplitudes

The use of the pinhole generally reduced the variance of closure phases and increased the amplitudes for the observed single stars. These results are compared to a set of simulated observations in Figure 11.

The closure phase standard deviation is reduced by 10% and 19% in the H band (1.6 μm) and the K_s band (2.2 μm), respectively. A larger reduction is expected at the longer wavelength, as the pinhole is relatively smaller and provides more aggressive spatial filtering. Although not displayed, the rms residuals of the data fit to their model values also reduced by 10%–20% indicating better calibrated data. The pinhole filtering increases the visibility of log-amplitude by 14% and 18% on the longest baselines at the H band and the K_s band, respectively. Each of these confirm that the spatial filter is reducing closure phase noise from sub-aperture redundancy.

In both cases the simulation predicts a larger reduction in variance, although the results across wavelengths are consistent. The simulation does not include the time-variation of AO residuals during a single exposure, particularly slow tip-tilt correction of the PALAO system (Bloemhof et al. 2001). This can produce several interferograms in a data set with large closure phase errors which cannot be improved by spatial filtering.

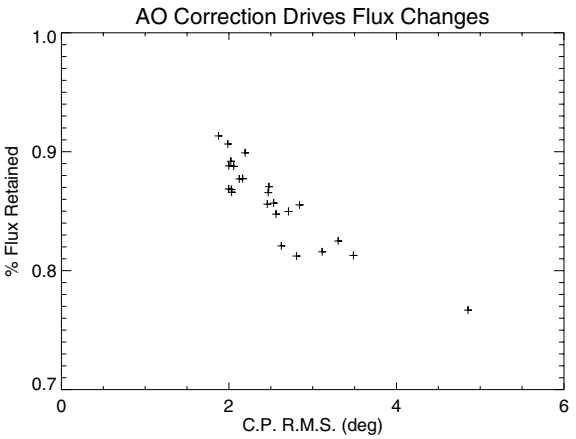


Figure 10. Drops in flux transmission through the pinhole are driven by changes in the PSF, which are tied to AO performance. The binary GJ 623 was resolved in K_s using 25 masking images through the pinhole. The images that produced the best-fitting closure phases (as measured by rms deviation from the model) also had the least flux blocked by the mask. Poor correction (lower Strehl ratio) displaces more flux into the outer halo of the PSF, which is then blocked by the pinhole. Poor correction also leads to larger closure phase errors. Thus, these two effects are strongly correlated. This trend is not caused by misalignment or movement of the target within the pinhole (see the text), but rather changes in AO correction (and lower Strehl ratio).

5.3. Binaries: Lower Closure Phase Variance

We observed four previously characterized binaries with well-defined orbits with and without the pinhole in several bandpasses: GJ 164 (Martinache et al. 2009), G 78-28 (Pravdo et al. 2006), GJ 623 (Martinache et al. 2007), and GJ 802B (Ireland et al. 2008). The characterized orbits were used to predict the location of each companion on the observing date using a Monte Carlo simulation to account for the uncertainty of the orbital parameters. Each target was observed to obtain 20 images (6 second exposures) in several bandpasses in 2009 September. Three of the four known binaries were successfully resolved (Table 1). The high-contrast companion to GJ 802B could not be resolved at the correct separation and contrast ratio. The quality of each set can be quantified by the rms of the difference between the data closure phases and the best-fitting closure phase model of a binary. The astrometry of the best-fitting binary model can also be compared to the astronomy predicted by the orbital parameters in the literature.

GJ 164 was imaged for three sets near dawn with the pinhole and one set without. All but the first sets of data suffered poor AO correction, were visibly less sharp, and had flux levels drop by nearly 75%. The binary could be well resolved in all three sets with the pinhole, but the latter sets yield much worse fits and larger systematic errors and are not used in the analysis. The fit parameters to GJ 164 share a degeneracy with a spurious set of parameters ($\Delta H = 0.030$, at nearly the same location), and limit the quality of parameter errors which can be derived from the fit. Instead, the magnitude contrast was held fixed to the values determined by Martinache et al. (2007). The observations of reference stars to calibrate GJ 623 varied in AO correction and quality between pinhole and non-pinhole measurements, and so this target is presented without calibration in order to compare the two performances.

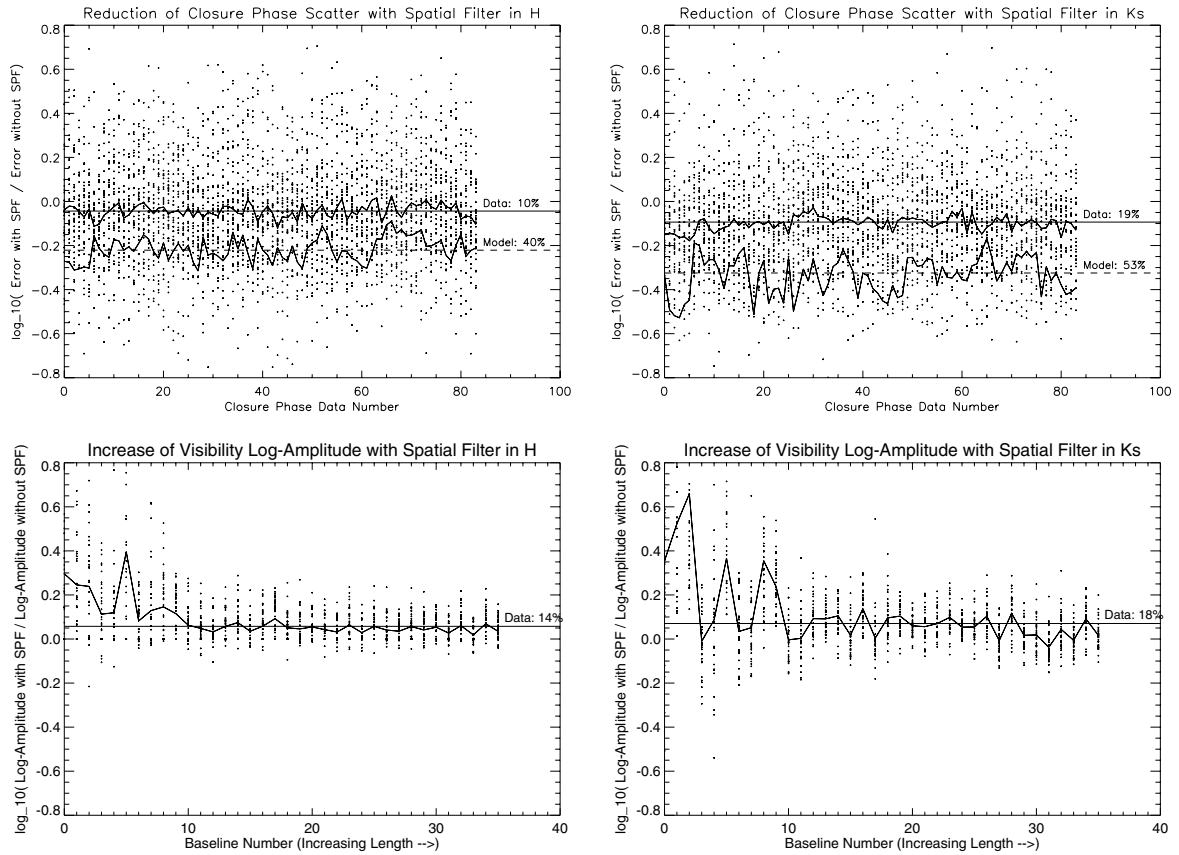


Figure 11. Closure phase standard deviation (scatter) is reduced and baseline visibility amplitude is increased when observed through the pinhole filter. Data points are drawn from observations of 26 single stars. Horizontal lines are the median of the data (solid) and the simulated experiment (dashed). Top row: closure phase scatter is reduced by 10% and 19% in H- and K_s-band measurements, respectively. Bottom row: visibility amplitude is increased by 14% and 18% in H and K_s bands, respectively. In all cases, the simulation (model) predicts a larger reduction in noise (see the Discussion).

The median closure phase scatter (stochastic errors) decreased for each observation by roughly 20%–25%, which indicates that the pinhole operated effectively to minimize closure phase errors from AO residuals. However, the data fits to binary models showed no decrease in rms residuals, whereas the fits to single-star data decreased by 10%–20%. Of the eight total measurements, four showed an increase in rms, although only two K_s-band observations increased by more than 10%. The observations that increased in rms are those in which either the companion was aligned far from the pinhole center or the companion is comparatively bright. We prefer to discuss this measurement in terms of the amount of unattributed systematic error that needs to be added in quadrature to the stochastic errors to yield the rms values, i.e., $(\mathrm{rms})^2 = \langle \sigma_\phi^2 + \sigma_{\mathrm{sys}}^2 \rangle$. This is listed in the next to last column of Table 1. A systematic component of 1.°0 specific to binary targets would account for the discrepancy of performance.

Misalignment of the target within the pinhole contributes to the increased residuals. Because each binary system could be resolved by direct imaging, the targets were aligned with their center of light as close to the pinhole center as could be accurately judged by eye. This necessarily placed each star off-center of the pinhole. The astrometric alignment of each component can be inferred from the center of the interferograms and is listed in Table 2. As discussed earlier, the alignment of the components is determined by the original placement of the observer and moves comparatively little during a set of images.

We simulated each binary with and within the pinhole displaced by the amounts in Table 2. The atmospheric seeing was tuned to fit Strehl ratio estimates from direct images taken throughout the night: 40% in K_s and Brγ bands, 15% in H and $CH4_s$ bands, and 10% in the J band. The astrometric jitter of the simulated images was consistent to those observed (and less than 20 mas per 20 images for most targets). Each simulation produced enough images until all errors reached a steady state; the estimates of the misalignment error are included in the last column of Table 1. Additional contributions from misaligned calibrators are not included.

Our calculation of the misalignment error requires knowledge of the absolute positions of the pinhole and target components, and an accurate measure of the PSF. Given the small jitter of targets, our calculation is limited by imperfect knowledge of the PSF. Even still, the quality of correction can fluctuate in timescales of minutes, as can be observed by viewing successive direct images or other observables such as pinhole transmission (Figure 10). Given these assumptions, misalignment contributes systematic components of 0.°4–0.°8 to the H- and K_s-band binary observations. We also note that simulations demonstrate large $\gtrsim$1.°0 alignment errors for slightly lower Strehl ratios ($\lesssim$10% in H and $\lesssim$20% in K_s). These challenges instead warrant the

Table 1
Observation of Known Binaries with and without Spatial Filter

Binary	Band	Separation (mas)	Position Angle (°)	Contrast Δ Mag	Best-fit rms (°)	Median σ_ϕ(°)	Cal	Estimated ϕ_{sys}(°)	Estimated σ_{align}(°)
GJ 164	H	48.7 ± 1.9	346.5 ± 1.7	[d]	4.03	2.11	Y	3.4	
	H (SPF)	48.8 ± 1.5	343.9 ± 0.9	[d]	3.33	1.99	Y	2.7	0.61
	Predicted[a]	49.4 ± 2.2	348.1 ± 10.1	1.835 ± 0.006					
	K_s	46.8 ± 1.2	340.3 ± 1.0	[d]	3.46	1.20	Y	3.2	
	K_s (SPF)	46.3 ± 0.7	341.2 ± 1.2	[d]	2.16	0.70	Y	2.9	0.42
	Predicted[a]	49.4 ± 2.2	348.1 ± 10.1	1.721 ± 0.097					
G 78-28	H	101.7 ± 0.3	359.2 ± 0.2	1.275 ± 0.010	2.78	1.16	Y	2.5	
	H (SPF)	102.6 ± 1.4	359.6 ± 0.2	1.323 ± 0.015	3.00	1.09	Y	2.8	0.68
	Predicted[b]	105.0 ± 5.7	359.3 ± 1.8	1.24 ± 0.07					
	K_s	103.4 ± 0.3	359.0 ± 0.1	1.221 ± 0.007	2.68	0.72	Y	2.6	
	K_s (SPF)	104.5 ± 0.2	359.3 ± 0.1	1.221 ± 0.005	3.60	0.56	Y	3.6	0.75
	Predicted[b]	105.0 ± 5.7	359.3 ± 1.8	1.14 ± 0.06					
	J	101.4 ± 0.6	358.7 ± 1.3	1.257 ± 0.021	7.60	3.12	Y	6.9	
	J (SPF)	100.8 ± 0.6	358.5 ± 0.2	1.263 ± 0.011	5.80	2.55	Y	5.2	0.24
	Predicted[b]	105.0 ± 5.7	359.3 ± 1.8	1.24 ± 0.07					
GJ 623	$CH4_s$	282.3 ± 1.4	176.3 ± 0.4	2.798 ± 0.050	3.38	1.90	N	2.8	
	$CH4_s$ (SPF)	277.3 ± 3.0	176.2 ± 0.2	2.829 ± 0.51	3.43	1.65	N	3.0	0.75
	Predicted[c]	279.1 ± 1.2	177.4 ± 0.7	2.860 ± 0.039					
	K_s	277.1 ± 0.4	176.5 ± 0.1	2.675 ± 0.014	2.35	0.45	N	2.3	
	K_s (SPF)	280.7 ± 0.2	176.6 ± 0.7	2.780 ± 0.10	1.9	0.30	N	1.9	0.46
	Predicted[c]	279.1 ± 1.2	177.4 ± 0.7	2.720 ± 0.014					
	Brγ	279.1 ± 0.3	176.2 ± 0.7	2.664 ± 0.012	0.94	0.42	N	0.8	
	Brγ (SPF)	279.4 ± 0.4	176.5 ± 0.1	2.778 ± 0.009	1.28	0.33	N	1.2	0.63
	Predicted[c]	279.1 ± 1.2	177.4 ± 0.7	2.720 ± 0.014					

Notes. Measurements of three known binary systems with and without the pinhole filter (SPF).
[a] Spatial filtering reduces the median scatter of closure phase measurements in all cases.
[b] The median residual between the data and best-fitting model does not increase when using the pinhole.
[c] An estimate of the systematic error due to a misalignment of 3 pixels. See the text for discussion.
[d] The target GJ 164 data were fit to a binary model holding fixed the mean contrast ratio given by Martinache et al. (2009). Note: Target GJ 802 is excluded from this table.

development of an observing strategy to calibrate alignment errors empirically.

The binary parameters measured with and without the pinhole, and those predicted by the system orbits, are in good agreement. In each case, the spatial filter data fit to slightly higher contrasts ratios (≈5%). This may indicate that the companion flux was skewed closer to the pinhole edge and its flux truncated by a few percent more, which is consistent with the measured locations of the objects within the pinhole (Table 2). No bias was found in the relative astrometry.

Generally speaking, the lower closure phase errors brought on by the pinhole filter ought to lead to proportionally lower errors in binary parameters and higher contrast of detection. The pinhole demonstrably decreases closure phase measurement error by 20%–25% and indicates that, if the systematic component can be minimized, the pinhole can provide an increase of contrast by 20%–25% using the same length of observation. However, the misalignment error limits the increase in precision. In practice, we scale the closure phase measurement errors until the χ^2 of the best fit is unity to account for any known systematic errors. Therefore, to maintain the increased precision in this analysis, the decrease in closure phase measurement error must be matched by a similar decrease in rms fit residual (or systematic component). Since the fit residuals did not improve, the precision in the binary parameters did not improve in all cases.

Table 2
Astrometry and Alignment of Targets within Pinhole

		Distance From Center (mas)		Estimated
Binary	Band	Primary	Secondary	σ_{align}
GJ 164	H	90 ± 50 (60)	140 ± 50 (60)	0.61
	K_s	50 ± 15 (40)	90 ± 15 (40)	0.42
G 78-28	H	90 ± 35 (50)	110 ± 35 (50)	0.68
	K_s	80 ± 10 (40)	70 ± 10 (40)	0.75
	J	100 ± 25 (45)	170 ± 23 (45)	0.24
GJ 623	$CH4_s$	130 ± 10 (35)	200 ± 10 (35)	0.75
	K_s	90 ± 10 (40)	200 ± 10 (40)	0.46
	Brγ	150 ± 10 (40)	190 ± 10 (40)	0.63

Notes. Alignment of targets within pinhole and estimated closure phase misalignment error. Position determined by center of interferograms; errors estimated from spread over 20 images. Values in parentheses include 40 mas uncertainty of the absolute pinhole position. Misalignment errors are calculated using the simulation of Section 4.3, assuming a Strehl of 15% in H and $CH4_s$, 45% in K_s and Brγ, and 10% in the J band.

6. SUMMARY OF RESULTS AND CONCLUSIONS

In this text, we discussed spatial wavefront variation on the scale of NRM sub-apertures, which contributes to closure phase errors we called sub-aperture redundancy noise. We proposed that quasi-static wavefront errors on the scale of the

sub-aperture, which evolve over hundreds of seconds or with telescope movement, give rise to the limitations of aperture masking calibration. Through simulations and direct observation we have shown that the use of a pinhole filter effectively reduces sub-aperture spatial variation and can reduce closure phase errors, increase visibility amplitudes, and lead to better calibration and higher detection contrasts.

With a simple premise, we estimate the optimal pinhole size to be $d_{\rm spf} \approx D/d_{\rm sub}$ in units of λ/D, or about $12\lambda/D$ for the Palomar nine-hole mask. We confirm this result using a detailed simulation of the Palomar aperture masking experiment and show that the operational pinhole size ranges from 11 to 14 λ/D. Smaller pinholes reject enough light from off-axis companions to alter the measured closure phases. Larger pinholes provide little spatial filtering.

Because the pinhole is not a perfect low-pass filter, the truncation of off-axis targets aliases power back into phase at sub-aperture scales. This introduces closure phase errors of its own, both systematic and stochastic. Using simulated observations over a wide range of corrections (Strehl ratio of 10%–100%), we show that a misalignment of the target by 200 mas introduces $0\overset{\circ}{.}3$–$1\overset{\circ}{.}0$ of systematic closure phase, which we have termed misalignment error. We propose several methods for the calibration of this component although none are undertaken in this experiment. As a result, the effectiveness of pinhole technique is limited by the accuracy with which targets can be aligned within the center of the pinhole.

We installed a pinhole of size 0.779 arcsec ($12\lambda/D$ at 1.6 μm) in the Slit wheel focal plane of the PHARO instrument on the Hale 200″ Telescope at Palomar Observatory. Using this pinhole we observed twenty single unresolved stars and three well-characterized binaries with and without the pinhole in a selection of standard near-infrared bands. The AO system provided moderate corrections: Strehl ∼5%–15% in the H band and ∼20%–40% in the K_s band.

Observations of unresolved stars showed that the pinhole reduced the stochastic variation of closure phases by 10%–25% from one image to the next, and visibility amplitudes increased. rms residuals to model fits decreased by a similar factor. Each shows that increasing sub-aperture coherence decreases closure phase redundancy noise. Lower fit residuals indicate that calibration improved, and that the aperture masking calibration limit is partially attributable to small-scale quasi-static wavefront errors. Furthermore, detection contrasts scales proportional to fit residuals, indicating an increase in detection contrast of 10%–25%.

The modest increase in closure phase signal to noise when the pinhole is used suggests that the spectrum of wavefront errors contains relatively little power on the smallest spatial-scales within the sub-aperture. Additionally, the smallest scale variations that cycle multiple times within a single sub-aperture tend to produce redundancy noise that averages out. The leading source of spatial redundancy noise, both quasi-static instrumental and residual atmospheric wavefront errors, is more likely low-order Zernike modes (i.e, "phase slopes") across the sub-apertures.

Simulations of these observations, when compared, predict a slightly larger benefit from the pinhole filtering. As discussed in Section 2.2, temporal variations of the baseline phase within a single exposure lead to a sort of temporal redundancy noise that cannot be removed by closure phases of the pinhole filter. Furthermore, the slow response of the Palomar AO tip-tilt mirror (5 Hz) gives rise to full aperture scale wavefront residuals which evolve on the order of a fraction of an exposure. The baseline phase changes these residuals induce contribute to the overall variation of the closure phases from one image to the next (though not calibration errors). The simulation did not include temporal variations of the AO residuals, and as a result gives the impression that the pinhole is more effective at reducing stochastic closure phase noise (50% versus 20%). Indeed, these results imply that temporal variation of the wavefront phase is a large source of closure phase variance with the current PALAO system.

We observed three well-characterized binary systems to determine how the pinhole influences the astrometric and photometric characterizations of discovered systems. The measured binary parameters were consistent with those acquired without the pinhole and those predicted by the measured orbital parameters of the system. No bias was found in the astrometry (i.e., binaries did not tend toward shorter separations with the pinhole); however, contrast ratios with the pinhole tended to be a few percent fainter. Our alignment technique placed the system center of light near the pinhole center, and the light of the companion is preferentially blocked by this arrangement. Better positioning of the target would likely remedy this bias.

Although closure phase stochastic errors decreased by 20%–25%, binary data acquired with the pinhole provided no better fits to data and suggest that binary targets observed with the pinhole contain an additional $1\overset{\circ}{.}0$ systematic error. After simulating these observations with and without the pinhole, we conclude that the systematic component arises from the off-center alignment of a resolved target. Simulations demonstrate large $\gtrsim 1\overset{\circ}{.}0$ alignment errors at low Strehl ratios ($\lesssim$10% in the H band and $\lesssim$20% in K_s).

Binary parameter errors and detection contrast both decrease in proportion to the closure phase errors, and indicate that one can achieve more precisely measured binary parameters and higher contrasts (20%–25%) with the pinhole.

Finally, we showed that multiplying the images by an apodized window function acts to reduce the effects of wavefront error similar to a spatial filter. Applying this post-processing method for spatial filtering allows easy tailoring of the shape of the filtering function, but only filters some form of the time-accumulated wavefront phase noise (not instantly). We found that the optimal window function has a HWHM of about $\lambda/d_{\rm sub}$, with smaller window functions favored for targets in which read noise is higher. By comparing the results of the two methods, it is clear that on-instrument filtering is superior because of the substantial temporal variation of the wavefront phase noise on sub-aperture scales.

7. DISCUSSION

7.1. A Strategy for Future Pinhole Observations

In light of this investigation, the pinhole can prove to be effective in two scenarios. When seeing or correction is poor, and closure phase observations do not reach the calibration limit, a pinhole filter will clearly provide higher signal-to-noise closure phases and more efficient data taking.

At the other extreme, the pinhole may be valuable for improving very high Strehl ratio ($\gtrsim$80%) observations with extreme-AO systems. The high spatial frequencies filtered by the pinhole contribute to a larger proportion of the overall wavefront errors in this scenario. The pinhole could be optimally tuned to the AO actuator size, much larger, and alignment requirements would be substantially relaxed. Most importantly, the aliasing of high spatial frequency content into sub-aperture scale phase

errors is a higher order effect of the wavefront (Poyneer & Macintosh 2004). At very high levels of correction, the aliased power is a smaller component of closure phase errors.

We can arrive at a rough estimate of a pinhole performance on extreme-AO systems by simulating a higher-order AO system, such as the 3368-actuator Palomar 3000 system (P3K; Dekany et al. 2007). As mentioned earlier, our simulation does not account for the sluggish tip-tilt correction of the current Palomar system and overpredicted the pinhole performance as a result. The P3K system will have upgraded response by using the current 241 actuator system to perform low-order correction, and thus provide more stable correction. Using the specifications of this system, we simulated a 10:1 binary separated by 150 mas in the H band under typical conditions (see Section 4.2). A pinhole of size $12\lambda/D$ decreased stochastic closure phase errors to 20% nominal values, i.e., a five-fold increase in signal to noise. When compared to our simulations of the current AO system (a two-fold increase), the pinhole can impact extreme-AO observations favorably.

For very high Strehl ratio observations, misalignment error will be lower and less sensitive to changes in correction (0.°19 in this example). Empirically calibrating misalignment by observing a reference star at the same position as the primary is only a partial solution (as it does not take into account the companion), likely limited by the pointing accuracy of the observer, and decreases the overall efficiency of the observations. Instead, observations should be taken by dithering the center of light around the pinhole center to minimize this systematic component. Using diametrically opposed target placements successfully reduced the misalignment component to 0.°11 in simulation, a reduction of nearly half.

Finally, one must also consider the optimal pinhole size to survey across the near-infrared bands. For example, if the effective range of pinhole sizes is narrow enough (e.g., 11–14 λ/D for this experiment) a single pinhole may not accommodate multiple wide band infrared filters. In those cases, multiple pinholes tuned to specific observation bands will be most effective.

7.2. Extreme-AO Aperture Masking Experiments

Next generation AO systems with aperture masks, such as Palomar 3000 (Dekany et al. 2007) and Gemini Planet Imager (Macintosh et al. 2008), provide actuator densities high enough to correct wavefront errors on sub-aperture scales. Design and optimization of these aperture masking experiments that balance sub-aperture size and number against AO correction and the limitations of systematics require a more detailed understanding of how sub-aperture wavefront errors impact closure phases, and where resources should be focused. Extreme-AO systems are anticipated to provide highly stable correction down to scales of 8 cm and thus may motivate smaller and additional sub-apertures, providing more spatial frequency coverage and resolution. This balance will depend on the spatial and temporal power spectra of the corrected wavefront residuals. These data should be accumulated and their impact on closure phase errors should be examined.

We thank the staff and telescope operators of Palomar Observatory for their support and many nights of assistance at the Hale Telescope. The authors appreciate the effort of the reviewers, who provided valuable advice on the clarity and depth of this text. This material is based upon work supported by the National Science Foundation under Grant No. AST-0705085. The Hale Telescope at Palomar Observatory is operated as part of a collaborative agreement between Caltech, JPL, and Cornell University. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, and is funded by the National Aeronautics and Space Administration and the National Science Foundation.

Facility: Hale

APPENDIX

PINHOLE FILTERING: INTEFEROMETRY

Observing through a pinhole alters the measured complex visibility of a point source even in the absence of any wavefront aberrations. This analysis assumes that the Fraunhofer approximation applies, i.e., that the electric field in the image plane is the Fourier transform of the phasor of the wavefront phase in the pupil frame. We assume no wavefront aberrations, including those from optical errors and central obscurations.

The optical path is shown in Figure 3. We denote the electric field at the telescope aperture with the subscript a and the electric field after spatially filtering with the subscript spf. The coordinates in the image plane are $\boldsymbol{\theta}$; all angles are measured in units of λ/D. We assume that the pinhole has angular diameter $d_{\rm spf}$ and is positioned at $\boldsymbol{\theta} = 0$. The point source is located at an angular position $\boldsymbol{\alpha}$ on the sky and $\boldsymbol{\alpha} = 0$ corresponds to perfect alignment of the source in the pinhole.

The wavefront phasor at the entrance of the aperture is

$$E_a(\boldsymbol{x}) = \Pi(\boldsymbol{x}/D)e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{x}/D},$$
$$\text{with } \begin{array}{ll}\Pi(\boldsymbol{x}/D) = 1 & \text{for } |\boldsymbol{x}/D| < \dfrac{1}{2},\\ \Pi(\boldsymbol{x}/D) = 0 & \text{elsewhere.}\end{array} \tag{A1}$$

The optics of the spatial filter truncate the electric field in the focal plane, which results in a convolution when the field is transformed back at a pupil plane. The spatially filtered electric field is then

$$E_{\rm spf}(\boldsymbol{x}) = [\Pi(\boldsymbol{x}/D)e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{x}/D}] \star \left[\frac{2J_1(\pi\boldsymbol{x}d_{\rm spf}/D)}{(\pi\boldsymbol{x}d_{\rm spf}/D)}\right], \tag{A2}$$

where the star represents convolution.

The use of an aperture mask, $M(\boldsymbol{x})$, containing a baseline $\boldsymbol{b}$ ultimately measures the spatial frequency $\boldsymbol{u} = \boldsymbol{b}/\lambda$ of the complex visibility. For clarity, we present this as a function of baseline:

$$V(\boldsymbol{b}) = \int d\boldsymbol{x}[M(\boldsymbol{x})M^*(\boldsymbol{x}+\boldsymbol{b})][E(\boldsymbol{x})E^*(\boldsymbol{x}+\boldsymbol{b})], \tag{A3}$$

where $E(\boldsymbol{x})$ is the electric field sampled by the mask.

Without the spatial filter, substitution for the field from Equation (A1) gives

$$V(\boldsymbol{b}) = T(\boldsymbol{b})e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{b}/D}, \quad \text{with } T(\boldsymbol{b}) = \int d\boldsymbol{x}M(\boldsymbol{x})M^*(\boldsymbol{x}+\boldsymbol{b}). \tag{A4}$$

With the spatial filter, the aperture mask samples $E_{\rm spf}(\boldsymbol{x})$ of Equation (A2). By recognizing that convolution with $e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{x}/D}$ leads to a Fourier transform, doing so before substitution yields

$$V_{\rm spf}(\boldsymbol{b}) = T(\boldsymbol{b})\int d\boldsymbol{\theta}\Pi[\boldsymbol{\theta}/d_{\rm spf}] \times e^{2\pi i\boldsymbol{\theta}\cdot\boldsymbol{b}/D} \times \left[\frac{2J_1(\pi|\boldsymbol{\theta}-\boldsymbol{\alpha}|)}{\pi|\boldsymbol{\theta}-\boldsymbol{\alpha}|}\right]^2. \tag{A5}$$

In the limit of an infinite pinhole, the integral converges to the value $T(\boldsymbol{b})e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{b}/D}$.

We can extend use of this approach to include arbitrary pupil-plane phase errors which arise prior to propagation through the pinhole system. In the presence of inhomogeneous phase in the pupil plane, $e^{i\phi(x)}$, Equation (A2) becomes

$$E_{\rm spf}(\boldsymbol{x}) = [\Pi(\boldsymbol{x}/D)e^{2\pi i\boldsymbol{\alpha}\cdot\boldsymbol{x}/D}e^{i\phi(\boldsymbol{x})}] \star \left[\frac{2J_1(\pi \boldsymbol{x} d_{\rm spf}/D)}{(\pi \boldsymbol{x} d_{\rm spf}/D)}\right]. \tag{A6}$$

Repeating the steps above, the complex visibility is written as

$$V_{\rm spf}(\boldsymbol{b}) = T(\boldsymbol{b}) \int d\boldsymbol{\theta}\Pi[\boldsymbol{\theta}/d_{\rm spf}] \times e^{2\pi i\boldsymbol{\theta}\cdot\boldsymbol{b}/D} \times {\rm PSF}[\boldsymbol{\theta} - \boldsymbol{\alpha}], \tag{A7}$$

with

$${\rm PSF}[\boldsymbol{\theta}] = \left|\int d\boldsymbol{x} e^{2\pi i\boldsymbol{\theta}\cdot\boldsymbol{x}/D}\Pi(\boldsymbol{x}/D)e^{i\phi(\boldsymbol{x})}\right|^2. \tag{A8}$$

REFERENCES

Bernat, D., Bouchez, A. H., Ireland, M., et al. 2010, ApJ, 715, 724
Bloemhof, E. E., Dekany, R. G., Troy, M., & Oppenheimer, B. R. 2001, ApJ, 558, L71
Cornwell, T. J. 1989, Science, 245, 263
Dekany, R., Bouchez, A., Roberts, J., et al. 2007, in Proc. Advanced Maui Optical and Space Surveillance Technologies Conference, ed. S. Ryan (Maui, HI: Maui Economic Development Board), 64
du Foresto, V. C., Ridgway, S., & Mariotti, J.-M. 1997, A&AS, 121, 379
Fabrycky, D. C., & Murray-Clay, R. A. 2010, ApJ, 710, 1408
Ferrari, A. 2007, ApJ, 657, 1201
Glindemann, A., Lane, R. G., & Dainty, J. C. 1993, J. Mod. Opt., 40, 2381
Haniff, C. A., Mackay, C. D., Titterington, D. J., Sivia, D., & Baldwin, J. E. 1987, Nature, 328, 694
Hayward, T. L., Brandl, B., Pirger, B., et al. 2001, PASP, 113, 105
Hinkley, S., Oppenheimer, B. R., Soummer, R., et al. 2007, ApJ, 654, 633
Hinkley, S., Oppenheimer, B. R., Soummer, R., et al. 2009, ApJ, 701, 804
Hinkley, S., Oppenheimer, B. R., Zimmerman, N., et al. 2011, PASP, 123, 74
Ireland, M. J., Kraus, A., Martinache, F., Lloyd, J. P., & Tuthill, P. G. 2008, ApJ, 678, 463
Jennison, R. C. 1958, MNRAS, 118, 276
Keen, J., Buscher, D., & Warner, P. 2001, MNRAS, 326, 138
Kraus, A., Covey, K., Liu, M., et al. 2009, Astro2010: The Astronomy and Astrophysics Decadal Survey, Science White Papers, No. 163
Kraus, A., Ireland, M., Martinache, F., & Lloyd, J. 2008, ApJ, 679, 762
Kraus, A. L., Ireland, M. J., Martinache, F., & Hillenbrand, L. A. 2011, ApJ, 731, 8
Lane, R. G., Glindemann, A., & Dainty, J. C. 1992, Waves Random Media, 2, 209
Linfield, R. P., Colavita, M. M., & Lane, B. F. 2001, ApJ, 554, 505
Lloyd, J. P., Martinache, F., Ireland, M. J., et al. 2006, ApJ, 650, L131
Lohmann, A. W., Weigelt, G., & Wirnitzer, B. 1983, Appl. Opt., 22, 4028
Macintosh, B. A., Graham, J. R., Palmer, D. W., et al. 2008, Proc. SPIE, 7015, 701518
Marois, C., Doyon, R., Racine, R., et al. 2005, J. R. Astron. Soc. Can., 99, 130
Martinache, F., Lloyd, J. P., Ireland, M. J., Yamada, R. S., & Tuthill, P. G. 2007, ApJ, 661, 496
Martinache, F., Rojas-Ayala, B., Ireland, M. J., Lloyd, J. P., & Tuthill, P. G. 2009, ApJ, 695, 1183
Poyneer, L., & Macintosh, B. 2004, J. Opt. Soc. Am. A, 21, 810
Pravdo, S. H., Shaklan, S. B., Wiktorowicz, S. J., et al. 2006, ApJ, 649, 389
Racine, R., Walker, G., Nadeau, D., Doyon, R., & Marois, C. 1999, PASP, 111, 587
Raymond, S. N., Barnes, R., Veras, D., et al. 2009, ApJ, 696, L98
Readhead, A., Nakajima, T., & Pearson, T. 1988, AJ, 95, 1278
Roddier, F. 1986, Opt. Commun., 60, 145
Shaklan, S. B., & Roddier, F. 1987, Appl. Opt., 26, 2159
Sivaramakrishnan, A., Koresko, C. D., Makidon, R. B., Berkefeld, T., & Kuchner, M. J. 2001, ApJ, 552, 397
Tatulli, E., Blind, N., Berger, J. P., Chelli, A., & Malbet, F. 2010, A&A, 524, 65
Tuthill, P., Lloyd, J., Ireland, M., et al. 2006, Proc. SPIE, 6272, 103
Veras, D., Crepp, J. R., & Ford, E. B. 2009, ApJ, 696, 1600
Ziad, A., Schöck, M., Chanan, G. A., et al. 2004, Appl. Opt., 43, 2316

IF YOU WANT TO PLAY IT SAFE ALL THE TIME, GO JOIN AN INSURANCE COMPANY.

ANDY WEIR, *THE MARTIAN*

Le projet SCExAO

Imagerie directe des planètes extrasolaires

L'observation en mode interférométrique à masque non redondant requiert des images non-saturées. Dès qu'une des sources du champ pointé sature localement le détecteur, l'image ne peut en effet plus exactement être décrite comme le résultat d'un produit de convolution (voir equation 1), ce qui en rend l'analyse par transformée de Fourier inutile. Le potentiel de détection à haut contraste de cette approche est donc fortement contraint par la dynamique du détecteur utilisé. En pratique, avec des limites de détection en contraste de six magnitudes (de l'ordre de 2×10^{-3}), les chances d'observer des planètes extrasolaires avec cette technique restent limitées à l'observation d'objets particulièrement jeunes, tels que le candidat compagnon planétaire dans le voisinage de l'étoile de type T-Tauri LkCa 15, membre de l'association Tau-Aur [21], à une distance orbitale d'environ 15 UA (voir la figure 22). Pour devenir sensible à la présence de compagnons à plus haut contraste, il faut changer d'approche et rajouter des optiques qui vont permettre de changer de gamme de dynamique. La contrepartie est que l'on ne peut (pour le moment) alors plus bénéficier des propriétés auto-étalonnées d'observables comme les clôtures de phase.

21. A. L. Kraus and M. J. Ireland. LkCa 15: A Young Exoplanet Caught at Formation? **ApJ**, 745:5, Jan. 2012

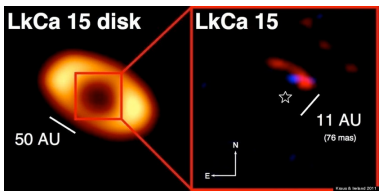


Figure 22: Image interférométrique de LkCa 15. A gauche : image submillimétrique, révélant la présence d'un disque. A droite : image IR (bande K (bleu) et L (rouge) de la partie intérieure du disque. Les structures persistantes observées dans ces images suggèrent l'existence d'un compagnon planétaire en cours d'accrétion.

La détection directe par imagerie des exoplanètes reste pour le moment une contribution mineure au grand nombre de découvertes qui ont été faites depuis 1995 (73 sur 3541). La difficulté est due au fait qu'en plus de résolution angulaire, ce type de détection requiert une très grande dynamique, entre 10^{-6} et 10^{-10} selon les applications, qui reste encore un véritable challenge technologique. Si l'on regarde comment sont distribués les demi-grand axes des $\sim$2000 planètes pour lesquelles il est connu, ce qui est présenté dans la figure 23, on voit que, malgré les biais observationnels de la vélocimétrie radiale et des transits photométriques, l'immense majorité des planètes attendues sont sur des orbites inférieures à 10 UA. Ceci suggère que pour maximiser le retour scientifique des campagnes d'observation de planètes extrasolaires, il faut mettre en oeuvre des solutions techniques permettant de sonder à haut contraste les très faibles séparations angulaires.

Les images spatialement résolues de ces exoplanètes permettent

d'aller au delà de la détection d'objets : une unique image offre en effet la possibilité d'appréhender une ou plusieurs planètes comme éléments constituants d'un système (voir des exemples en figure 24). La possibilité d'isoler la contribution lumineuse d'un compagnon planétaire dans une image permet enfin d'envisager en faire une analyse spectroscopique, qui nous permettra de statuer sur l'environnement physique que leurs éventuelles surfaces offrent. Bien que l'interférométrie couvre finalemant assez bien le domaine de séparation angulaire correspondant au gros de la distribution des demi-grand axes, on sait qu'en pratique, avec des limites de détection en contraste de six magnitudes (de l'ordre de 2×10^{-3}), les chances d'observer des planètes extrasolaires avec cette technique restent limitées à l'observation d'objets particulièrement jeunes, tel que LkCa 15.

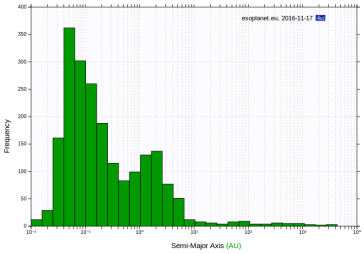


FIGURE 23: Distribution des demi-grand axes des orbites des planètes extrasolaires connues à la mi-novembre 2016. L'axe horizontal, en échelle logarithmique, couvre un domaine allant de 10^{-2} à 10^{4} UA.

La coronographie

LA CORONAGRAPHIE S'IMPOSE rapidement comme l'alternative logique à l'interférométrie pour atteindre le haut contraste. La coronographie est maintenant une technique bien comprise, qui a considérablement évolué depuis le design original de Lyot, initialement adapté à l'observation de l'environnement Solaire et qui se décline aujourd'hui en un large panel de variantes, tels que l'APLC (Coronographe de Lyot à Pupille Apodisée), le PIAA [22] (Phase Induced Amplitude Apodization) dont il sera plusieurs fois question dans ce chapitre, et le coronographe vecteur-Vortex [23]. Le coronographe idéal est un système optique qui supprime l'intégralité de la lumière d'une source placée sur l'axe et laisse passer l'intégralité de la lumière des sources situées au delà d'une limite de séparation angulaire qui dépend de la solution technique retenue [24]. Cette limite, appelée IWA (Inner Working Angle), au vu des propriétés anticipées des planètes à imager, a intérêt à être aussi petite que possible. La position exacte de cette limite (conventionnellement placée lorsque la transmission atteind 50 % de son maximum), va dépendre du niveau de contraste offert par le coronographe, des spécificités de la pupille de l'instrument et de la largeur de la bande spectrale couverte.

La coronographie, de façon générale, est fortement dépendante de la qualité de la surface d'onde qui l'alimente : l'extinction coronagraphique haute performance n'est en effet possible que si le contrôle de front d'onde garantit que la lumière de la source à atténuer est bien dirigée vers les fonctions actives du système optique, en particulier le masque focal : toute aberration résiduelle se manifeste par de la lumière diffractée à l'extérieur du coeur de la PSF, qui n'est plus interceptée par le masque plan focal. Ainsi, s'il est possible sur le papier de concevoir un système optique avec une atténuation de 10 à 11 ordres de grandeur, garantir la performance de ce système en condition d'observation depuis le sol reste un challenge auquel les instruments récents ont seulement partiellement répondu.

22. O. Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging. **A&A**, 404:379–387, June 2003

23. D. Mawet, E. Serabyn, K. Liewer, R. Burruss, J. Hickey, and D. Shemo. The Vector Vortex Coronagraph: Laboratory Results and First Light at Palomar Observatory. **ApJ**, 709:53–57, Jan. 2010

24. O. Guyon, E. A. Pluzhnik, M. J. Kuchner, B. Collins, and S. T. Ridgway. Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs. **ApJS**, 167:81–99, Nov. 2006

Le haut contraste depuis le sol

Plusieurs instruments sol dédiés à l'imagerie directe de planètes et de disques sont maintenant en service [25] [26] . Deux gros projets en particulier : le "Spectro-Polarimetric High-contrast Exoplanet REsearch" (SPHERE) au VLT et le "Gemini Planet Imager" (GPI) sur Gemini Sud , occupent majoritairement le paysage scientifique de cette communauté. D'autres projets comme P1640 à Palomar [27] et l'instrument "Subaru Coronagraphic eXtreme Adaptive Optics" (SCExAO) [28], complémentent ces poids lourds par une approche plaçant la R&D au coeur de la philosophie de projet. La recette de fabrication de GPI et SPHERE a respecté une architecture très rationnelle, extrapolant le modèle de la génération précédente d'instruments imageurs post-OA, qui sépare deux fonctions distinctes :

1. une boucle de contrôle de front d'onde de haut ordre (dite adaptive optique extrême - ou XAO) dont la fonction est de stabiliser et d'améliorer la qualité de la surface d'onde, grâce aux informations produites par un analyseur de surface d'onde dédié.
2. une caméra d'imagerie haut contraste (équipée d'un ou plusieurs coronographes) dont le rôle est d'atténuer la source brillante sur l'axe et de révéler les structures faibles dans son voisinage direct.

Les incroyables images produites par ces instruments (voir figure 24) ayant conduit à la découverte de disques et de compagnons en orbite autour d'étoiles jeunes, semblent valider le modèle de conception. Il faut cependant rappeler que ces images, avec un contraste effectif de l'ordre de $\sim 10^6$ sont des produits qui ont bénéficié de traitements *a posteriori* sophistiqués permettant de regagner environ deux ordres de grandeur en contraste : les images brutes de ces instruments restent, malgré le contrôle de front d'onde par la boucle XAO et le coronographe, généralement spécifié pour un contraste brut plus élevé, encore essentiellement dominées par des speckles et structures de diffraction non-filtrées.

Seul un petit nombre de planètes réunit les conditions de masse, d'âge, de séparation orbitale et de distance à notre système Solaire, pour pouvoir être directement imagées aujourd'hui avec de telles performances. Malgré tout, grâce aux grands diamètres des télescopes d'aujourd'hui et de demain, les instruments au sol n'ont pas encore épuisé le filon. Je vais, dans la suite de ce chapitre, défendre l'idée qu'une amélioration significative des performances brutes des instruments est possible aujourd'hui, si le modèle de conception intègre mieux ensemble les fonctions de contrôle de surface d'onde et d'imagerie haut contraste. Cette meilleure intégration, se traduisant par un contrôle du front d'onde utilisant en plus de la XAO, le feedback de l'image focale scientifique, a fait l'objet de mon travail de R&D sur l'instrument SCExAO.

25. J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, J. Antichi, A. Baruffolo, P. Baudoz, A. Berton, A. Boccaletti, M. Carbillet, J. Charton, R. Claudi, M. Downing, P. Feautrier, E. Fedrigo, T. Fusco, R. Gratton, N. Hubin, M. Kasper, M. Langlois, C. Moutou, L. Mugnier, J. Pragt, P. Rabou, M. Saisse, H. M. Schmid, E. Stadler, M. Turrato, S. Udry, R. Waters, and F. Wildi. SPHERE: A 'Planet Finder' Instrument for the VLT. *The Messenger*, 125, Sept. 2006
26. B. Macintosh, J. R. Graham, P. Ingraham, Q. Konopacky, C. Marois, M. Perrin, L. Poyneer, B. Bauman, T. Barman, A. S. Burrows, A. Cardwell, J. Chilcote, R. J. De Rosa, D. Dillon, R. Doyon, J. Dunn, D. Erikson, M. P. Fitzgerald, D. Gavel, S. Goodsell, M. Hartung, P. Hibon, P. Kalas, J. Larkin, J. Maire, F. Marchis, M. S. Marley, J. McBride, M. Millar-Blanchaer, K. Morzinski, A. Norton, B. R. Oppenheimer, D. Palmer, J. Patience, L. Pueyo, F. Rantakyro, N. Sadakuni, L. Saddlemyer, D. Savransky, A. Serio, R. Soummer, A. Sivaramakrishnan, I. Song, S. Thomas, J. K. Wallace, S. Wiktorowicz, and S. Wolff. First light of the Gemini Planet Imager. *Proceedings of the National Academy of Science*, 111:12661–12666, Sept. 2014
27. B. R. Oppenheimer, C. Beichman, D. Brenner, R. Burruss, E. Cady, J. Crepp, L. Hillenbrand, S. Hinkley, E. R. Ligon, T. Lockhart, I. Parry, L. Pueyo, E. Rice, L. C. Roberts, J. Roberts, M. Shao, A. Sivaramakrishnan, R. Soummer, G. Vasisht, F. Vescelus, J. K. Wallace, C. Zhai, and N. Zimmerman. Project 1640: the world's first ExAO coronagraphic hyperspectral imager for comparative planetary science. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8447 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2012
28. N. Jovanovic, F. Martinache, O. Guyon, C. Clergeon, G. Singh, T. Kudo, V. Garrel, K. Newman, D. Doughty, J. Lozi, J. Males, Y. Minowa, Y. Hayano, N. Takato, J. Morino, J. Kuhn, E. Serabyn, B. Norris, P. Tuthill, G. Schworer, P. Stewart, L. Close, E. Huby, G. Perrin, S. Lacour, L. Gauchet, S. Vievard, N. Murakami, F. Oshiyama, N. Baba, T. Matsuo, J. Nishikawa, M. Tamura, O. Lai, F. Marchis, G. Duchene, T. Kotani, and J. Woillez. The Subaru Coronagraphic Extreme Adaptive Optics System: Enabling High-Contrast Imaging on Solar-System Scales. **PASP**, 127:890–910, Oct. 2015

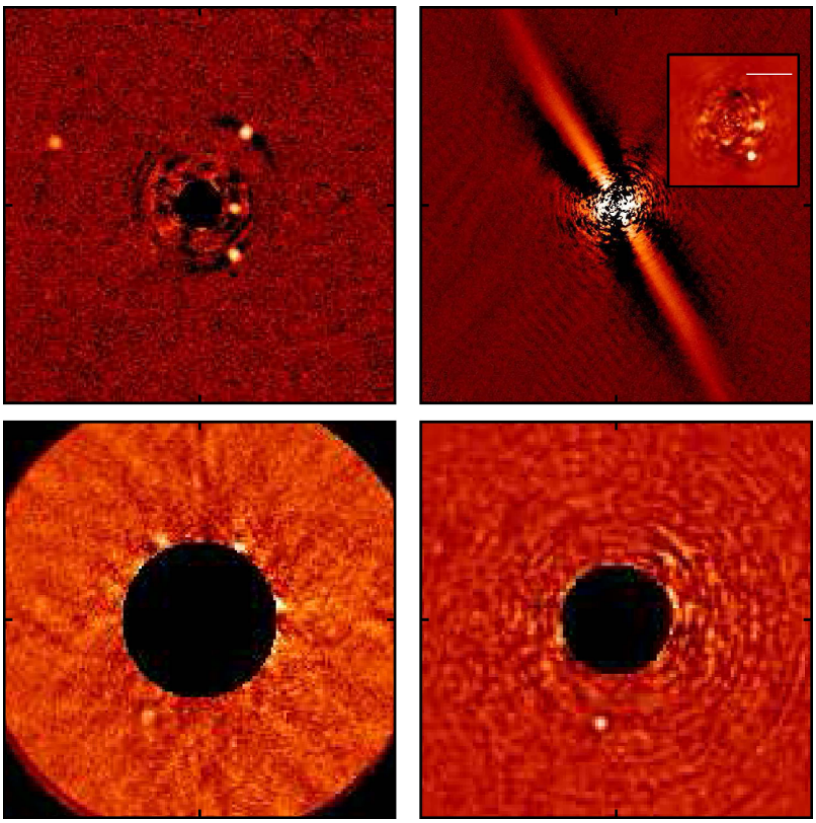


FIGURE 24: Etat de l'art de l'imagerie des exoplanètes avec optique adaptative extrême : les systèmes HR 8799, HD95086 et 51 Eri hébergent un ou plusieurs compagnons planétaires de masse comprise entre cinq et dix fois la masse de Jupiter. Pour ces images, qui sont le résultat de traitement à posteriori de type imagerie différentielle angulaire, la limite de détection en contraste est de l'ordre de 10^{-6}.

SCExAO

Contrastant avec des projets construits par des consortia, un projet indépendant développé par un groupe de trois semi-permanents et d'étudiants au Subaru Telescope, a pu faire sourire : c'est pourtant cette perspective qui m'a fait, après trois ans passés à Cornell, retourner travailler au Subaru Telescope, pour mettre au point ce qui est aujourd'hui l'instrument SCExAO (Subaru Coronagraphic Extreme Adaptive Optics), en service au foyer Nasmyth infrarouge du Subaru Telescope.

La spécificité du projet SCExAO a été d'offrir un environnement de R&D appliqué, centré sur la mise à disposition d'un instrument modulaire et évolutif, utilisant le retour d'expérience des observations et les progrès continus de la technologie pour améliorer ses performances. Ce mode de fonctionnement est très différent de l'approche qui fait norme aujourd'hui, basée sur la définition d'un cahier des charges figé aux fonctions réparties par un consortium, en réponse à un appel d'offre spécifique. Alors que ce modèle est certainement justifié lorsqu'il s'agit de fabriquer un outil de production scientifique de masse sur un créneau ne présentant que très peu d'inconnues, le retour d'expérience des projets XAO tend à valider le parti pris de SCExAO.

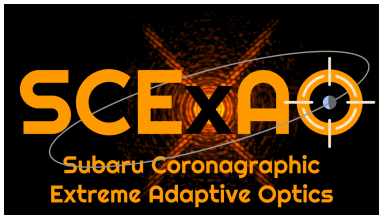


FIGURE 25: Logo de l'instrument SCExAO

En effet, l'incroyable qualité de la correction apportée par les systèmes XAO (voir la figure 24, extraite de [29]) a permis de mettre en

29. B. P. Bowler. Imaging Extrasolar Giant Planets. **PASP**, 128(10):102001, Oct. 2016

évidence plusieurs effets instrumentaux indésirables jusque là invisibles, induits par l'observatoire lui même, et qui imposent à l'instrument haut contraste, un changement de stratégie pour pouvoir s'en affranchir :

— la persistence d'aberrations induites par les flexions mécaniques dans le télescope couplés à la diffraction induite par les multiples optiques, conduit à l'apparition de speckles quasi-statiques dans les images.

— des imprécisions et de l'hystérèse [30] dans l'entraînement mécanique des télescopes, se matérialisant par des élongations de la PSF (vibrations), en particulier au moment du transit et près du zénith.

— l'apparition d'aberrations [31] induites par de forts gradients de température près de la structure porteuse du miroir secondaire des télescopes, particulièrement visibles lorsque le vent tombe, un phénomène appelé "low-wind effect" (LWE).

— les aberrations résiduelles de bas-ordre, auxquelles l'optique adaptative amont n'est pas optimalement sensible mais qui affectent fortement la réjection coronographique, en particulier aux faibles séparations.

— les imperfections dans le contrôle de l'infrastructure comme les compensateurs de réfraction atmosphérique différentielle (ADC) [32], qui se traduit par une PSF allongée.

30. J. Lozi, O. Guyon, N. Jovanovic, G. Singh, S. Goebel, B. Norris, and H. Okita. Characterizing and mitigating vibrations for SCExAO. In *Adaptive Optics Systems V*, volume 9909 of **Proc. SPIE**, page 99090J, July 2016

31. J.-F. Sauvage, T. Fusco, M. Lamb, J. Girard, M. Brinkmann, A. Guesalaga, P. Wizinowich, J. O'Neal, M. N'Diaye, A. Vigan, D. Mouillet, J.-L. Beuzit, M. Kasper, M. Le Louarn, J. Milli, K. Dohlen, B. Neichel, P. Bourget, P. Haguenauer, and D. Mawet. Tackling down the low wind effect on SPHERE instrument. In *Adaptive Optics Systems V*, volume 9909 of **Proc. SPIE**, page 990916, July 2016

32. P. Pathak, O. Guyon, N. Jovanovic, J. Lozi, F. Martinache, Y. Minowa, T. Kudo, H. Takami, Y. Hayano, and N. Narita. A High-precision Technique to Correct for Residual Atmospheric Dispersion in High-contrast Imaging Systems. **PASP**, 128(12):124404, Dec. 2016

A CES EFFETS MAINTENANT CONNUS, on peu s'attendre à devoir se confronter, à chaque amélioration des performances, à des détails encore non complètement pris en considération : différence entre longueur d'onde de l'analyse de surface d'onde et celle de l'image ; retard systématiques de correction résultant en l'apparition de speckles non seulement persistants, mais aussi incohérents, conduisant à des biais de détection. La caractéristique commune de ces effets est qu'ils échappent à la structure de contrôle de surface d'onde, parce qu'ils sont à la marge du régime spécifique de fréquence spatiale et temporelle pour lesquels l'analyseur de surface d'onde a été optimisé. La faible sensitivité intrinsèque à ces modes, combinée au problème de chemin optique non-commun, encourage à utiliser l'image scientifique comme outil de diagnostic.

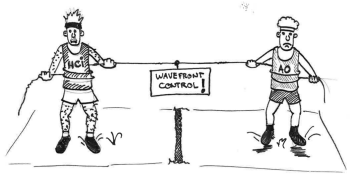


FIGURE 26: Imagerie Haut Contraste (HCI) vs Optique Adaptative (AO) : la bataille pour le contrôle de front d'onde a démarré !

SI ON VEUT ALLER AU DELÀ DE CET EXISTANT, on ne peut plus se permettre de considérer séparement les fonctions de contrôle de front d'onde et d'imagerie haut contraste, qu'il faut au contraire intégrer en une seule fonction spécifique. Le fil directeur de mon travail de R&D sur SCExAO, a été, autant que faire se peut, de développer les techniques pratiques qui utilisent directement l'image produite par le détecteur scientifique pour contrôler le miroir déformable qui contrôle la surface d'onde. Mon ambition sur le long terme est de pousser cette idée à son paroxysme et d'arriver à un système d'op-

tique adaptative sans analyseur de surface d'onde autre que l'image produite par la caméra scientifique.

Le PIAA au coeur de SCExAO

Le projet SCExAO a véritablement démarré à Subaru en 2008, quelques cinq ans après la première publication [33] sur le concept de PIAA (Phase Induced Amplitude Apodization). Fort des succès prometteurs de ce concept étudié d'abord en simulation [34] et puis en laboratoire [35], la construction et l'intégration d'optiques PIAA adaptées à la pupille du Subaru Telescope a commencé un peu avant ma prise de fonction dans le projet.

L'ambition initiale était de bénéficier de l'infrastructure offerte par le contexte : un système d'optique adaptative amont appelé AO188 déjà en service au foyer Nasmyth IR du télescope [36], une caméra infrarouge appelée HiCIAO [37] utilisée dans le cadre d'un grand relevé de détection de planètes extrasolaires et de disques, et d'insérer entre ces deux éléments : un miroir déformable de type MEMS à grand nombre d'actionneurs (1024 dans la version originale, 2500 dans sa version actuelle) pour faire de la modulation fine et à haut ordre du front d'onde en sortie de l'AO188, et un coronographe PIAA, pour pouvoir faire du haut contraste à 10^{-6} avec un IWA de l'ordre de 1.5 λ/D.

CONTRAIREMENT AUX APPLICATIONS SPATIALES pour lesquelles il avait été initialement pensé, une apodisation de type PIAA dans un instrument pour le Subaru Telescope doit pouvoir composer avec une pupille particulièrement inhospitalière au haut contraste. Une particularité du Subaru Telescope est l'exploitation d'instruments très grand champ, au foyer primaire : la pupille au foyer Nasmyth est impactée par la mécanique prévue pour supporter le poids d'un instrument primaire et présente une large obstruction centrale (30 % du diamètre du télescope) ainsi que d'épaisses araignées ($\sim$25 cm). Ces caractéristiques ayant un impact majeur sur la figure de diffraction produite par le télescope, même en l'absence d'aberrations, une solution technique a été mise au point pour atténuer l'impact de ces structures, utilisant le principe de la reconfiguration de pupille. Deux éléments optiques faisaient initialement partie de cette solution, illustrée par la figure 27 :

— un complexe assemblage de prismes appelés SRP (Spider Removal Plate), dont le rôle était de translater vers l'intérieur les quadrants formés par les araignées du télescope et ainsi limiter leur impact.

— un système d'apodisation à deux lentilles fortement asphériques, le PIAA, qui réussit le tour de force d'apodiser la pupille et de faire disparaître l'obstruction centrale du télescope, avec une transmission avoisinant les 90 %.

33. O. Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging. **A&A**, 404:379–387, June 2003

34. F. Martinache, O. Guyon, E. A. Pluzhnik, R. Galicher, and S. T. Ridgway. Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronograph. II. Performance. **ApJ**, 639:1129–1137, Mar. 2006

35. O. Guyon, E. Pluzhnik, F. Martinache, J. Totems, S. Tanaka, T. Matsuo, C. Blain, and R. Belikov. High-Contrast Imaging and Wavefront Control with a PIAA Coronagraph: Laboratory System Validation. **PASP**, 122:71–84, Jan. 2010

36. Y. Minowa, Y. Hayano, S. Oya, M. Watanabe, M. Hattori, O. Guyon, S. Egner, Y. Saito, M. Ito, H. Takami, V. Garrel, S. Colley, T. Golota, and M. Iye. Performance of Subaru adaptive optics system AO188. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2010

37. K. W. Hodapp, R. Suzuki, M. Tamura, L. Abe, H. Suto, R. Kandori, J. Morino, T. Nishimura, H. Takami, O. Guyon, S. Jacobson, V. Stahlberger, H. Yamada, R. Shelton, J. Hashimoto, A. Tavrov, J. Nishikawa, N. Ukita, H. Izumiura, M. Hayashi, T. Nakajima, T. Yamada, and T. Usuda. HiCIAO: the Subaru Telescope's new high-contrast coronographic imager for adaptive optics. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Aug. 2008

La reconfiguration de la pupille par ces optiques transforme complètement la pupille en un faisceau apodisé, produisant une image focale sur l'axe ne présentant plus que de très faibles structures de diffraction résiduelles (voir les PSFs sur l'axe résultantes dans la figure 28). A l'heure d'aujourd'hui, le PIAA reste un élément clé de SCExAO. La SRP ne fait par contre plus partie des optiques de l'instrument. Alors que le PIAA procède à un remapping continu de la distribution d'intensité dans la pupille, la SRP (dont la construction était difficile) opère une transformation brutale, à l'origine de discontinuités de la surface d'onde. Pour compenser cette apodisation moins idéale, le Lyot stop du coronographe a simplement été modifié pour prendre en compte les araignées.

Cette brève description des optiques de remapping me permet d'introduire la publication suivante[38], qui présente en détails les tests de ces optiques et de leurs inverses respectifs, en particulier les propriétés d'imagerie hors-axe.

38. J. Lozi, F. Martinache, and O. Guyon. Phase-Induced Amplitude Apodization on Centrally Obscured Pupils: Design and First Laboratory Demonstration for the Subaru Telescope Pupil. **PASP**, 121:1232–1244, Nov. 2009

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **121**:1232–1244, 2009 November
© 2009. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

Phase-Induced Amplitude Apodization on Centrally Obscured Pupils: Design and First Laboratory Demonstration for the Subaru Telescope Pupil

JULIEN LOZI, FRANTZ MARTINACHE, AND OLIVIER GUYON

National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720; frantz@naoj.org

Received 2009 March 17; accepted 2009 September 11; published 2009 October 21

ABSTRACT. High-contrast coronagraphic imaging is challenging for telescopes with central obstructions and thick spider vanes, such as the Subaru Telescope. We present in this article the first laboratory demonstration of a high-efficiency PIAA-type coronagraph on such a pupil, using coronagraphic optics which will be part of the Subaru Coronagraphic Extreme-AO (SCExAO) system currently under assembly. Lossless pupil apodization is performed by a set of aspheric PIAA lenses specifically designed to also remove the pupil's central obstruction, coupled with a spider removal plate (SRP) which removes spider vanes by translating four parts of the pupil with tilted plane-parallel plates. An "inverse-PIAA" system, located after the coronagraphic focal plane mask, is used to remove off-axis aberrations and deliver a wide field of view. Our results validate the concept adopted for the SCExAO system, and show that the Subaru Telescope pupil can be properly apodized for high-contrast coronagraphic imaging as close as $\approx 1\ \lambda/D$ with no loss of sensitivity. We also verify that off-axis aberrations in the system are in agreement with theory, and that the inverse PIAA system recovers a wide usable field of view for exoplanet detection and disks imaging.

Online material: color figure

1. INTRODUCTION

While the existence of large numbers of extrasolar planets around solar type stars has been unambiguously demonstrated by radial velocity (RV), transit, and microlensing surveys, attempts at their direct imaging with AO-equipped large telescopes remain largely unsuccessful. Due to modest AO performance and lack of high-performance coronagraphs, the planet-rich inner parts of solar systems (including the habitable zones) are still out of the reach of current imaging surveys, which are only sensitive to planets beyond $\sim 0.2''$. Several recent surveys (Lafrenière et al. 2007; Biller et al. 2007; Kasper et al. 2007) indicate that massive planets on such wide orbits are rare, although recent observations have found candidates around the young A stars HR 8799 (Marois et al. 2008), β Pictoris (Lagrange et al. 2008), and Fomalhaut (Kalas et al. 2008).

Many coronagraph concepts have been developed over the last 10 years, most of them derived from Bernard Lyot's original design (Lyot 1939). While most designs theoretically offer very high contrast, the performance is in practice severely compromised by residual wavefront errors after correction by the AO system (Racine et al. 1999; Guyon 2005; Cavarroc et al. 2006) and by the optical layout of telescopes, especially the presence of a central obscuration and spider vanes due to the secondary mirror (Sivaramakrishnan & Lloyd 2005; Martinez et al. 2008).

Yet a new generation of high-contrast imaging instruments is coming online: GPI on Gemini South (Macintosh et al. 2008), SPHERE on VLT (Beuzit et al. 2006), and SCExAO/HiCIAO on Subaru (Hodapp et al. 2008). They all use so-called extreme AO systems which employ deformable mirrors (DM) with large number of fast actuators and new wavefront control techniques to provide the best raw contrast ($\sim 10^{-5}$) accessible from the ground in the near-infrared.

In the first phase of the SCExAO project (spring, 2010), a high-performance PIAA coronagraph will be implemented with a 1024 actuator MEMs-based deformable mirror (DM) as an upgrade that will feed Subaru's newly commissioned coronagraphic imager instrument HiCIAO (Hodapp et al. 2008) in the context of the Subaru Strategic Exploration of Exoplanets and Disks (SEEDS) campaign. One of the key scientific drivers of HiCIAO is to use direct imaging to investigate the presence of Jovian-mass planets around young stars. Current HiCIAO observations in angular differential imaging (ADI) mode using Subaru's AO system (AO188) can only efficiently probe for companions at angular separations greater than $\sim 0.5''$ and reach maximum sensitivity around $1''$. For instance, Lafrenière et al. (2007) respectively report H-band contrast sensitivities of 9.5 mag and 12.9 mag for these separations. SCExAO was designed to complement this search parameter space and allow to probe the innermost parts of extrasolar planetary systems. It is essentially an optical bench that will replace HiCIAO's current fore-optics (standard Lyot coronagraph), for follow-up observations of challenging targets that would benefit from better AO corrections.

At a given angular separation, the achieveable raw contrast level in an extreme-AO system is mostly driven by the speed and accuracy of the wavefront control system. The coronagraphic solution we report in this article and adopted for SCExAO was designed to achieve a 10^{-6} raw contrast at $1\ \lambda/D$ angular separation in the absence of wavefront aberrations at the input of SCExAO. The system is optimized for the H band (1.6 μm).

In the first installment of the SCExAO bench, the SCExAO DM will be used to calibrate the slowly varying aberrations in order to reduce the level of quasi-static and slow-moving speckles (Guyon et al. 2009). Fast correction (1 kHz correction frequency) will be provided by the 188 actuator Subaru curvature system on the front of SCExAO. A subsequent upgrade of the system will include an internal fast visible wavefront sensor that will feed the SCExAO DM, and actually run in extreme AO mode. The raw contrast is then expected to improve to 10^{-4} (detailed performance estimates are indicative rather than predictive, and subject to performance validation in a laboratory prototype currently under design). The $100\times$ to $1000\times$ gain between raw contrast and detection limit will be achieved through time averaging of speckles in the focal plane, augmented by calibration and differential imaging techniques. Consequently, the coronagraph system must be designed so that it does not create fixed speckles much above the 10^{-6} contrast level.

While static phase errors in the coronagraph can be removed by the DM, speckles or diffraction features due to amplitude errors are harder to remove by using phase on the DM: they can require a large stroke and can only removed from one side of the focal plane in a relatively narrow spectral bandwidth. The SCExAO system is therefore designed to minimize such features by amplitude apodization of the pupil (including removing the large central obstruction) with a phase-induced amplitude apodization (PIAA; Guyon 2003; Guyon et al. 2005) and removal of the pupil spider vanes by a custom spider removal plate (SRP). This paper shows that these techniques do not compromise throughput or angular resolution, and do allow high-contrast imaging as close as $1\ \lambda/D$ from the optical axis.

We introduce in § 2 the design of these two key optical components of the SCExAO system and in § 3 present laboratory results that validate their basic functionality. We also show that the static aberrations these components introduce are within the correction range of the SCExAO deformable mirror, and that they will therefore not affect the overall SCExAO performance. Finally, our laboratory demonstration shows how the pupil remapping distorts images of off-axis sources (§ 4). We demonstrate, however, that these aberrations are well compensated by an inverse PIAA.

The tests reported in this article were conducted with no wavefront control system, and we therefore provide no measurement of contrast performance for the SCExAO system at this time. The implementation of a wavefront control system within SCExAO will be the next step of development for this project. High-contrast results with an integrated PIAA coronagraph + wavefront control system are reported in a separate paper (Guyon et al. 2009).

2. LOSSLESS APODIZATION OF A CENTRALLY OBSCURED BEAM

With a central obscuration and spider vanes, the pupil of the large ground-based telescope is incompatible with most high-performance coronagraph concepts, although a few notable exceptions exist (see, for example, Roddier & Roddier 1997; Soummer 2005).

On the Subaru Telescope pupil, the size of the central obstruction (30%, linear) and spider thickness (22 cm) require that the coronagraph is designed to remove the diffraction features they create in the focal plane. The spiders, if left uncorrected, create four spikes at $\sim 10^{-3}$ contrast. These spikes are especially problematic at small angular separation, where they cover most of the position angle space. The central obscuration, at 30%, creates its own set of diffraction rings with a 10^{-3} contrast level at the peak of the central obstruction's first ring (approximately $10\ \lambda/D$). Any coronagraph designed to offer contrast better than 10^{-3} on Subaru Telescope therefore needs to take into account the spiders and central obstruction. We note that several coronagraph concepts can be designed to mitigate the effects of such features. For example, in Lyot-type coronagraphs, the pupil Lyot plane can mask most of the diffraction due to the central obstruction and spiders. This is usually achieved at the expense of coronagraph throughput (for example, a larger fraction of the pupil is masked by the Lyot mask). Martinez et al. (2008) show that this option is suitable for a contrast goal of 10^{-5}, a spider thickness equal to 1.8% of the pupil diameter and a coronagraph with a $2.4\ \lambda/D$ inner working angle. We have not carefully evaluated this option for SCExAO, but we note that it may be difficult to implement given the large spider thickness (2.8% of the telescope diameter) on Subaru, and the SCExAO goal to keep diffraction features due to spiders at or below the 10^{-6} contrast level at $1\ \lambda/D$ from the optical axis.

The small inner working angle of the SCExAO coronagraph makes this Lyot mask approach much less efficient, as the small focal plane mask diameter creates a correspondingly large halo around the spiders ($2.4\times$ larger than for a $2.4\ \lambda/D$ inner working angle coronagraph), so a larger area would need to be masked by the Lyot stop. A variant, proposed by Abe et al. (2006), consists of inserting a complex amplitude filter in the pupil Lyot plane to produce a fully cleared output pupil, which can improve the throughput.

In this article, we demonstrate the feasibility of a different approach: our coronagraph is designed to remove (or at least greatly reduce) both the spiders and central obstruction by geometrically remapping the pupil. Indeed, the classical apodization technique used on most coronagraphs reduces the effective diameter of the telescope, therefore affecting the inner working angle (to $\sim 4\ \lambda/D$) as well as the overall throughput.

SCExAO uses a PIAA-based coronagraph, which preserves both qualities. Compared to classical apodization, this "boost" in resolution allows to probe for the presence of companions at angular separations close to λ/D.

First, a device based on geometrical optics removes the spider vanes by translating the four "pupil quadrants" they define toward the center of the pupil. This device, called the Spider Removal Plate (SRP), as well as its manufacturing are introduced in detail in § 2.1. The beam is then apodized using a Hybrid Phase-Induced Amplitude Apodization scheme (Pluzhnik et al. 2006), compatible with the presence of a central obscuration. The apodizer is presented in § 2.2.

2.1. A Plate to Remove the Spider Vanes

Our approach, illustrated in Figure 1, is to translate each of the four parts of the beam with a single tilted plate of glass, in a manner reminiscent of the technique of pupil densification (Labeyrie 1996). The translations are computed to fill the gap due to the spiders. The spider vanes of the Subaru Telescope are 224 mm thick, for a total pupil diameter of 7.92 m. To account for alignment errors between the SRP and the pupil, we have chosen to design the SRP for slightly thicker (250 mm) spider vanes. The coronagraph we are assembling for Subaru is designed for an input beam of 17.96 mm in diameter. At this scale, the gap due to the spider vanes is therefore 0.453 mm. The SRP consists of four tilted plane-parallel plates, each translating a part of the pupil inwards, as shown in Figure 1. It can be best described as a "pyramid-shaped rooftop" of constant thickness. All four plates were cut from the same plane-parallel plate (a.k.a. optical window), to guarantee, within tolerances, a constant thickness.

Commercially available optical windows have, for typical specifications, a parallelism better than 5″, about 2.5×10^{-5} radians. Across a 1 cm long piece cut from such a window, the maximum wedge is going to be 0.25 μm ($\sim\lambda/7$ in H band) which is well within the comfort zone of a deformable mirror. All four plates of or custom component were cut from a single window of thickness $e = 15$ mm and parallelism better than 3″ and the geometry of the cuts and assembly was optimized to avoid pasting together areas which originated from distant parts of the original window.

To translate a quadrant, each of the four segments of the SRP is tilted by an angle α that is a function of the amplitude of the desired translation. δx and δy respectively are the horizontal and vertical translation of the pupil quadrants (cf. Fig. 1), and β is the angle between the horizontal axis x and the top right vane of the Subaru Telescope pupil. The goal of the SRP being to completely fill the spider gap $\gamma = 0.453$ mm, δx and δy must verify

$$\delta x \sin\beta + \delta y \cos\beta = \gamma. \quad (1)$$

Tilting a SRP segment by an angle α not only displaces the corresponding quadrant by δ, it also introduces an optical path

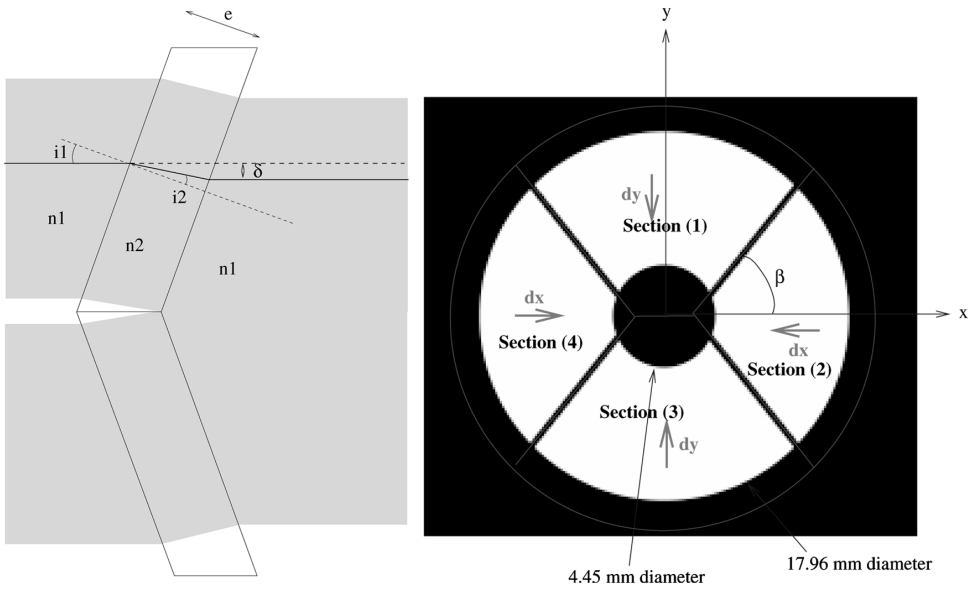


FIG. 1.—*Left panel:* Cross-section of the spider removal plate. The translation of the beam δ is a function of the tilt angle i_1 as well as the index and the thickness e of the window. *Right panel:* Geometry of the Subaru Telescope pupil. The SRP translates quadrants 1 and 3 inward by the amount δy, and quadrants 2 and 4 by δx. See the electronic edition of the *PASP* for a color version of this figure.

difference (OPD) Δ relative to a hypothetical undeflected beam. The relation between the tilt angle α (identical to i_1 in Fig. 1), the deplacement δ, and the OPD Δ come from equations of geometric optics (cf. left panel, Fig. 1)

$$\sin i_1 = n \sin i_2 \quad \text{and}$$
$$\delta = e \sin(i_1 - i_2)/\cos i_2,$$
$$\Delta = \frac{e}{\cos i_2}[n - \cos(i_1 - i_2)].$$

where n is the refractive index of the glass (index of air considered to be exactly one). In the small angle approximation, these relations reduce to:

$$\delta \approx e\alpha(n-1)/n \tag{2}$$

$$\Delta \approx e(n-1)(1+\alpha^2/2n). \tag{3}$$

For the SRP to be useful, one must ensure that the OPD is the same for all four quadrants at all wavelengths, so that the wavefront of an on-axis source remains continuous at all wavelengths after the remapping. This can only be achieved when the four plates have simultaneously the same tilt angle α. The solution to equation (1) is therefore to have $\delta x = \delta y = \delta$.

For a window of thickness $e = 15$ mm and index $n = 1.443$ (fused silica for $\lambda = 1.6$ μm), each plate needs to be tilted by an angle $\alpha = 5.004°$ (cf. eq. (2)). To guarantee the continuity of the wavefront on-axis after remapping within $\lambda/10$, one can derive equation (3) relative to α and find that the tolerance for the OPD translates into a constraint on the tilt angle of each plate

$$\sigma_\alpha = \frac{n\sigma_\Delta}{(n-1)e\alpha}, \tag{4}$$

which gives a tolerance of 0.02° for the tilt angle. Figure 2 shows a picture of the SRP that was assembled for the SCExAO system.

The refractive index of fused silica being a monotonous decreasing function of the wavelength, a simple way to ensure that the SRP completely fills the spider vanes over the entire H band (1.485–1.785 μm) is to optimize it for the "red side" of the band, which is less deflected than the "blue side." The two main nonideal effects of the SRP are:

1. The amplitude of the lateral shifts is chromatic, and the pupil size at the SRP output is therefore wavelength dependent. This effect is however small (0.3% over the entire H band) and can be accounted for in the coronagraph design.

2. The output wavefront for off-axis sources contains "steps" which produce aberrations. These aberrations limit the useful field of view and will be quantified in § 4.1. One solution to this problem is to include a "reverse SRP" behind the coronagraph but in front of the science camera to recover a wide field of view.

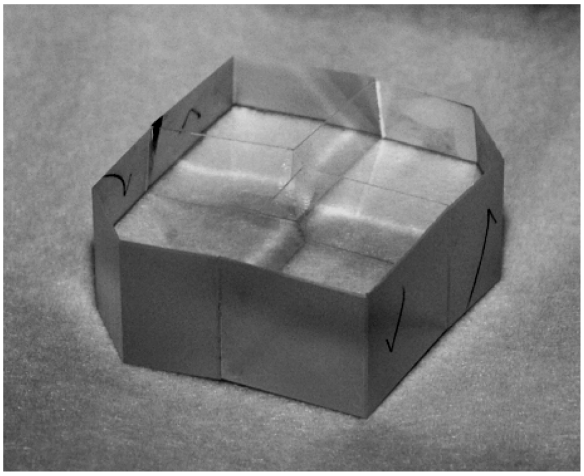


FIG. 2.—Assembled spider removal plate. The plate is 15 mm thick. Note that since the spider vane angle β is 51.75° and not 45°, the SRP is not continuous at the interfaces between the four plates and there are steps running along them.

After the SRP, the beam has no spiders, but still has a central obstruction.

2.2. PIAA with a Central Obscuration

Although not part of the original design of Lyot (1939), apodization of the pupil has been demonstrated to provide a major enhancement of most coronagraphs (Aime & Soummer 2003). The best apodizing function depends on the pupil geometry as well as the type of coronagraph, but prolate spheroidal functions in general have been shown to be natural apodizers for Lyot-type coronagraphs (Soummer et al. 2003), and solutions can be found for centrally obscurated pupils (Soummer 2005).

The simplest way of performing an apodization is to insert a mask whose radial transmission profile follows such a prolate spheroidal function, the so-called classical pupil apodization (CPA). While extremely robust, and insensitive to moderate tip-tilt residuals, this approach has two main drawbacks: the throughput is low (as low as $\sim$0.1 for a 10^{-10} contrast; Kasdin et al. 2003; Guyon et al. 2006), and the effective pupil diameter is reduced by a factor approximately equal to the square root of the throughput (due to the fact that apodizers remove the light mostly at the edges of the pupil), which translates into a loss of angular resolution.

The Phase-Induced Amplitude Apodization (PIAA; Guyon 2003) addresses these issues and apodizes the beam using a very different approach. It uses a set of two tailored optics working in pair: inserted in the pupil plane, the first changes the distribution of light, while the second collimates the beam for an on-axis source. While this idea has already been studied in great details (Guyon 2003; Guyon et al. 2005; Martinache et al. 2006), this

article presents a design adapted to the presence of a central obscuration.

Pluzhnik et al. (2006) have already shown that a PIAA/CPA hybrid system is an attractive design that eases the manufacturability of the first of the two PIAA optics near the outer edge of the pupil. This is the design we have adopted for the SCExAO system. These PIAA optics are designed as a refractive system (lenses instead of mirrors). While such a choice would clearly be a risky option for a space-based coronagraph aimed at reaching very high contrast ($\sim 10^{-9}$), refractive optics can safely be used in coronagraphs for ground-based telescopes.

The set of PIAA lenses is an afocal system that leaves the pupil diameter unchanged. A low-dispersion material was chosen (CaF_2) to keep chromatic aberrations small. Such a refractive PIAA system has several advantages: its circular symmetry makes it easier to manufacture and it is quite compact (96 ± 0.1 mm long for a $\sim$20 mm beam). Radial sag profiles of the PIAA lenses are presented in Figure 3. The profile for L_1 exhibits four main features:

1. A central flat circular area corresponding to what is left of the central obscuration after the SRP (cf. § 2.1).
2. A convex (convergent) ring that densifies the inner part of the pupil and concentrates the distribution of light toward the center of the pupil.
3. A concave (divergent) ring that dilutes the outer part of the pupil. The telescope beam only illuminates L_1 out to 8 mm (vertical dashed line on Fig. 3), which corresponds to the bottom of the well, where the curvature radius is smallest.
4. An external flat ring that provides a convenient way of mounting the lens.

The exact shape of the lens around the outer edge of the beam is critical to the beam shaping, and greatly conditions the efficiency of the downstream coronagraph. Yet because the curvature radius of the lens is the smallest in that region, it is also where the biggest manufacturing difficulty arises.

In order to relax the manufacturing tolerances, the PIAA was designed to provide a partial apodization only, to later be completed by a binary mask (cf. § 2.3). Also, while apparently superfluous, the righthand side of the well in the profile of L_1 acts as a safeguard, making the system independent of slight changes of scale in the pupil that typically accompany realignment of upstream optics. Thus, an oversized pupil that would otherwise induce light leaks has its outermost part deflected away from the beam, preserving the performance at an acceptable (a few percent) cost in throughput.

The sag profile for L_2 is complementary to L_1 and exhibits three main features:

1. A concave (divergent) central part that collimates the light that was densified by part [2] of L_1.
2. A convex (convergent) ring that collimates the light that was remapped by part [3] of L_1.
3. A flat external ring to facilitate mounting.

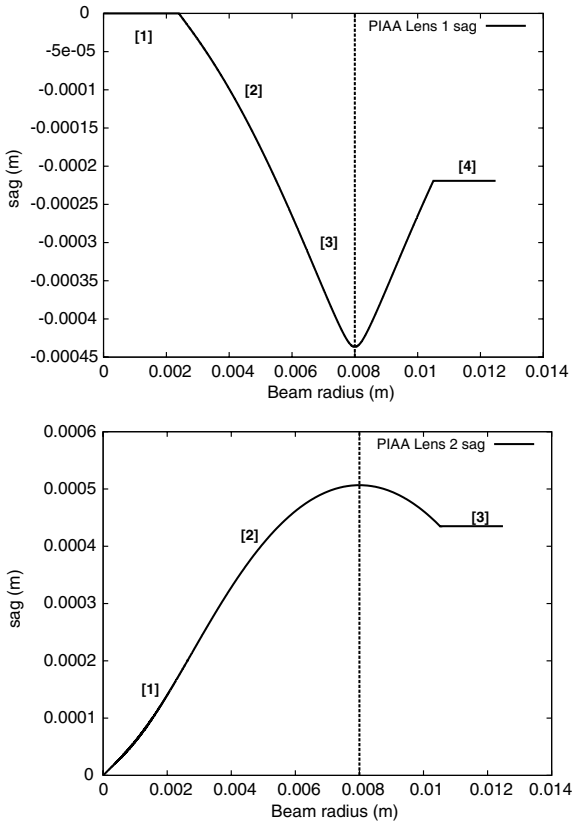


FIG. 3.—Radial sag profile for the two PIAA lenses. On both panels, the *vertical dashed line* localizes the outer edge of the pupil (radius 8 mm). Labels on the two curves refer to main features of the lenses discussed in the text.

Figure 4 shows a picture of the manufactured PIAA lenses before they are tested in the lab. The lenses are designed to be separated by 96 mm exactly at $\lambda = 1.6$ μm, and adapted for a 16 mm-diameter beam.

2.3. Binary Mask to Complete the Apodization

As explained in § 2.2, the PIAA leaves a small amount of light (0.82% of the central surface brightness) in the outer part of the apodized pupil. To be effective, the apodization therefore needs to be completed by a CPA mask (Pluzhnik et al. 2006) in order to turn the (still) hard outer edge of the pupil into a smooth transition. To perform this complementary apodization, we chose to use a binary mask (Kasdin et al. 2005) with a series of narrow opaque rings blocking light. With this approach, the apodizer is achromatic and its transmission variations do not introduce phase variations on the pupil.

The position and width of the rings is optimized to best approximate the ideal continuous apodization profile shown in Figure 5 as the "Apodizer transmission" curve. Several

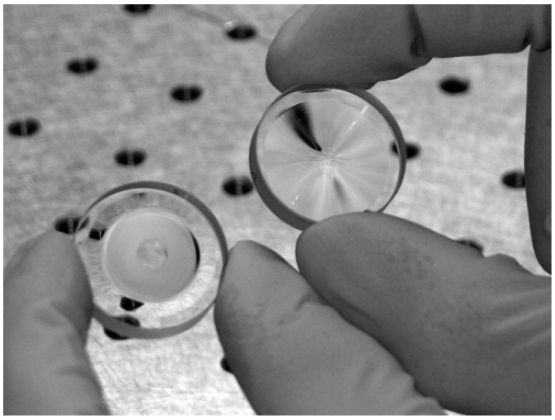


FIG. 4.—PIAA lenses used to apodize the pupil of the Subaru telescope. L_1 (*left*) apodizes the beam and removes the central obstruction. L_2 (*right*) collimates the apodized beam.

constraints were imposed on the design to ensure manufacturability: no ring should be less then 1.6 μm wide and the gap between consecutive opaque rings should be no less than 15 μm. The resulting design is composed of 155 opaque rings for a total apodizer diameter of 11.09 mm (defined by the outer edge of the last opening between opaque rings). The apodizer was manufactured by lithography on a transmissive substrate. Figure 6 shows a microscope image of the mask.

The minimum width of the rings was limited to 1.6 μm to ease manufacturing and avoid subwavelength features which could create strong chromatic and polarization effects. We have not done a diffraction analysis of the mask to quantify such effects, but we note that (1) a similar design with 0.8 μm wide rings working in the visible did not show significant deviation from Fraunhofer diffraction theory (Guyon et al. 2009), and (2) the rings are mostly affecting the outer part of the pupil where there is little incoming light.

The postapodizer throughput over the 11.09 mm diameter is 95.4%, but due to the narrow rings in the apodizer, some of the light transmitted is diffracted at large angles. The "effective" throughput of the apodizer is therefore lower, at 92.3%, and would be equal to the "raw" throughput if the apodizer were continuous instead of binary. An additional 4.8% transmission is lost because the beam is clipped by the apodizer at 69.3% of the nominal 16 mm beam diameter. This clipping is equivalent to clipping the outer 2.5% of the input telescope beam at the entrance of the PIAA system, and relaxes pupil alignment tolerances at the expense of a 4.8% throughput loss and a 2.5% angular resolution loss. The combined throughput from (1) absorption in the apodizer, (2) diffractive losses in the apodizer, and (3) beam clipping is 87.5%. We note that most of these losses could be reduced with a more aggressive design with

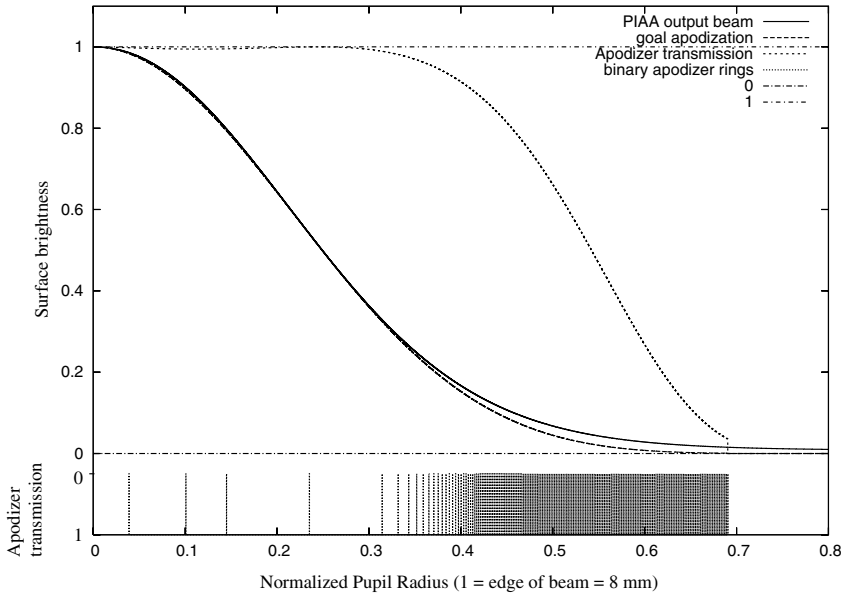


FIG. 5.—Binary mask transmission profile and discretization. The goal apodization (dashed line) is the result of the multiplication of the PIAA output beam profile (*solid line*) by the apodizer transmission profile (*short dashed line*). The rings in the binary apodizer (bottom) become very tightly spaced in the outer part of the beam, effectively blocking a larger fraction of the incoming light.

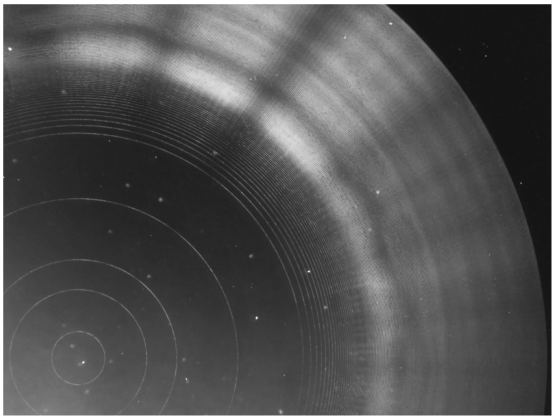


FIG. 6.—Image of the binary mask taken with a microscope. The transmission is about 10% on the edge of the pupil.

tighter tolerances, but we decided for this first on-sky PIAA system to keep the design more conservative.

3. LABORATORY TESTING

3.1. Optical Layout

Because of its specific design (incoming and outcoming beams are collimated, and the aspheric surfaces face each other), the PIAA can work over a broad range of wavelengths. The distance between the two PIAA lenses can compensate for the change of refraction index when changing wavelength. This allowed us to test the optics in visible light, using a standard laser diode ($\lambda = 633$ nm) as our light source, although the system is optimized for the H band ($\lambda = 1.6$ μm).

We have chosen to test the remapping optics (PIAA optics and SRP) with visible light for convenience. The optics were designed for near-IR wavelength, but their functionality can be tested with visible light, although near-IR optimized antireflection coatings create ghosts. We note that a high-contrast coronagraphic imaging test would require near-IR light in addition to wavefront correction with a deformable mirror, but in the absence of wavefront control, testing at a wavelength almost 3 times shorter than nominal provides much higher sensitivity to aberrations as well as alignment errors.

Except for the wavefront control system that is not tested in this laboratory demonstration, the test bed is close to the system to be installed onto the Nasmyth platform of the Subaru Telescope, after the 188 actuator AO system and before HiCIAO (Hodapp et al. 2008). Figure 7 provides an overview of the system. In the diagram (top), a 17 mm collimated beam comes horizontally from the lower-right corner, and goes through a mask simulating the Subaru pupil. The beam then goes through the SRP (see § 2.1), the PIAA (see § 2.2), and the binary mask (cf. § 2.3). At this point, the beam is suitably shaped for high-performance coronagraphy.

The light reflects off a flat mirror and travels from bottom to top through a first achromatic doublet that provides the first focal plane where a coronagraphic mask is inserted, and another achromatic doublet that provides another pseudopupil plane downstream.

The light then reflects off the top-left flat mirror and goes toward the left, through an exact replica of the first PIAA system, only mounted backward. The motivation for this will be discussed in § 4.2. A two-doublet system makes an image on a CCD camera that can be placed either in the pupil or the image plane.

Each element (cf. lower image in Fig. 7) is mounted on a motorized translation stage to be removed from the optical train in a few seconds. This provides good flexibility, and permits the testing of multiple combinations for diagnostics, to the extent that even if all "active" elements (SRP and PIAA) are taken off the beam, the system still works as a perfectly functional nonapodized Lyot coronagraph.

3.2. Image Acquisition and Processing

All acquisitions are made by a 1600×1200 pixel CCD camera (7.4 μm square pixels). A optical density of 2 is used to avoid saturation. Each image is cleaned by removing bias and hot pixels, and for each simulation, we align and coadd 20 acquisitions to reduce readout and photon noises.

The final two-doublet system mentioned in § 3.1 was chosen to provide sampling better than Nyquist in the image plane, while still allowing us to image the pupil with a sufficient number of pixels to observe the effect of the PIAA and the SRP. With these "active" elements out of the beam, the retained configuration provides 2.3 pixels per λ/D and 390 pixels in diameter for the pupil of the telescope.

3.3. Effect of the SRP

Figure 8 demonstrates the effect of the SRP, whose design and specifications are discussed in § 2.1. The figure compares images of the pupil (left column) and the corresponding point-spread functions (PSFs, right column) taken with and without the SRP in the beam.

The effect of the suppression of the spider vanes is clearly visible as the diffraction spikes present in the standard imaging mode PSF (top right image of Fig. 8) are diminished when the SRP is introduced into the beam (bottom right image of same figure). Thinner spiders are still present, however, and contribute to weak diffraction spikes in the images. This incomplete suppression of the spiders can be attributed to the finite thickness of the glue holding the quadrants together in the SRP.

Since no wavefront control system was used during these tests, nothing compensates for the aberrations introduced by the SRP, which being imperfect, necessarily deteriorates the quality

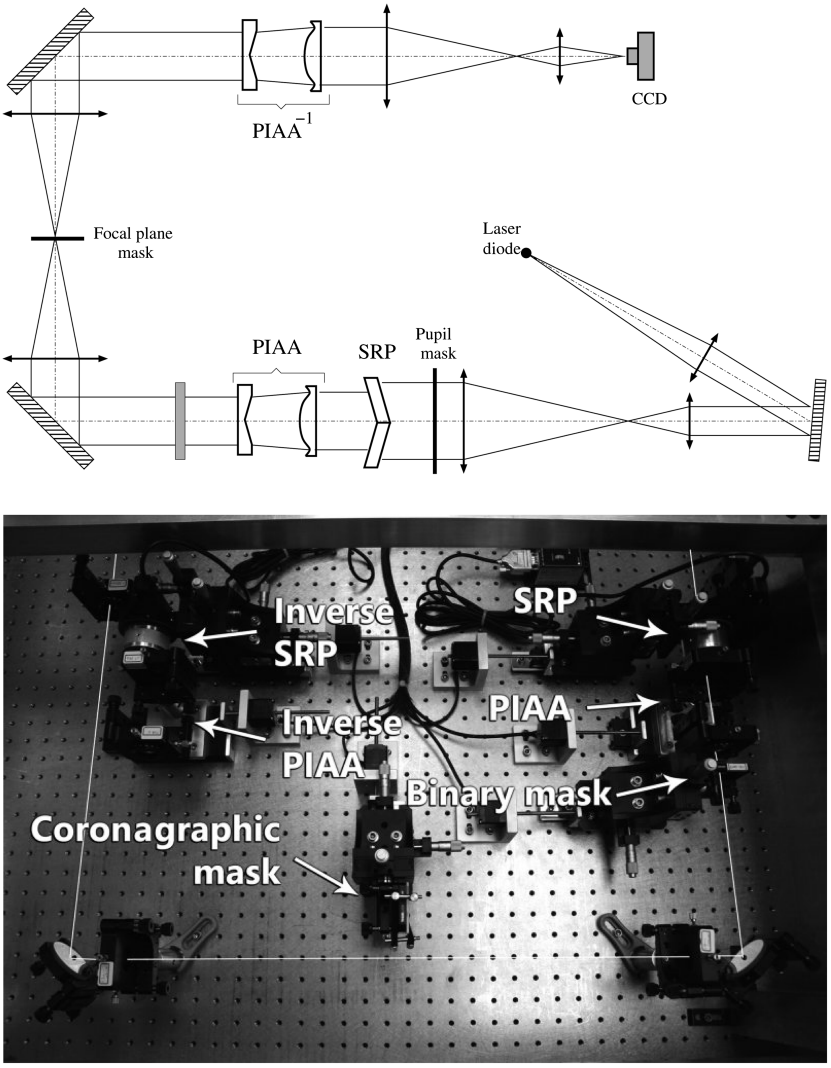


FIG. 7.—*Top:* Optical layout used for the tests reported in this article. *Bottom:* Photograph of the actual experiment: the light is comes from the top right and goes into the SRP, the PIAA lenses, and the binary mask. A lens focalizes the beam on the coronagraphic mask, and the beam then goes into the inverse PIAA, then the inverse SRP. The beam then goes to the CCD camera.

of the PSF. In order to offer a basis for comparison, the central column of Figure 8 presents simulated versions of the same images. Each PSF simulation was calculated using the square root of its corresponding pupil image for amplitude and constant phases. Without the SRP, a comparison of simulated and experimental images is convincing evidence that the relaying lenses and mirrors are sufficiently well aligned not to introduce any obvious aberrations.

When inserting the SRP, the deterioration of the image quality is nonnegligible. Repeated measurements of a Strehl ratio $S = 0.48$ with the SRP ($S = 0.98$ without) offer an estimate of the rms aberration it introduces through

$$S = \exp -(2\pi\sigma/\lambda)^2, \tag{5}$$

which for $\lambda = 633$ nm gives a rms wavefront aberration $\sigma = 86$ nm . For reference, for $\lambda = 1.6$ μm, the same component would provide a Strehl ratio $S = 0.89$.

This wavefront error, since it is static, does not have an impact on the final PSF quality as long as (1) it is measured by

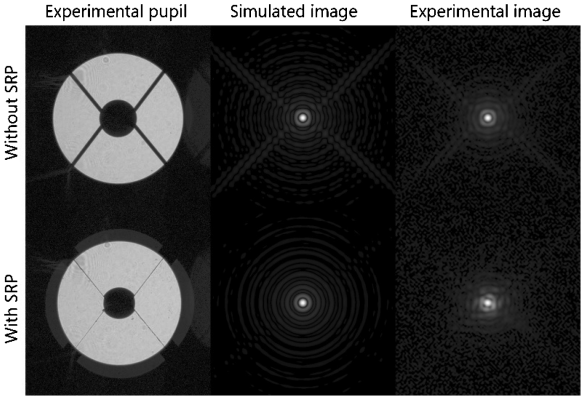


FIG. 8.—Effect of the SRP on the pupil. *Top* (*left to right*): Image taken with the camera in a location conjugated with the pupil plane, showing the mask simulating the Subaru telescope pupil, corresponding simulated image, experimental image. *Bottom:* The same images with the SRP inserted into the beam (cf. § 3.3 for details).

the wavefront sensor, and (2) it is within the correction range of the deformable mirror. The SCExAO architecture will include a wavefront sensor at the science focal plane, and the SCExAO's DM stroke is about 1 μm. Since the remapping introduced by the SRP will turn a continuous wavefront error into a sharp step, a DM located after the SRP cannot easily correct such wavefront errors. The SCExAO's DM will therefore be located before the SRP.

Because it is index dependent (see eq. [4]), the displacement δ is a function of the wavelength and will go from $\delta = 0.402$ mm for $\lambda = 1.6$ μm to $\delta = 0.410$ mm for $\lambda = 633$ nm.

We note that being optimized for the near-infrared, the coating of the SRP is partially reflective (transmission $T = 0.89$) at the test wavelength. The pupil image with the SRP (Fig 8, lower left) shows that each quadrant of the SRP creates a ghost toward the outside of the pupil. These ghosts will be considerably fainter in the near-IR wavelength for which the coatings have been designed."

For now, the tip-tilt of the SRP is controlled by two manual micrometer drives, with an accuracy of a few arcseconds. In the final system, remotely controlled actuators with higher accuracy will be used.

3.4. Pupil Apodization

The role of the apodization is to turn the hard-edge pupil, responsible for the presence of rings in the PSF, into the ideal prolate spheroidal that will maximize the coronagraphic null. In this system, the apodization is a three-step process that combines the SRP, the PIAA, and the binary mask.

Figure 9 plots pupil radial profiles measured with our experiment using different combinations of these three elements, to illustrate how each affects the pupil.

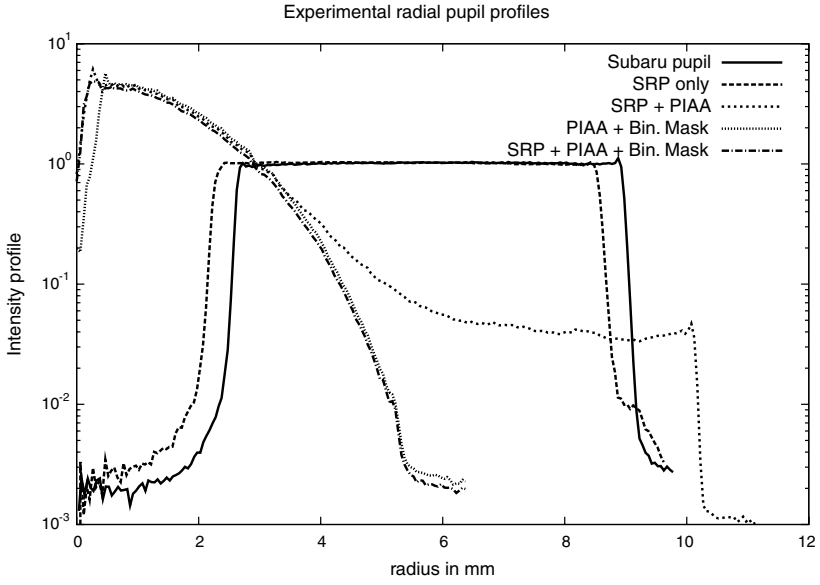


FIG. 9.—Radial intensity profiles of the pupil in the presence of different combinations of the active remapping elements: SRP, PIAA, and binary mask. The mean illumination of the pupil in the absence of remapping (*solid curve*) is taken as reference. Note that unlike Fig. 5, the y-axis uses a logarithmic scale.

The effect of the SRP on the pupil profile is minimal. To fill the gap due to the spider vanes, it moves the four quadrants inward by 0.4 mm. As the pupil radius is ~9 mm, this translates into the 4.5% pupil size reduction observed between the solid and dashed line profiles in Figure 9. Note that the central obscuration shrinks by the same amount. Section 4.2 will show how this apparent reduction of the pupil diameter does not also translate into a penalty in angular resolution.

The PIAA has a much more dramatic effect on the pupil. It takes the light that was uniformly distributed in the original pupil and concentrates it toward the center, filling up the void left by the central obscuration. The PIAA therefore turns the hard-edge (solid-line) profile into a bright lobe (curve labeled SRP + PIAA in Fig. 9) surrounded by a weak (~30 times fainter than originally), slightly oversized uniform background. The binary mask completes the apodization. It essentially substracts the background (additional attenuation by a factor ~10) and puts the finishing touches to the central lobe, to maximize the rejection by the occulting mask.

While looking at Figure 9, the reader may be under the impression that the binary mask sabotages the hard work put in by the PIAA and the SRP, by reducing the effective diameter of the pupil by a factor of 2. The reader should be reassured by two things: (1) the vertical scale on the plots of Figure 9 is logarithmic and this apparent reduction of the pupil diameter is much less significant than it appears, and (2) the binary mask is located downstream of the PIAA, which leaves very little light on the edge of the pupil, where absorption by the binary mask is greatest. Overall, the transmission of the system is 92.3%, excluding losses due to coatings, and beam clipping introduced to relax pupil-alignment tolerances. To achieve the same apodization profile on a nonobstructed pupil, a CPA would have a 17.5% throughput.

Figure 10 shows images of the illumination of the apodized pupil taken with the SRP in and out of the beam, as well as corresponding focal plane images. The presence of the spider vanes in the top left quadrant offers a way of visualizing the remapping of the pupil performed by the PIAA: not only do the spider vanes get thinner toward the center of the pupil as the distribution of complex amplitude gets concentrated, but because they do not exactly converge toward the geometric center, the spider vanes are distorted by the remapping. The images in the middle column show the corresponding simulated PSFs, calculated from the experimental distribution of amplitude in the pupil with a perfectly corrected wavefront (i.e., phase = 0). These images demonstrate the effect of the SRP as it reduces the diffraction spikes of the spiders by a factor of 10^3. The images in the right column are their experimentally obtained counterparts. Although essential features of the PSFs are reproduced, in the absence of wavefront control, the performance is obviously not as good as the simulations predict. The elliptic (astigmatic) experimental PSF obtained with the SRP can be attributed to an error of alignment at the time of data acquisition. Nevertheless, the fact that in both cases these PSFs contain an pseudo-Airy disk shows that the rms of aberrations is below $\lambda/4$, that is, 150 nm. The DM the SCExAO project will use offers a peak-to-valley wavefront correction amplitude of 3 μm. The correction of this static 150 nm aberration will therefore not represent a large fraction of the mirror stroke, which would otherwise limit the SCExAO performance.

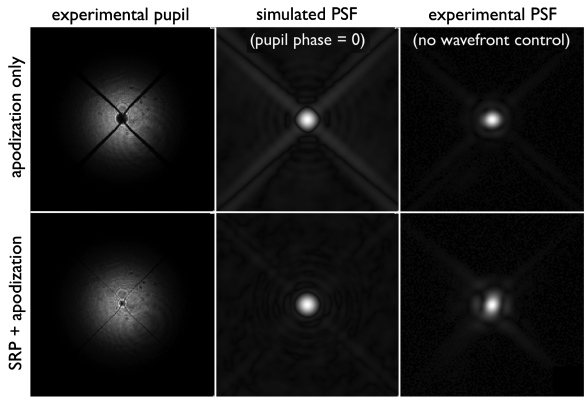


FIG. 10.—Apodization of the pupil: the theory and the experiment without and with SRP. Results in the image plane.

4. OFF-AXIS IMAGING PERFORMANCES

It is a basic rule of conventional optics never to alter the pupil if the basic object-image convolution relation is to be preserved. With components such as the SRP and the PIAA, the system presented in this article obviously violates this rule, and this section explores the consequences of their respective actions on the field of view.

By design, SCExAO aims at complementing HiCIAO's parameter space toward small angular separations. A strong constraint on the extent of the extreme AO field of view on the actual system will come from the DM used for the wavefront correction. Indeed, the finite number N of actuators projected across the pupil of the telescope sets the outer working angle (OWA) of such a system to $(N/2) \times (\lambda/D)$. The SCExAO DM offers, at most, $N = 32$ such actuators. On the 8 meter Subaru Telescope observing in H band, this OWA is therefore limited to $0.6''$ $(16\lambda/D)$.

4.1. SRP Field of View: Simulations and Results

Being fairly achromatic and inducing no light loss, the remapping of the pupil performed by the SRP can be undone, after the coronagraphic mask, by an inverse system, and two SRPs were manufactured in that scope (cf. Fig. 7). After testing, however, one of the two turned out to be of lesser quality, which suggested that we should investigate how much the best

SRP would alter the "wide-field" imaging capabilities of the SCExAO system.

Indeed, as explained in § 2.1, each quadrant of the SRP adds an OPD that is a function of the tilt angle α for an on-axis source (cf. eqs. [3] and [5]). The design, however, guarantees, within tolerances, that the OPD is the same for all quadrants, so that the wavefront of an on-axis source is continuous after the SRP.

Although the tilt angle is the same, the azimuth of the tilt is different for each quadrant. As a consequence, for an off-axis source, the incidence angle will be different, resulting in a different OPD for each quadrant. At a given wavelength λ, this will translate into a phase offset $\Delta\phi$ that can be calculated using equation (5), and for α, substituting $\alpha + \theta$, with $\theta \ll \alpha$ (a legitimate approximation, since $\alpha = 5°$)

$$\Delta\phi = \frac{2\pi e(n-1)}{n\lambda}\alpha \cdot \theta. \tag{6}$$

The dependence on the tilt azimuth is expressed in the use of the dot product $\alpha \cdot \theta$. We have used this equation to predict how the off-axis affects the quality of the PSF. Figure 11 presents the result of these predictions and compares them to laboratory images taken at different off-axis angles with the SRP only. The agreement between the simulations and the data over a fairly large field of view ($20\lambda/D$) is excellent and incidentally confirms the validity of the proposed approximations.

One sees that for angular separations up to $10\lambda/D$, the SRP produces an image with a pseudo-Airy disk. To better understand the off-axis PSF degradation by the SRP, Figure 11 also plots the energy encircled within the pseudo-Airy disk for the relevant cases. Theoretical values of this quantity for a nonobstructed and a 30% obstructed telescope like Subaru are respectively 85% and 69%. Because the SRP incidentally also reduces the size of the obscuration, simulations show that on-axis, the encircled energy is 72%. It degrades by an acceptable 7% for sources at a $5\lambda/D$, and by an unfortunate 30% at $10\lambda/D$. Experimental images and measurements conform to these simulations, although the on-axis performance is reduced to 41%. This lesser performance can be attributed to the absence of wavefront control in this experiment.

Over a sufficiently large ($10\lambda/D$ in diameter) field of view, the image quality appears very satisfactory, despite the absence of wavefront control, within which high-contrast imaging is performed: an inverse SRP is not necessary in this observing mode. For noncoronagraphic "wide field" observations, however, the SRP should be removed from the beam.

4.2. Off-Axis Imaging with a PIAA

For an off-axis source, the pupil-remapping the PIAA performs introduces large aberrations to the wavefront, which translate into very unsually shaped PSFs. The matter is not new and has been extensively described by Guyon et al. (2005) in the case of a nonobstructed aperture. The novelty in the PIAA that will equip the SCExAO system is that it must accomodate the presence of a central obscuration, and as presented in § 3.4, the PIAA lenses used in combination with the SRP manages to suppress it completely. However, this requires a somewhat brutal remapping, with dramatic consequences on the PSF which finds itself pineapple-shaped at separations greater than $10\lambda/D$ (cf. Fig. 12).

Fortunately, just like with any pupil remapping (Guyon & Roddier 2002; Guyon 2003), the aberrations the PIAA introduces can entirely be corrected using an exact copy of the PIAA, only positioned backward, which restores the original pupil and provides wide field of view imaging capabilities. In practice, this so-called inverse PIAA is located in the pseudo-pupil plane located after the occulting mask (cf. Fig. 7). While this idea of using an inverse system to recover wide field of view capabilities after a PIAA coronagraph has often been

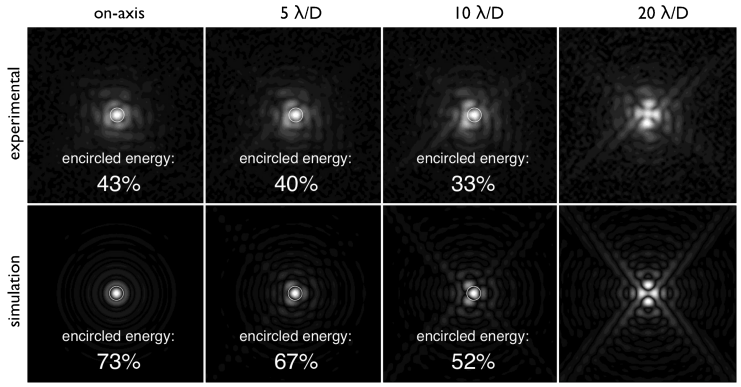


FIG. 11.—Comparison between simulated and experimental images of an off-axis source, from 0 to $20\lambda/D$ with a $5\lambda/D$ step. The field of view of each image is $\pm 12\lambda/D$.

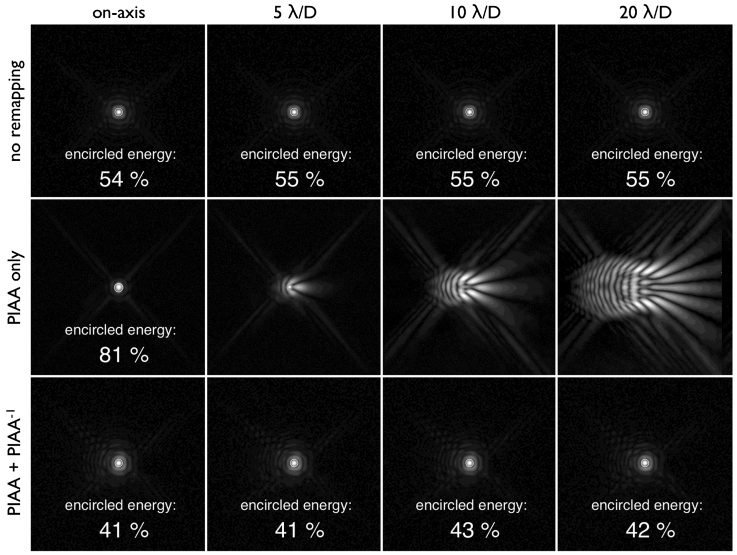


FIG. 12.—Experimental images of sources observed on and off-axis in the absence of SRP. The results of three configurations are shown. *Top row:* no apodization; *middle row:* PIAA + binary mask; *bottom row:* PIAA + binary mask + inverse PIAA. The huge off-axis aberrations the PIAA introduces are satisfactorily compensated by the inverse PIAA as the whole combination system provides diffraction-limited images.

mentioned in PIAA-related publications (Guyon 2003; Guyon et al. 2005; Martinache et al. 2006; Pluzhnik et al. 2006), this article reports the first successful implementation of such an inverse system.

Figure 12 shows a series of off-axis images taken with and without the inverse PIAA and compares them to their equivalent without any remapping at all (except for the SRP). These images show that the inverse PIAA (bottom row) efficiently corrects the huge aberrations the PIAA introduces in the first place (middle row), and turns the pineapple-shaped off-axis images into more conventional Airy-like images, virtually identical to the ones the system produces with no beam apodization at all (top row).

Just as was done for the SRP (cf. § 4.1), Figure 12 also shows the energy encircled within the pseudo-Airy disk of the PSFs (except for the pineapples). The theoretical value of this quantity for the Subaru Telescope pupil is 69%. In the absence of remapping (first row), the experimentally measured encircled energy is 55%. Despite the absence of wavefront control, the apodization by the PIAA manages to bring it a little over 80% on-axis. With the inverse PIAA, the encircled energy stabilizes a little over 40%, whatever the angular position of the source. This 20% loss (relative to the initial 55%) can be attributed to imperfection of the optics. Nevertheless, this result, achieved in the absence of wavefront control is remarkable: the obtention of an actual Airy disk after the inverse PIAA shows that the wavefront residuals rms is better than $\lambda/4$. Observations in the near-infrared and the deformable mirror on the actual SCExAO system should significatively improve this performance.

5. CONCLUSION

The Subaru Coronagraphic Extreme AO system is an instrument to be installed onto the Subaru Telescope, as an upgrade to the recently commissioned infrared camera HiCIAO (Hodapp et al. 2008). The SCExAO system offers the means to probe the innermost part of planetary systems by using technologies optimized for high contrast at small angular separation.

Because it is constrained by the fundamental limits of diffraction and the practical limits of actual telescopes, each high angular resolution imaging system is the result of a multivariable optimization process: contrast versus angular separation versus field of view. While nothing fundamental prevents a PIAA-based coronagraph from achieving high contrast simultaneously with a small inner working angle and over a very wide field of view, practical limits, like a large central obscuration and thick spider vanes, impose some level of specialization for high performance. The SRP + PIAA-based architecture of the SCExAO project was chosen to open a unique niche of high contrast with a small inner working angle for an 8 m telescope. While focusing on a $10\lambda/D$ diameter field of view may seem restrictive, the recent results of other high angular resolution techniques working on a very small field of view, like aperture masking interferometry (Martinache et al. 2009), which SCExAO will also implement, show that such specialization pays off.

In that scope, this article introduces two critical components of the system: a new design for the PIAA apodizer that accomodates the presence of a central obscuration, and a

compound phase plate that suppresses the spider vanes present in most modern telescope designs. The intensive tests conducted on these advanced optical components, at a wavelength 3 times shorter than nominal, demonstrate satisfactory performance despite the absence of wavefront control. We are now integrating these optics with the wavefront control system necessary to produce high-contrast images.

The SCExAO system, in its first installment, is designed to provide high-contrast imaging in the 1 to $4\lambda/D$ separation range in imaging mode only (no spectroscopy). By using achromatic near-IR lenses for relay optics and a low-dispersion material (CaF_2) for the PIAA lenses, chromatic effects in SCExAO are mitigated. Polychromatic operation of SCExAO (for spectroscopy) has, however, not been explored at this point, and will likely require redesign of several components in the optical train (most importantly to compensate for the linear dependence of the diffraction limit λ/D with wavelength in the focal plane). The effect of SCExAO residual chromatic effects on simultaneous differential imaging (SDI) has also not been evaluated, but is believed to be manageable over the $\sim$1% wide spectral band covered by each of the SDI filters. We note that SCExAO does not rely on SDI for differential detection: thanks to the high-order SCExAO DM, coherence-based techniques to separate residual speckles from true sources (as demonstrated in Guyon et al. 2009) will be the the primary speckle calibration mode.

REFERENCES

Abe, L., Murakami, N., Nishikawa, J., & Tamura, M. 2006, A&A, 451, 363

Aime, C., & Soummer, R. 2003, in EAS Publ. Ser., ed. C. Aime, & R. Soummer, 8, 79

Beuzit, J.-L., Feldt, M., & Dohlen, K., et al. 2006, Messenger, 125, 29

Biller, B. A., Close, L. M., & Masciadri, E., et al. 2007, ApJS, 173, 143

Cavarroc, C., Boccaletti, A., Baudoz, P., Fusco, T., & Rouan, D. 2006, A&A, 447, 397

Guyon, O. 2003, A&A, 404, 379

———. 2005, ApJ, 629, 592

Guyon, O., Pluzhnik, E. A., Galicher, R., Martinache, F., Ridgway, S. T., & Woodruff, R. A. 2005, ApJ, 622, 744

Guyon, O., Pluzhnik, E. A., Kuchner, M. J., Collins, B., & Ridgway, S. T. 2006, ApJS, 167, 81

Guyon, O., Pluzhnik, E. A., & Martinache, F., et al. 2009, PASP, submitted

Guyon, O., & Roddier, F. 2002, A&A, 391, 379

Hodapp, K. W., Suzuki, R., & Tamura, M., et al. 2008, Proc. SPIE, 7014, 701419

Kalas, P., Graham, J. R., & Chiang, E., et al. 2008, Science, 322, 1345

Kasdin, N. J., Vanderbei, R. J., Littman, M. G., & Spergel, D. N. 2005, Appl. Opt., 44, 1117

Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G. 2003, ApJ, 582, 1147

Kasper, M., Apai, D., Janson, M., & Brandner, W. 2007, A&A, 472, 321

Labeyrie, A. 1996, A&AS, 118, 517

Lafrenière, D., Doyon, R., & Marois, C., et al. 2007, ApJ, 670, 1367

Lagrange, A., Gratadour, D., & Chauvin, G., et al. 2008, ArXiv e-prints

Lyot, B. 1939, MNRAS, 99, 580

Macintosh, B. A., Graham, J. R., & Palmer, D. W., et al. 2008, Proc. SPIE, 7015, 701518

Marois, C., Macintosh, B., & Barman, T., et al. 2008, ArXiv e-prints

Martinache, F., Guyon, O., Pluzhnik, E. A., Galicher, R., & Ridgway, S. T. 2006, ApJ, 639, 1129

Martinache, F., Rojas-Ayala, B., Ireland, M. J., Lloyd, J. P., & Tuthill, P. G. 2009, ApJ, 695, 1183

Martinez, P., Boccaletti, A., & Kasper, M., et al. 2008, A&A, 492, 289

Pluzhnik, E. A., Guyon, O., & Ridgway, S. T., et al. 2006, ApJ, 644, 1246

Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, PASP, 111, 587

Roddier, F., & Roddier, C. 1997, PASP, 109, 815

Sivaramakrishnan, A., & Lloyd, J. P. 2005, ApJ, 633, 528

Soummer, R. 2005, ApJ, 618, L161

Soummer, R., Aime, C., & Falloon, P. E. 2003, A&A, 397, 1161

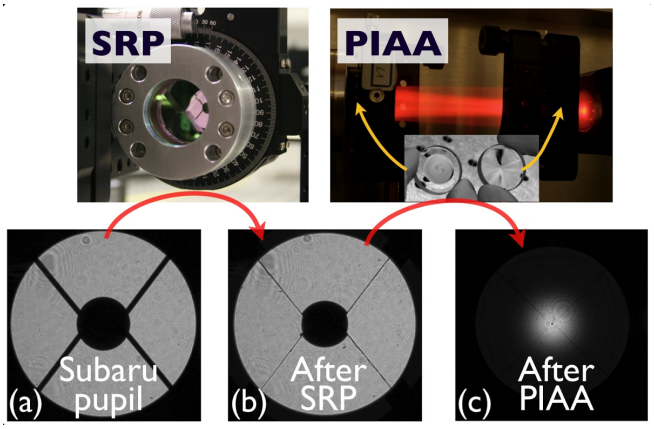


FIGURE 27: Adaptation de la pupille du Subaru Telescope à l'imagerie haut contraste avec la SRP (Spider Removal Plate) et le PIAA (Phase Induced Amplitude Apodization) pour éliminer le problème de la diffraction respectivement introduite par les araignées du télescope et les bords francs du miroir primaire du télescope. La rangée des trois images en bas compare la pupille originale du télescope (a) à la sortie de la SRP seule (b) et à celle de la combinaison SRP + PIAA (c). Grâce à ces optiques, la pupille a perdu toutes les caractéristiques qui la rendent impropre à l'imagerie haut contraste.

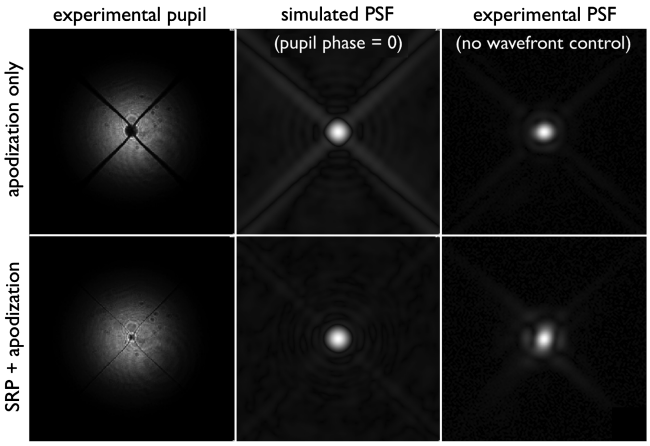


FIGURE 28: Pupilles de sortie et PSFs sur l'axe produites par une illumination uniforme de pupille de Subaru après remapping par les optiques PIAA seule (rangée du haut) et PIAA + SRP (rangée du bas).

SCExAO : architecture de l'instrument

Dans sa version originale, SCExAO devait former une unité destinée à être placée entre l'AO188 du télescope Subaru et la caméra scientifique HiCIAO, utilisée dans le cadre du relevé SEEDS[39]. A l'intérieur, se trouvait :

— un miroir déformable (DM) placé aussi près que possible d'un plan pupille en sortie de l'AO188 du télescope pour contrôler le front d'onde.

— une lame dichroïque séparant (à ∼900 nm) le visible de l'infrarouge.

— une voie scientifique infrarouge, guidant la lumière à travers les différentes optiques constituantes du PIAAC (ie incluant en plus du PIAA d'entrée, un PIAA inverse restorant le champ de vue).

— une caméra regardant la lumière rejetée par le masque plan focal, servant à monitorer l'alignement du faisceau sur le coronographe, faisant l'objet de la thèse de G. Singh.

— une voie technique visible pour un analyseur de surface d'onde (options : pyramide et courbure non-linéaire, thèse de C. Clergeon) contrôlant le DM commun à toutes les voies.

— une voie supplémentaire d'imagerie courte pose dans le visible, exploitant une généralisation de l'idée de "Lucky Imaging" faisant

39. M. Tamura. Subaru Strategic Exploration of Exoplanets and Disks with HiCIAO/AO188 (SEEDS). In T. Usuda, M. Tamura, and M. Ishii, editors, *American Institute of Physics Conference Series*, volume 1158 of *American Institute of Physics Conference Series*, pages 11–16, Aug. 2009

l'objet de la thèse de V. Garrel.

— une sortie collimatée, relayant la pupille sur la fenêtre d'entrée du cryostat de la caméra HiCIAO.

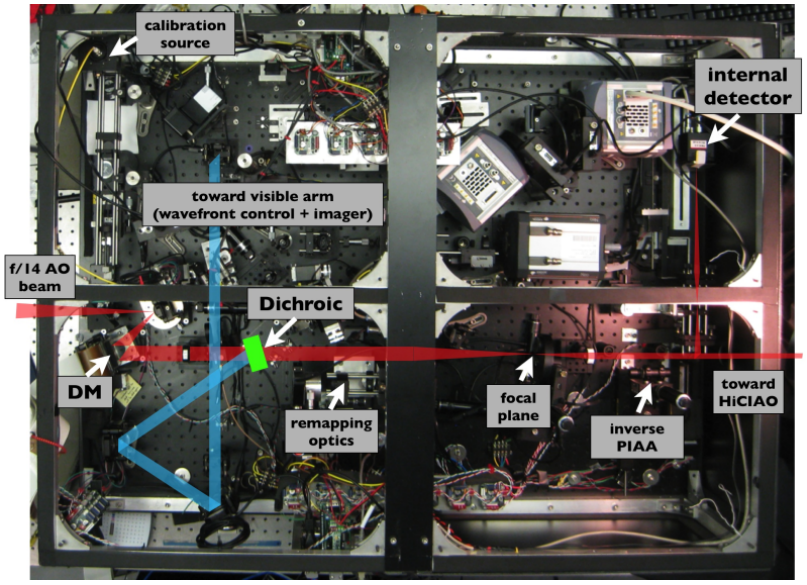


FIGURE 29: Photo vue de dessus de l'instrument SCExAO en 2011. L'entrée de la lumière se fait par la gauche : l'optique adaptative AO188 alimente SCExAO avec un faisceau convergent à f/13.92. Un miroir de champ et le miroir déformable (alors placé dans un plan intermédiaire entre l'image et la pupille), sont sur des montures motorisées permettant de contrôler l'alignement du faisceau avec le reste de l'instrument. Après la dichroïque, le faisceau IR traverse le banc en ligne droite, passant au travers des différentes optiques du PIAAC : PIAA, masque plan focal et PIAA inverse, avant de sortir en direction de la caméra HiCIAO. Pour le travail en laboratoire, un miroir de pick-off placé juste avant la sortie redirige le faisceau vers une caméra interne basse sensitivité.

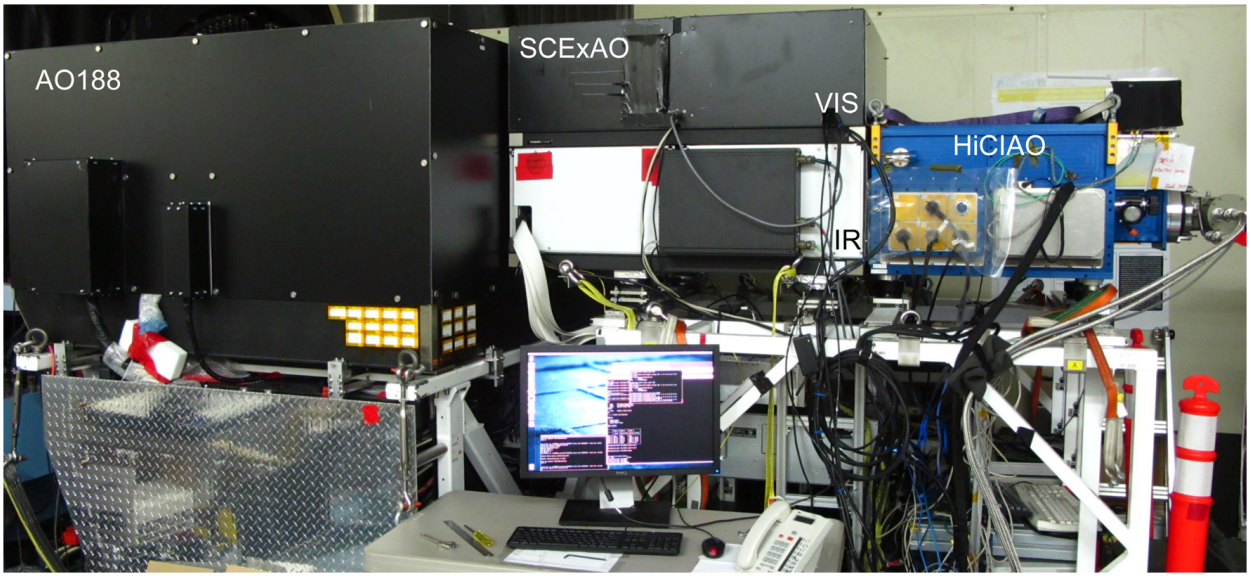


FIGURE 30: SCExAO sur la plateforme Nasmyth du Subaru Telescope. Alimenté par la gauche par l'AO188, SCExAO est maintenant constitué de deux tables optiques empilées (IR en bas, visible en haut). En sortie du banc IR, HiCIAO est utilisé comme imageur.

La figure 29 propose une photo annotée vue de dessus de ce banc original, construit sur la base d'une table optique de 1200x900 mm autour de laquelle, une cage (tube de section carré de 25mm en aluminium) a été rajoutée pour le protéger. Une des contraintes directrices de ce design inhabituel où le DM n'est pas exactement en plan pupille était la minimisation du nombre d'optiques à traverser entre l'AO188 et HiCIAO, de façon à minimiser l'impact de SCExAO en transmission. Le positionnement du DM hors plan pupille s'est avéré problématique, avec une instabilité de l'alignement de pupille

du télescope, une utilisation sub-optimale des actionneurs et une limitation du champ effecivement couvert. Cette configuration a cependant été mise à profit pour mettre au point les premières expériences de contrôle de speckles dans des images coronographiques avec PIAA [40]. Profitant d'une mise à jour attendue du DM pour un modèle avec un plus grand nombre d'actionneurs, le schéma optique global a été complètement revu, avec des optiques supplémentaires relayant la pupille sur le DM, et deux bancs optiques empilés (voir la figure 30) : un visible et un infrarouge, couplés par un système de périscope. Doubler la surface de tables optiques a évidemment permis de considérablement aérer le design du système. Cet espace a au fur et à mesure, été mis à profit pour intégrer de nouvelles fonctionnalités, en partie en provenance de collaborations externes proposant de nouveaux modules pour SCExAO.

40. F. Martinache, O. Guyon, C. Clergeon, and C. Blain. Speckle Control with a Remapped-Pupil PIAA Coronagraph. **PASP**, 124:1288–1294, Dec. 2012

La figure 31 ([41]) offre une vue d'ensemble de la composition de l'instrument ainsi réorganisé. Les fonctionnalités offertes sont décrites dans l'article inclus à cette section. Outre la supervision globale de ce changement majeur d'architecture et un suivi particulier sur la distribution de l'alimentation électrique dans l'instrument, le travail de R&D que je vais décrire par la suite concerne le banc IR et les outils à mettre en place pour améliorer le potentiel d'imagerie haut contraste d'un instrument observant depuis le sol.

41. N. Jovanovic, F. Martinache, O. Guyon, C. Clergeon, G. Singh, T. Kudo, V. Garrel, K. Newman, D. Doughty, J. Lozi, J. Males, Y. Minowa, Y. Hayano, N. Takato, J. Morino, J. Kuhn, E. Serabyn, B. Norris, P. Tuthill, G. Schworer, P. Stewart, L. Close, E. Huby, G. Perrin, S. Lacour, L. Gauchet, S. Vievard, N. Murakami, F. Oshiyama, N. Baba, T. Matsuo, J. Nishikawa, M. Tamura, O. Lai, F. Marchis, G. Duchene, T. Kotani, and J. Woillez. The Subaru Coronagraphic Extreme Adaptive Optics System: Enabling High-Contrast Imaging on Solar-System Scales. **PASP**, 127:890–910, Oct. 2015

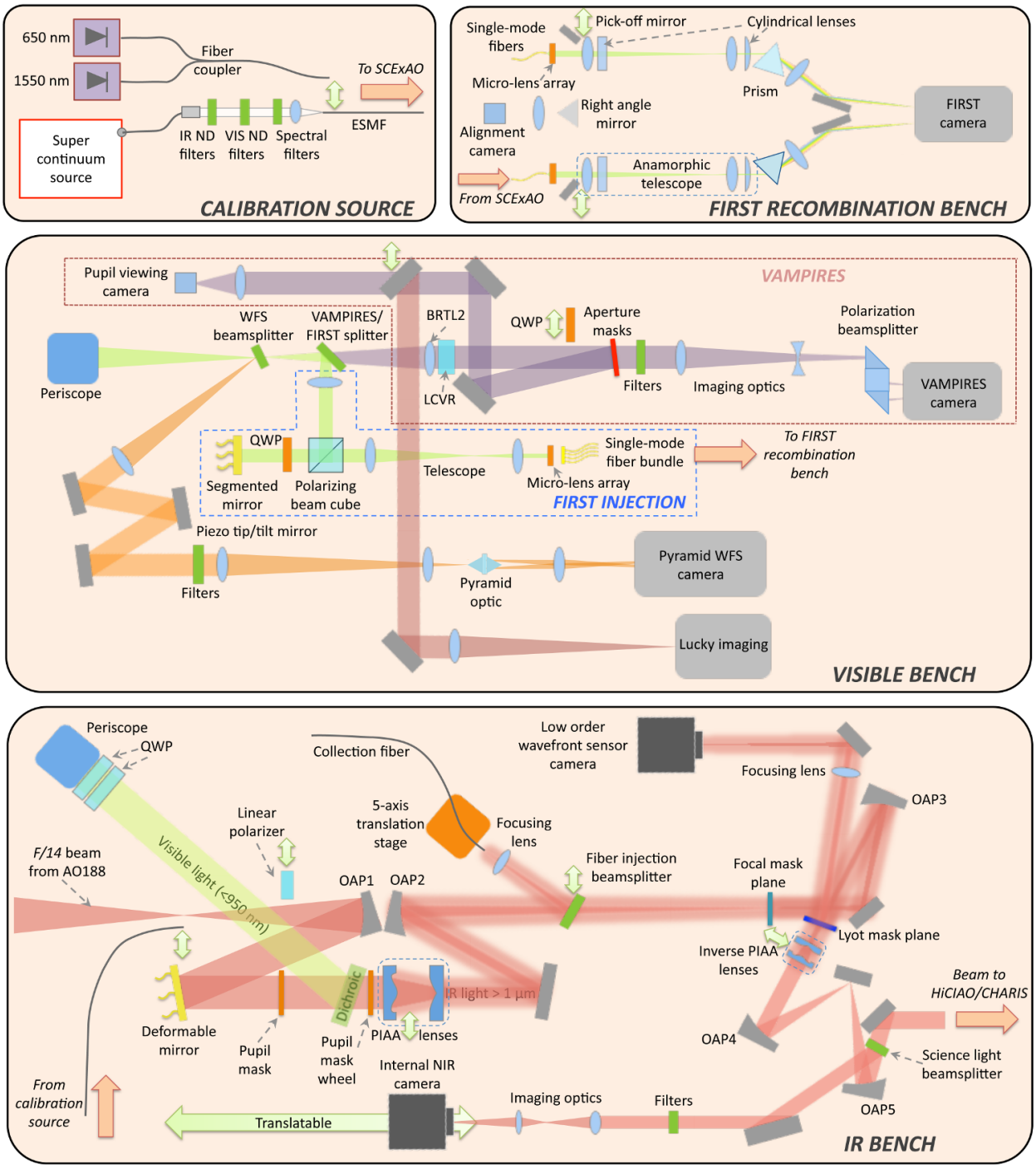


FIGURE 31: Design optique de SCExAO. L'instrument est constitué de deux bancs optiques superposés (infrarouge et visible), avec une séparation dichroïque à 900 nm. Les deux bancs partagent quelques optiques, dont le miroir déformable (DM), utilisé pour le contrôle de front d'onde. Mon travail a principalement concerné le banc infrarouge, qui a été dimensionné par le choix de l'architecture d'un coronagraphe de type PIAA.

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **127**:890–910, 2015 September
© 2015. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

The Subaru Coronagraphic Extreme Adaptive Optics System: Enabling High-Contrast Imaging on Solar-System Scales

N. JOVANOVIC,[1,2] F. MARTINACHE,[3] O. GUYON,[1,4,5] C. CLERGEON,[1] G. SINGH,[1,6] T. KUDO,[1] V. GARREL,[7] K. NEWMAN,[5,8] D. DOUGHTY,[1,5] J. LOZI,[1] J. MALES,[4,9] Y. MINOWA,[1] Y. HAYANO,[1] N. TAKATO,[1] J. MORINO,[10] J. KUHN,[11] E. SERABYN,[11] B. NORRIS,[12] P. TUTHILL,[12] G. SCHWORER,[6,12] P. STEWART,[12] L. CLOSE,[4] E. HUBY,[6,13] G. PERRIN,[6] S. LACOUR,[6] L. GAUCHET,[6] S. VIEVARD,[6] N. MURAKAMI,[14] F. OSHIYAMA,[14] N. BABA,[14] T. MATSUO,[15] J. NISHIKAWA,[10] M. TAMURA,[10,16] O. LAI,[1,7] F. MARCHIS,[17] G. DUCHENE,[18,19] T. KOTANI,[10] AND J. WOILLEZ[20]

Received 2015 February 19; accepted 2015 June 23; published 2015 August 31

ABSTRACT. The Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) instrument is a multipurpose high-contrast imaging platform designed for the discovery and detailed characterization of exoplanetary systems and serves as a testbed for high-contrast imaging technologies for ELTs. It is a multiband instrument which makes use of light from 600 to 2500 nm, allowing for coronagraphic direct exoplanet imaging of the inner $3\lambda/D$ from the stellar host. Wavefront sensing and control are key to the operation of SCExAO. A partial correction of low-order modes is provided by Subaru's facility adaptive optics system with the final correction, including high-order modes, implemented downstream by a combination of a visible pyramid wavefront sensor and a 2000-element deformable mirror. The well-corrected NIR (y-K bands) wavefronts can then be injected into any of the available coronagraphs, including but not limited to the phase-induced amplitude apodization and the vector vortex coronagraphs, both of which offer an inner working angle as low as $1\lambda/D$. Noncommon path, low-order aberrations are sensed with a coronagraphic low-order wavefront sensor in the infrared (IR). Low noise, high frame rate NIR detectors allow for active speckle nulling and coherent differential imaging, while the HAWAII 2RG detector in the HiCIAO imager and/or the CHARIS integral field spectrograph (from mid-2016) can take deeper exposures and/or perform angular, spectral, and polarimetric differential imaging. Science in the visible is provided by two interferometric modules: VAMPIRES and FIRST, which enable subdiffraction limited imaging in the visible region with polarimetric and spectroscopic capabilities respectively. We describe the instrument in detail and present preliminary results both on-sky and in the laboratory.

Online material: color figures

[1] National Astronomical Observatory of Japan, Subaru Telescope, 650 North A'Ohoku Place, Hilo, HI, 96720; jovanovic.nem@gmail.com.
[2] Department of Physics and Astronomy, Macquarie University, Sydney NSW 2109, Australia.
[3] Observatoire de la Cote d'Azur, Boulevard de l'Observatoire, Nice 06304, France.
[4] Steward Observatory, University of Arizona, Tucson, AZ 85721.
[5] College of Optical Sciences, University of Arizona, Tucson, AZ 85721.
[6] LESIA, Observatoire de Paris, 5 Place Jules Janssen, Meudon 92195, France.
[7] Gemini Observatory, c/o AURA, Casilla 603, La Serena, Chile.
[8] NASA Ames Research Center, Moffett Field, CA 94035.
[9] NASA Sagan Fellow.
[10] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Japan.
[11] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109.
[12] Sydney Institute for Astronomy (SIfA), Institute for Photonics and Optical Science (IPOS), School of Physics, University of Sydney, NSW 2006, Australia.
[13] Département d'Astrophysique, Géophysique et Océanographie, Université de Liège, 17 Allée du Six Août, 4000 Liège, Belgium.
[14] Division of Applied Physics, Faculty of Engineering, Hokkaido University, Kita-13, Nishi-8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan.
[15] Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan.
[16] Department of Astronomy, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan.
[17] Carl Sagan Center at the SETI Institute, Mountain View, CA 94043.
[18] Astronomy Department, University of California, Berkeley, CA 94720-3411.
[19] University Grenoble Alpes & CNRS, Institut de Planetologie et d'Astrophysique de Grenoble (IPAG), Grenoble F-3800, France.
[20] European Southern Observatory (ESO), Karl-Schwarzschild-Str. 2, Garching 85748, Germany.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

1. INTRODUCTION

The field of high-contrast imaging is advancing at a great rate with several extreme adaptive optics systems having come online in 2014, including the Gemini Planet Imager (GPI) (Macintosh 2014), the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE) (Beuzit et al. 2008), and the focus of this work, the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) system which join the already running P1640 (Dekany et al. 2013). These systems all share a similar underlying architecture: they employ a high-order wavefront sensor (WFS) and a deformable mirror (DM) to correct for atmospheric perturbations enabling high Strehl ratios in the near-infrared (NIR) ($>90\%$), while a coronagraph is used to suppress on-axis starlight downstream. The primary motivation for such instrumentation is the direct detection of planetary mass companions at contrasts of 10^{-5}–10^{-6} with respect to the host star, at small angular separations (down to 1–$5\lambda/D$) from the host star.

The era of exoplanetary detection has resulted in $\sim$1500 planets so far confirmed (Han et al. 2014). The majority of these were detected via the transit technique with instruments such as the *Kepler Space Telescope* (Borucki et al. 2010). The radial velocity method (Mayor & Queloz 1995) has also been prolific in detection yield. Both techniques are indirect in nature (the presence of the planet is inferred by its effect on light from the host star) and hence often deliver limited information about the planets themselves. It has been shown that it is possible to glean insights into atmospheric compositions via techniques such as transit spectroscopy (Charbonneau 2001), whereby star light from the host passes through the upper atmosphere of the planet as it propagates to Earth, albeit with limited signal-to-noise ratio. The ability to directly image planetary systems and conduct detailed spectroscopic analysis is the next step toward understanding the physical characteristics of their members and refining planetary formation models.

To this end, so far <50 substellar companions have been directly imaged (see Fig. 3 in Pepe et al. [2014]). The challenge lies in being able to see a companion, many orders of magnitude fainter, at very small angular separations ($<1''$), from the blinding glare of the host star. Indeed, the Earth would be $>10^9\times$ fainter than the sun if viewed from outside the solar system in reflected light. Although these levels of contrast cannot be overcome from ground-based observations at small angular separations ($<0.5''$), it is possible to circumvent this by imaging the thermal signatures instead and targeting bigger objects. Indeed, all planets imaged thus far were large Jupiter-like planets (which are brightest) detected at longer wavelengths (in the near-IR H and K-bands and the mid-IR L and M-bands) in thermal light (a subset of detections include Kraus & Ireland [2012]; Lagrange et al. [2009]; Marois et al. [2008]). To overcome the glare from the star which results in stellar photons swamping the signal from the companion, adaptive optics systems (AO) are key (Lagrange et al. 2009). Although angular differential imaging is the most commonly used technique for imaging planets thus far (Marois et al. 2008), coronagraphy (Lafreniere et al. 2009; Serabyn et al. 2010a) and aperture-masking interferometry (Kraus & Ireland 2012) have also been used to make detections. With the direct detection of the light from the faint companion itself, spectroscopy becomes a possibility and indeed preliminary spectra have been taken for some objects as well (Barman et al. 2011; Oppenheimer et al. 2013).

In addition to planetary spectroscopy, how disks evolve to form planetary systems is a key question that remains unanswered. Thus far coronagraphic imagers like HiCIAO at the Subaru Telescope have revealed intricate features of the inner parts of circumstellar disks using polarization differential imaging (under the SEEDS project [Tamura et al. 2009]). These solar-system scale features include knots and spiral density waves within disks like MWC758 and SAO 206462 (Grady et al. 2013; Muto et al. 2012). How such features are affected by or lead to the formation and evolution of planets can only be addressed by high-contrast imaging of the inner parts (up to 15 AU from the star) of such disks. To address the lack of information in this region, high-contrast imaging platforms equipped with advanced wavefront control and coronagraphs are pushing for smaller inner working angles (IWA). In the limit of low wavefront aberrations, currently achieved with AO systems operating in the near-IR, coronagraphs are the ideal tool for imaging the surrounding structure/detail as they are unrivaled in achievable contrast. Both the contrast and IWA are dependent on the level of wavefront correction available. With wavefront corrections typically offered by facility adaptive optics (AO) systems on 5–10-m class telescopes ($\sim$30–40% in H-band), previous generations of coronagraphic imagers, such as the Near-Infrared Coronagraphic Imager (NICI) on the Gemini South telescope (Artigau 2008), were optimized for an IWA of 5–$10\lambda/D$. However, with extreme AO (ExAO) correction offering high Strehl ratio and stable pointing, GPI and SPHERE have been optimized for imaging companions down to angular separations of $\sim 3\lambda/D$ (>120 mas in the H-band). SCExAO utilizes several more sophisticated coronagraphs including the Phase-Induced Amplitude Apodization (PIAA) (Guyon 2003) and vector vortex (Mawet et al. 2010), which drive the IWA down to just below $1\lambda/D$. At a distance of 100 pc, the PIAA/vortex coronagraphs on SCExAO would be able to image from 4 AU outward (approximately the region beyond the orbit of Jupiter). Further, in the case of the HR8799 system, the IWA would be 1.6 AU (the distance of Mars to our Sun) making it possible for SCExAO to image the recently hypothesized fifth planet at 9 AU (mass between 1 and 9 Jupiter masses) (Gozdziewski & Migaszewski 2014) if it is indeed present as predicted.

Despite the state-of-the-art IWA offered by these coronagraphs in the near-IR, the structure of disks and the distribution of planets at even closer separations than $1\lambda/D$ will remain inaccessible with coronagraph technology alone. This scale is scientifically very interesting as it corresponds to the inner parts of the solar system where the majority of exoplanets have been

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

found to date based on transit and radial velocity data (Han et al. 2014). To push into this regime, SCExAO uses two visible wavelength interferometric imaging modules known as VAMPIRES (Norris et al. 2012a) and FIRST (Huby et al. 2012). VAMPIRES is based on the powerful technique of aperture masking interferometry (Tuthill et al. 2000), while FIRST is based on an augmentation of that technique, known as pupil remapping (Perrin et al. 2006). Operating in the visible part of the spectrum, the angular resolution of these instruments on an 8-m class telescope approaches a territory previously reserved for long baseline interferometers (15 mas at $\lambda = 700$ nm) and expands the type of target that can be observed to include massive stars. Although the modules operate at shorter wavelengths where the wavefront correction is of a lower quality, interferometric techniques allow for subdiffraction limited imaging (10 mas resolution using $\lambda/2D$ criteria as conventionally used in interferometry) even in this regime, albeit at lower contrasts ($\sim 10^{-3}$), making optimal use of the otherwise discarded visible light. Despite the lower contrasts, aperture-masking interferometry has already delivered faint companion detections at unprecedented spatial scales (Kraus & Ireland 2012). Each module additionally offers a unique capability. For example, the polarimetric mode of VAMPIRES is designed to probe the polarized signal from dusty structures such as disks around young stars and shells around giant stars (Norris et al. 2012b, 2015) at a waveband where the signal is strongest. This is a visible analog of that offered by the SAMPol mode on the NACO instrument at the VLT (Lenzen et al. 2003; Norris et al. 2012b; Tuthill et al. 2010). FIRST, on the other hand, offers the potential for broadband spectroscopy and is tailored to imaging binary systems and the surface features of large stars. Such capabilities greatly extend SCExAO beyond that of a regular ex-AO facility.

Finally, with a diffraction-limited point-spread-function (PSF) in the near-IR, and a large collecting area, SCExAO is ideal for injecting light into fiber-fed spectrographs such as the Infrared Doppler instrument (IRD) (Tamura et al. 2012). In addition, this forms the ideal platform for exploring photonic-based technologies such as photonic lanterns (Leon-Saval et al. 2013) and integrated photonic spectrographs (Cvetojevic et al. 2012) for next-generation instrumentation.

The aim of this publication is to outline the SCExAO instrument and its capabilities in detail and offer some preliminary results produced by the system. In this vein, § 2 describes the key components of SCExAO while § 3 highlights the functionalities and limitations of the instrument. § 4 outlines plans for future upgrades and the paper concludes with a summary in § 5.

2. THE ELEMENTS OF SCEXAO

In order to understand the scientific possibilities and limitations of the SCExAO instrument, it is important to first understand the components and their functionalities. To aid the discussion, a system level diagram of SCExAO is shown in Figure 1, and an image of the instrument at the Nasmyth platform is shown in Figure 2. A detailed schematic of the major components is shown in Figure 3. The components and functionalities of SCExAO have been undergoing commissioning as will be outlined throughout this publication. A summary of

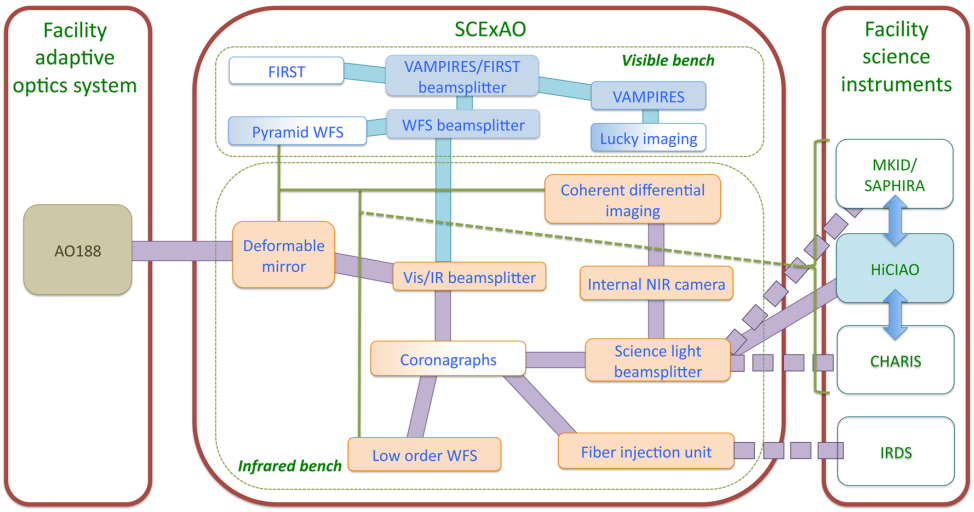


FIG. 1.—System level flow diagram of the SCExAO instrument. *Thick purple and blue lines* depict optical paths while thin green lines signify communication channels. *Dashed lines* indicate that a connection does not currently exist but it is planned for the future. *Filled boxes* indicate commissioning status. *Solid boxes*: commissioned; *graded boxes*: partially commissioned; *white background*: not commissioned. See the electronic edition of the *PASP* for a color version of this figure.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

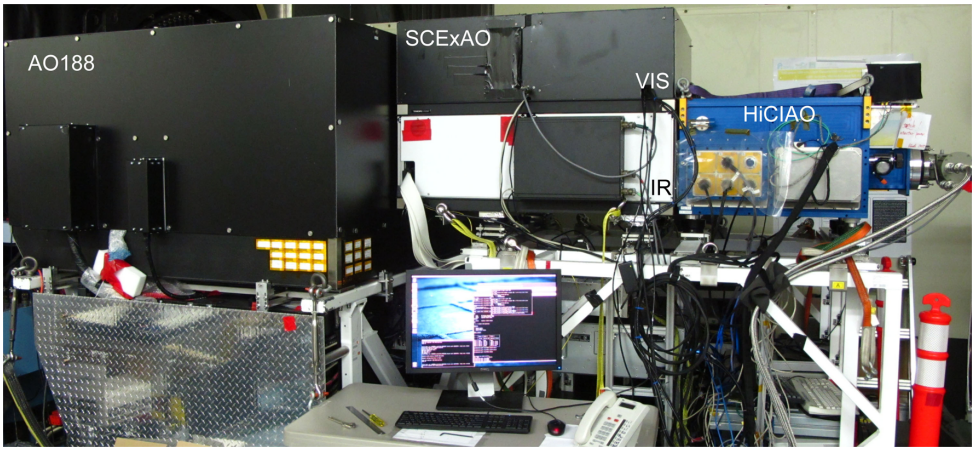


FIG. 2.—Image of SCExAO mounted at the Nasmyth IR platform at Subaru Telescope. To the left is AO188, which injects the light into SCExAO (center), and HiCIAO is shown on the right. The FIRST recombination bench is not shown for visual clarity. See the electronic edition of the *PASP* for a color version of this figure.

the commissioning status of each mode of operation or module of SCExAO can be found in the tables.

The main aim of SCExAO is to exploit the well-corrected wavefront enabled by the high-order WFS to do high-contrast imaging with light across a broad spectrum: from 600 to 2500 nm. As such, there are a number of instrument modules within SCExAO that operate in different wavebands simultaneously while the coronagraph is collecting data. This hitchhiking mode of operation enables maximum utilization of the stellar flux, which allows for a more comprehensive study of each target.

2.1. SCExAO at a Glance

The SCExAO instrument consists of two optical benches mounted on top of one another, separated by ~350 mm. The bottom bench (IR bench) hosts the deformable mirror, coronagraphs, and a Lyot-based low-order wavefront sensor (LLOWFS) while the top bench (visible bench) hosts the pyramid WFS, VAMPIRES, FIRST, and lucky imaging (see Fig. 3). The benches are optically connected via a periscope.

The light from the facility adaptive optics system (AO188) is injected into the IR bench of SCExAO and is incident on the 2000 element deformable mirror (2k DM) before it is split by a dichroic into two distinct channels: light shorter than 940 nm is reflected up the periscope and onto the top bench while light longer than 940 nm is transmitted. The visible light is then split by spectral content by a range of long- and short-pass dichroics which send the light to the pyramid WFS (PyWFS) and visible light science instruments. The PyWFS is used for the high-order wavefront correction and drives the DM on the IR bench. The VAMPIRES and FIRST modules utilize the light not used by the PyWFS. Lucky imaging/PSF viewing makes use of light rejected by the aperture masks of VAMPIRES.

The IR light that is transmitted by the dichroic on the IR bench propagates through one of the available coronagraphs. After the coronagraphs, the light reflected by the Lyot stop is used to drive a LLOWFS in order to correct for the chromatic and noncommon errors (such as tip/tilt) between the visible and IR benches (Singh et al. 2014). The light transmitted by the coronagraphs is then incident on the science light beamsplitter which determines the spectral content and exact amount of flux to be sent to a high frame rate internal NIR camera as compared to a science grade detector such as the HAWAII 2RG in the HiCIAO instrument and soon to be commissioned CHARIS. The internal NIR camera can then be used to drive various coherent differential imaging algorithms.

2.2. Detailed Optical Design

The instrument is designed to receive partially corrected light from the facility adaptive optics system, AO188 (188 actuator deformable mirror). The beam delivered from AO188 converges with a speed of $f/14$. Typical H-band Strehl ratios are ~30–40% in good seeing (Minowa et al. 2010). The beam is collimated by an off-axis parabolic mirror (OAP1, $f = 255$ mm) creating an 18 mm beam. Details of the OAPs can be found in the appendix. The reflected beam is incident upon the 2k DM, details of which are in § 3.1.1. The surface of the DM is placed one focal length from OAP1 which conjugates it with the primary mirror of the telescope (i.e., it is in a pupil plane). Once the beam has reflected off the DM, it is incident upon a fixed pupil mask which replicates the central obstruction and spiders of the telescope, albeit slightly oversized. This mask is permanently in the beam (both on-sky and in the laboratory)

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

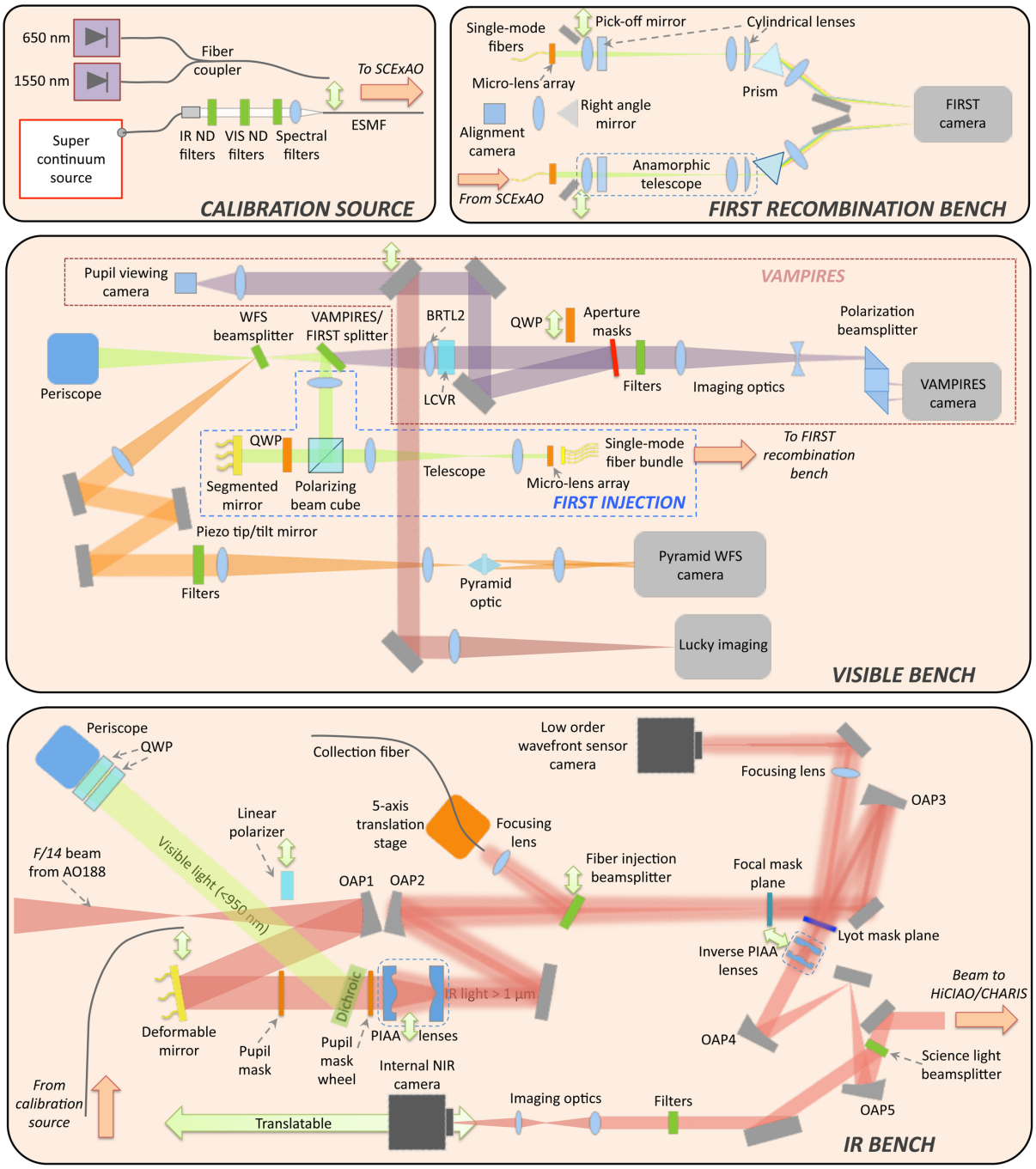


FIG. 3.—Schematic diagram of the SCExAO instrument. Top box (left): Portable calibration source layout. Top box (right): FIRST recombination bench. Middle box: layout of the visible optical bench which is mounted on top of the IR bench. Bottom box: IR bench layout. *Dual head green arrows* indicate that a given optic can be translated in/out of or along the beam. *Orange arrows* indicate light entering or leaving the designated bench at that location. See the electronic edition of the *PASP* for a color version of this figure.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

so that response matrices collected for the various wavefront sensors with the internal light source in SCExAO can be used on-sky as well. It is positioned as close to the pupil plane as possible (~70 mm away from the DM) and forms the primary pupil for the instrument. The image/pupil rotator in AO188 is used to align the telescope pupil with the fixed internal mask when on-sky (i.e., SCExAO operates with a fixed pupil and rotating image when observing).

Immediately following the mask is a dichroic beamsplitter (50 mm diameter, 7 mm thick) which reflects the visible light (<940 nm) and transmits the IR (>940 nm). In the transmitted beam path, there is a mask wheel after the dichroic which hosts numerous masks including the shaped pupil coronagraphic mask. The automated mounts for the phase-induced amplitude apodization (PIAA) coronagraph lenses are adjacent to the mask wheel. They too were placed as close to the DM pupil as possible. The PIAA lenses themselves will be described in detail, for now it is important to note that they can be retracted from the beam entirely. A flat mirror is used to steer the light onto OAP2 which focuses the beam ($f = 519$ mm). A variety of coronagraphic focal plane masks, used to suppress the on-axis star light, are housed in a wheel which also has three axes of translation in the focal plane (the masks are outlined in § 3.2). OAP3 recollimates the beam to a 9 mm diameter beam ($f = 255$ mm). A wheel with Lyot plane masks is situated in the collimated beam such that the masks are conjugated with the pupil plane. The Lyot wheel can be adjusted in lateral alignment via motorized actuators. Light diffracted by the focal plane mask and reflected from the Lyot stop is imaged onto a nonscience grade (i.e., relatively high noise), high frame rate detector which is used for low-order wavefront sensing. The light transmitted by the Lyot stop is next incident upon the inverse PIAA lenses (detailed in § 3.2). They are mounted on stages which allow motorized control of the lateral positioning and are conjugated to the PIAA lenses upstream and can also be retracted from the beam entirely. OAP4 and OAP5 reimage the telescope pupil for the HiCIAO/CHARIS camera. On its way to the camera, the beam is intercepted by the science light beamsplitter which is used to control the flux and spectral content sent to a high frame rate internal NIR camera and the facility science instruments. Detectors in a high-contrast imaging instrument are among the most important components, and therefore, the performance of those used in SCExAO is summarized in Table 1.

The light directed to the internal NIR camera passes through a filter wheel for spectral selectivity. The content of the filter wheel and the science light beamsplitter wheels are reported in Table 2. An image is formed on the internal NIR camera via a pair of IR achromatic lenses (see appendix for details). The sampling and field-of-view on this and other cameras is summarized in Table 3. The internal NIR camera is mounted on a translation stage allowing it to be conjugated to any location between the focal and pupil planes.

The light reflected by the dichroic on the bottom bench (which splits the visible and IR channels) is directed toward a periscope which transports it to the upper bench. An achromatic lens (50 mm diameter, $f = 500$ mm) is mounted in the periscope to reimage the pupil onto the top bench. A wheel hosts a range of dichroic beamsplitters at the focus of the beam on the top bench to select the spectral content to be directed toward the PyWFS (named WFS beamsplitter; content listed in Table 2). The light reflected is collimated by an achromatic lens ($f = 200$ mm) and the pupil image is located on a piezo-tip/tilt mirror used to modulate the PyWFS. The reflected beam is routed to focusing optics which form a converging beam with $f/40$ (a combination of achromatic lenses with focal lengths of $f = 400$ and $f = 125$ mm are used for this). A double pyramid prism shaped optic (Esposito et al. 2010) is placed at the focus of the beam such that the vertex is on the optical axis. An image of the resulting four pupils is generated via one additional lens ($f = 300$ mm) on the detector. The PyWFS is discussed in more detail in § 3.1.2.

The light which is not directed to the WFS is split between two modules: VAMPIRES and FIRST. The basic optical layouts are outlined here and details in regards to specifics, including

TABLE 1
DETECTOR CHARACTERISTICS USED WITHIN SCEXAO

Detector name	Technology	Detector size (pixels)	Pixel size (μm)	Read-out noise (e^-)	Full frame (Frame rate [Hz])	Operating wavelengths (nm)	Manufacturer (Product name)
Internal NIR camera	InGaAs (CMOS)	320×256	30	114	170	900–1700	Axiom Optics (OWL SW1.7HS)
LOWFS camera	InGaAs (CMOS)	320×256	30	140	170	900–1700	Axiom Optics (OWL SW1.7HS)
HiCIAO/CHARIS	HgCdTe (CMOS)	2048×2048	18	15–30	<3	850–2500	Teledyne (HAWAII 2RG)
PyWFS (previous)	Si (sCMOS)	2560×2160 (128×128)	6.5	1.2	100 (1700)	350–1000	Andor (Zyla)
PyWFS (current)	Si (EMCCD)	240×240 (120×120)	24	<0.3	2000 (3700)	350–1000	Firstlight Imaging (OCAM2K)
Lucky imaging	Si (EMCCD)	512×512	16	<1	35	300–1000	Andor (iXon3 897)
VAMPIRES	Si (EMCCD)	512×512	16	<1	56	300–1000	Andor (iXon Ultra 897)
FIRST	Si (EMCCD)	512×512	16	<1	31.9	300–1000	Hamamatsu (ImagEM C9100-13)

NOTES. —The values in the table are taken from manufacturer specifications/measurements. Values in parentheses indicate the subframed (previous PyWFS) or binned (current PyWFS) image size and corresponding frame rate used. Read-out noise quoted for EMCCDs is with gain applied. Quantum efficiencies for these detectors are shown in Table 7.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

TABLE 2
FILTER AND BEAMSPLITTER WHEEL CONTENTS

Slot	Internal camera filters	Science light beamsplitters	Wavefront sensor beamsplitter	VAMPIRES/FIRST splitter
1	Open	AR-coated window (900–2600 nm)	Open	Silver mirror
2	T = y-band	$T, R = 50\%$ (900–2600 nm)	Silver mirror	$T, R = 50\%$, 600–950 nm
3	T = J-band	$T = 10\%$, $R = 90\%$ (900–2600 nm)	$T, R = 50\%$, 600–950 nm	$T < 700$ nm, $R > 700$ nm
4	T = H-band	$T = 90\%$, $R = 10\%$ (900–2600 nm)	$T < 650$ nm, $R > 650$ nm	$T > 700$ nm, $R < 700$ nm
5	50 nm bandpass at 1600 nm	$T > 90\%$ (1400–2600 nm), $R > 95\%$ (900–1400 nm)	$T < 700$ nm, $R > 700$ nm	AR-coated window (600–950 nm)
6	–	$T > 90\%$ (1900–2600 nm), $R > 95\%$ (900–1900 nm)	$T < 750$ nm, $R > 750$ nm	Open
7	–	Gold mirror	$T < 800$ nm, $R > 800$ nm	–
8	–	–	$T < 850$ nm, $R > 850$ nm	–
9	–	–	$T > 750$ nm, $R < 750$ nm	–
10	–	–	$T > 800$ nm, $R < 800$ nm	–
11	–	–	$T > 850$ nm, $R < 850$ nm	–

NOTES. —The values in the table are based on the final measurements made by the manufacturer. T, transmission; R, reflection; AR, anti-reflection. Items 5 and 6 of the science light beamsplitter have not been delivered yet.

calibration and performance, are given in other publications (Huby et al. 2012; Norris et al. 2015). For engineering purposes, a gray beamsplitter is used to divide the light between VAMPIRES and FIRST ($50/50$), but can be swapped for several other optics if the science demands it (contents of the VAMPIRES/FIRST splitter wheel are shown in Table 2).

The light transmitted by the splitter is used by the VAMPIRES module. It is first collimated (by BRTL2) and then passes through a series of optics including a liquid crystal variable retarder (LCVR), pupil masks, and spectral filters before being focused by a combination of a converging and diverging lens onto a low noise detector. A polarization beamsplitter, consisting of a polarizing beam cube and three right-angled prisms, is placed in the beam after the final focusing lens and is used to spatially separate the two orthogonally polarized components on the detector. The VAMPIRES module combines aperture-masking interferometry with polarimetry. The sparse aperture masks are housed in the pupil wheel where an assortment of masks with various throughputs and Fourier coverage can be found (please see Norris et al. [2015] for details). The module operates on 50 nm bandwidths of light selected within the 600–800 nm range, via a set of spectral filters in order to maintain fringe visibility, while maximizing the signal-to-noise ratio. At 650 nm, VAMPIRES can achieve an angular resolution of 8.4 mas (Fizeau criteria) and has a field-of-view in masking mode which ranges from 80 mas to 460 mas depending on the mask selected (larger fields-of-view, ∼1–2″ are possible in normal imaging mode, no mask). Rather than simply blocking the unwanted light from the pupil, the masks are reflective mirrors with transmissive subapertures so that the unwanted light is redirected to a pupil-viewing camera (PtGrey, Flea, FL3-U3-13S2M-CS) which allows for fine alignment of the masks with the pupil. The pupil-viewing mode of VAMPIRES is only used when aligning the masks. To utilize the light when pupil-viewing is not being used a mirror is translated into the beam to direct the light to the Lucky imaging module/point-spread-function (PSF) viewer (described in detail in § 3.3).

TABLE 3
PLATE SCALE AND FIELD-OF-VIEW

Detector	Sampling (mas/pixel)	Field-of-view (″)
Internal NIR camera	12.1 ± 0.1	4×3
HiCIAO	8.3 ± 0.1	$>10 \times 10$
VAMPIRES (M)	–	0.08–0.46
VAMPIRES (NM)	6.0	2×1
FIRST	–	0.1

NOTES. —All values in this table were measured off-sky by moving the calibration source laterally in the input focal plane (AO188 focal plane) and determining this motion in pixels on the detector. This motion is converted to a plate scale based on the knowledge of the AO188 plate scale which is well known to be $1.865''$/mm. This method yields consistent values to those obtained by looking at astrometric binaries. M, with aperture masks; NM, no masks. Range of field-of-view is dependent on choice of mask. Please see Norris et al. (2015) for more details.

The high angular resolution imaging capability of VAMPIRES is boosted by an advanced polarimetric capability. This gains its strength from the multitiered differential calibration scheme which is utilized. First, two quarter wave plates (HWP) mounted in front of the periscope on the bottom bench can be used to compensate for the birefringence induced by the mirrors in AO188 and SCExAO. The setting of these plates is done by careful calibration beforehand. Fast polarization modulation (∼50 Hz) comes from the LCVR which is switched between every image. The analyzer splits the signal into two distinct interferograms on the detector with orthogonal polarizations. Finally, a HWP positioned in front of AO188, used for HiCIAO polarimetric imaging in the NIR, was determined to work well in the visible and is used as the main polarization switching component in VAMPIRES. The quarter wave plate (QWP) before the aperture mask wheel and the polarizer on the bottom bench are used for calibrating the polarization systematics with the internal calibration source and can be swung into the beam when off-sky (they are not used on-sky). For more details about

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

the nested differential calibration procedure please refer to Norris et al. (2015).

The light reflected by the beam splitter is sent to the FIRST module. It is collimated before entering a polarizing beamsplitter cube. The main beam is reflected at 90° onto a 37-element segmented mirror (Iris AO [Helmbrecht 2011]) which is conjugated to the pupil plane. A QWP placed between the polarizing beamsplitter and the segmented mirror is used to rotate the polarization so that the beam passes through the cube when reflected off the segmented mirror. A beam-reducing telescope compresses the beam so that a single segment of the mirror has a one-to-one correspondence with a microlens in the microlens array (MLA) used for injecting into the bundle of single-mode polarization maintaining fibers (see Fig. 3). This architecture allows the segmented mirror to fine-tune the coupling into each of the fibers with small tip/tilt control. Currently 18 of the 30 available fibers (2 sets of 9 fibers each) are used and they transport the light to a separate recombination bench (see Fig. 3) where the interferograms are formed and data collected. A description of the recombination bench is beyond the scope of this work. The instrument offers an angular resolution of 9 mas at 700 nm with a 100 mas field-of-view ($\sim 6\lambda/D$ in radius). In addition, broadband operation from 600 to 850 nm, with a spectral resolving power of 300, offers a new avenue to maximizing data output while using standard bispectrum analysis techniques (and hence precision/contrast) while simultaneously allowing spectra to be collected.

FIRST has had several successful observing campaigns on close binary stars at Lick Observatory. It has now made the move from the Cassegrain focus of the Shane telescope to the gravitationally invariant Nasmyth platform of Subaru Telescope, which minimizes mechanical flexure and hence instrumental instabilities. A more comprehensive description of the instrument including the recombination bench, how it works, and the initial science results is presented in Huby et al. (2012, 2013).

Despite all the advanced wavefront control, interferometers and coronagraphs in SCExAO, the performance of the system is highly dependent on the stability of the PSF. Vibrations can plague high-contrast imaging testbeds via flexure, windshake of mirrors, or moving parts (e.g., cryocoolers). For this reason several key efforts have been made to address PSF stability in SCExAO. First, four elastomer-based vibration isolators (Newport, M-ND20-A) are used to support the SCExAO bench at the Nasmyth platform. These elastomers have a natural frequency 9–12 Hz and are used to damp high-frequency mechanical vibrations from the surrounding environment. Second, the mounts for the OAP-based relay optics in the instrument were custom made from a single piece to minimize the drift of the PSF. Finally, the mounting scheme for the HiCIAO cryocooler was rebuilt such that the pump was connected to the dewer by springs and metal bellows only (soft connections with low spring constants for maximum damping). This reduced the tip/tilt jitter observed on the NIR internal camera by a factor of 10 at 10 Hz and up to 100 at a resonant frequency of 23 Hz (Jovanovic et al. 2014a). These steps have improved overall PSF stability against long-term drifts and vibrations above a few Hz. However, there are small residual resonances which will be addressed with a linear quadratic Gaussian (LQG) controller (Poyneer et al. 2014) implemented to the LLOWFS loop in future.

2.3. Internal Calibration

An important feature of any high-contrast testbed is the ability to internally calibrate it off-sky. The SCExAO instrument can be aligned/calibrated with its internal calibration source. The source can be seen in Figure 3 and consists of a standalone box which houses a super continuum source (Fianium–Whitelase micro) for broadband characterization, and two fiber-coupled laser diodes (675 and 1550 nm) for alignment. The light from the super continuum source is collimated and passes through a series of wheels which house neutral density filters for both visible and IR wavelengths as well as a selection of spectral filters. The light is coupled into an endlessly single-mode photonic crystal fiber (NKT photonics - areoGUIDE8) which transports the light to the SCExAO bench. The fiber is mounted on a translation stage and can be actuated into the focus of the AO188 beam (see Fig. 3) when internal calibration is performed or the instrument is not at the telescope. The endlessly single-mode fiber is ideal for this application as it offers a diffraction-limited point source at all operating wavelengths of the super continuum source and SCExAO (600–2500 nm).

The effects of atmospheric turbulence can also be simulated. This is achieved by using the DM to create a large phase-screen with the appropriate statistical fluctuations. A single-phase screen Kolmogorov profile is used where the low spatial frequencies can be attenuated, to mimic the effect of an upstream AO system like AO188. A 50×50 pixel subarray (corresponding to the actuators of the DM) is extracted from the larger phase screen. By scanning the subarray across the phase map, a continuous and infinite sequence of phase screens can be generated. The amplitude of the RMS wavefront map, the magnitude of the low spatial frequency modes, and the speed of the subarray passing over the map (i.e., windspeed) are all free parameters that can be adjusted. This simulator is convenient as it allows great flexibility when characterizing SCExAO modules. Note that, due to the limited stroke of the DM (which is discussed in § 3.1.1), the turbulence simulator cannot be used to simulate full seeing conditions but provides a level of wavefront perturbation that is representative of post AO188 observing conditions. Finally, although the simulator can provide turbulence with the correct spatial structure pre- or post-AO, it does not take into account the temporal aspects of the correction provided by an upstream AO system (i.e., AO corrects low temporal frequencies leaving only higher frequencies).

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

2.4. Instrument Throughput

To plan observing schedules and determine the limitations of the instrument, the throughput was accurately characterized. As SCExAO has many branches, Tables 4 and 5 summarize the measured internal system throughput for each. Although not explicitly listed, the throughputs include all flat mirrors required to get the light to a given module. In addition, Table 6 highlights the measured throughputs of the optics upstream from SCExAO which includes the effects of the atmosphere, telescope, AO188 HWP, atmospheric dispersion compensator (ADC), AO188 optics, as well as HiCIAO. To determine the total throughput to a given detector plane of the instrument, one should first use Table 2 to select the appropriate beamsplitter, and then find the corresponding throughput for that branch from Table 4 or 5. This value should then be multiplied by the relevant throughputs in Table 6. To convert to photoelectrons, the quantum efficiencies for the detectors used in SCExAO are listed in Table 7. Finally, the throughputs of the focal plane masks used for the coronagraphs are not shown in Table 4 as they are specifically designed to attenuate the light. Indeed, the throughput of these masks depends on the distance from the optic axis, and for this, we refer the reader to literature such as Guyon (2005).

From the tables, we can determine the throughput to several focal planes within SCExAO as an example of the system performance and a demonstration of how to use the tables. One commonly used observing mode involves directing the majority

TABLE 4

THROUGHPUT OF THE VARIOUS ARMS OF THE IR CHANNEL OF SCExAO MEASURED IN THE LABORATORY FROM THE INTERNAL CALIBRATION SOURCE

Path (elements)	Band	Throughput (%)
From calibration source input to…		
Internal NIR camera		
(OAPs, DM, SLB2)	y	35.3
(OAPs, DM, SLB3)	y	60.6
(OAPs, DM, SLB4)	y	4.5
(OAPs, DM, SLB7)	y	71.1
(OAPs, DM, SLB2)	J	27.2
(OAPs, DM, SLB3)	J	53.7
(OAPs, DM, SLB4)	J	5.9
(OAPs, DM, SLB7)	J	50.4
(OAPs, DM, SLB2)	H	32.8
(OAPs, DM, SLB3)	H	56.2
(OAPs, DM, SLB4)	H	9.4
(OAPs, DM, SLB7)	H	63.4
Facility science instruments		
(OAPs, DM, SLB1)	y	72.1
(OAPs, DM, SLB2)	y	34.9
(OAPs, DM, SLB3)	y	7.6
(OAPs, DM, SLB4)	y	67.2
(OAPs, DM, SLB1)	J	66.9
(OAPs, DM, SLB2)	J	33.4
(OAPs, DM, SLB3)	J	6.2
(OAPs, DM, SLB4)	J	60.5
(OAPs, DM, SLB1)	H	65.2
(OAPs, DM, SLB2)	H	31.8
(OAPs, DM, SLB3)	H	5.0
(OAPs, DM, SLB4)	H	58.2
Single-mode injection		
(OAPs, DM, FIB)	J	77.0
(OAPs, DM, FIB)	H	78.3
Components in isolation		
(PIAA + Binary Mask + IPIAA)	y	53.2
(PIAA + Binary Mask + IPIAA)	J	52.2
(PIAA + Binary Mask + IPIAA)	H	52.5

NOTES. —OAPs, off-axis parabolic mirrors used to relay the light; DM, the window and mirror of the DM; SLB, science light beamsplitter; the number designates the slot (see Table 2); FIB, fiber injection beamsplitter.

TABLE 5

THROUGHPUT OF THE VARIOUS ARMS OF THE VISIBLE CHANNEL OF SCExAO MEASURED IN THE LABORATORY FROM THE INTERNAL CALIBRATION SOURCE

Path (elements)	Band	Throughput (%)
From calibration source input to…		
Pre-WFSBS		
(PTTB)	R	57.8
	I	65.4
	z	65.6
PyWFS camera		
(PTTB + PyWFS optics WFSBS 2)	R	43.0
	I	47.1
	z	47.6
(PTTB + PyWFS optics, WFSBS 3)	R	20.5
	I	23.0
	z	25.2
(PTTB + PyWFS optics, WFSBS 4, 9, 10 or 11)	R	42.4
(PTTB + PyWFS optics, WFSBS 4, 5, 6 or 7)	I	49.1
(PTTB + PyWFS optics, WFSBS 4, 5, 6, 7 or 8)	z	48.7
VAMPIRES camera		
(PTV + WFSBS 3)	R	12.6
	I	15.3
	z	15.2
(PTV + WFSBS 5, 6, 7, or 8)	R	25.3
(PTV + WFSBS 8)	I	27.5
(PTV + WFSBS 9, 10, or 11)	z	30.6)
FIRST input		
(PTF + WFSBS 3)	R	12.6
	I	15.3
	z	13.4
(PTF + WFSBS 5, 6, 7 or 8)	R	27.1
(PTF + WFSBS 8)	I	29.0
(PTF + WFSBS 9, 10 or 11)	z	28.6

NOTES. —OAP1, the first off-axis parabolic mirror used to relay the light; DM, the window and mirror of the DM; WFSBS, wavefront sensor beamsplitter; the number designates the slot (see Table 2); BRTL1: beam-reducing telescope lens 1, which is located in the periscope. Path to top bench (PTTB): OAP1, DM, dichroic, QWP ×2, periscope mirrors, BRTL1. Path to VAMPIRES (PTV): OAP1, DM, dichroic, QWP ×2, periscope mirrors, BRTL1 & 2, 50/50 BS, and focusing lenses. Path to FIRST (PTF): OAP1, DM, dichroic, QWP ×2, periscope mirrors, BRTL1, 50/50 BS, and collimating lens. Note that the throughput is only quoted to the input of FIRST. Also, the throughputs of the aperture masks for VAMPIRES were not included; please see Norris et al. (2015) for these values.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

TABLE 6

THROUGHPUT OF THE ATMOSPHERE, TELESCOPE, AO188 HWP, ATMOSPHERIC DISPERSION COMPENSATOR (ADC), AO188 OPTICS, AND HICIAO AS A PERCENTAGE

Band	Atmos.	Tele.	HWP	ADC	AO188	HiCIAO
R	91	82	93	95	68	–
I	96	75	95	95	69	–
z	97	79	95	95	69	–
y	98	80	94	95	72	–
J	95	91	96	94	77	87
H	97	92	94	92	79	78
K	92	93	92	70	82	82

NOTES. —The atmospheric transmission in R, I, Z bands is estimated from Buton et al. (2012). The throughput of the atmosphere in the y-, J-, H-, and K-bands assumes an airmass of 1.5 and precipitable water vapor content of 1.6 mm (data taken from Gemini Observatory website) (Lord 1992). The throughput of the telescope is calculated from the reflectivity of the material used for the coating and does not include effects like diattenuation from M3, for example. The throughputs for the HWP and ADC were measured on-sky. The throughputs for R, I, Z, and y band for AO188 have been calculated from materials. J-, H-, and K-band throughputs for AO188 were measured.

of the postcoronagraphic light toward HiCIAO for H-band imaging. In this case, one would select the 90% transmitting beamsplitter in the science light beamsplitter wheel (Slot 4 from Table 2). In this case, the throughput of SCExAO would be 9.4% to the internal NIR camera and 58.2% to HiCIAO (from Table 4). To determine the total throughput from the top of the atmosphere, one would need to multiply the SCExAO throughputs by those listed for the atmosphere (97%), telescope (92%), HWP (94%), ADC (92%), AO188 (79%), and HiCIAO optics/filters (78%) by those listed in Table 6, which yields values of 4.5% and 27.7%, respectively. From here, one would need to take into account the losses of the focal plane and Lyot masks which are coronagraph-specific. Finally, it is possible to use the quantum efficiencies listed in Table 7 to convert the throughput of the instrument into the number of photoelectrons on the HiCIAO detector for a given magnitude target.

In addition to flux at the detector plane, the throughput to the primary wavefront sensor is a useful piece of information for appropriate target selection. Since the PyWFS does most of its sensing at 850–900 nm, one could operate with an 800 nm short-pass filter in the wavefront sensor beamsplitter wheel (slot 7 from Table 2). The throughput of SCExAO to the WFS detection plane with this splitter would be 48.7% in the z-band (from Table 5). The associated throughput as measured from the top of the atmosphere would be 23.2% (including the z-band throughputs from Table 6). This relatively low throughput is attributed to the fact the PyWFS is a long way downstream of the telescope and even within SCExAO. In the future, if SCExAO were to be rebuilt, one could imagine moving the WFS to be the first element after the telescope and ADC which would no doubt improve the throughput by a factor of 2 at least. This would allow the PyWFS to operate with high performance on a 1 magnitude fainter star.

TABLE 7

QUANTUM EFFICIENCY AS A FUNCTION OF WAVEBAND FOR THE DETECTORS IN SCEXAO AS A PERCENTAGE

Detector	R	I	z	y	J	H	K
Internal NIR camera	–	–	–	58	73	73	–
LOWFS camera	–	–	–	58	73	73	–
HiCIAO/CHARIS	–	–	–	–	46	68	68
PyWFS (current)	98	96	77	–	–	–	–
Lucky imaging	93	74	47	–	–	–	–
VAMPIRES	93	74	42	–	–	–	–
FIRST	91	70	42	–	–	–	–

NOTES. —Values were taken from manufacturer specifications.

3. THE FUNCTIONALITIES OF SCEXAO

Key to the successful implementation of a high-performance coronagraph or interferometer for high-contrast imaging is a wavefront control system to correct for both atmospheric as well as instrumental aberrations. Wavefront control comprises two primary components: sensing and correction. The former is taken care of by a wavefront sensor, a device designed to map the spatial profile of the phase corrugations, while the latter by an adaptive element such as a DM. In this section, we describe how the elements of SCExAO function.

3.1. Wavefront Control

3.1.1. Wavefront Correction

The deformable mirror is the proverbial heart of any adaptive optics system. The 2k DM used in SCExAO (Fig. 4) is manufactured by Boston Micromachines Corporation based on MEMS technology. The DM is enclosed in a sealed chamber in order to control its environment. The critical parameter to control is the humidity level and is kept below 20% as advanced aging of MEMS actuators has been observed when high voltages (180 V is the maximum that can be applied), required to actuate individual elements, are applied in a moist environment (Shea et al. 2004). A desiccant is used to filter the compressed air of moisture. A circuit monitors both humidity and pressure and is set up to interlock the power to the DM based on these two metrics (see Appendix).

The chamber window is optimized for transmission across the entire operating range of SCExAO, 550–2500 nm (antireflection-coated, IR-fused silica). The DM's silicon membrane is gold coated and hosts 2000 actuators within the 18 mm diameter beam (there are others outside the pupil but they are not connected). The membrane is bonded to the ~250 μm square-shaped actuators, which are on a pitch of 400 μm. This means that there are 45 and 45.5 actuators across the 18 mm beam in the vertical and horizontal directions, respectively. The number of actuators across the DM defines the highest spatial frequency

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

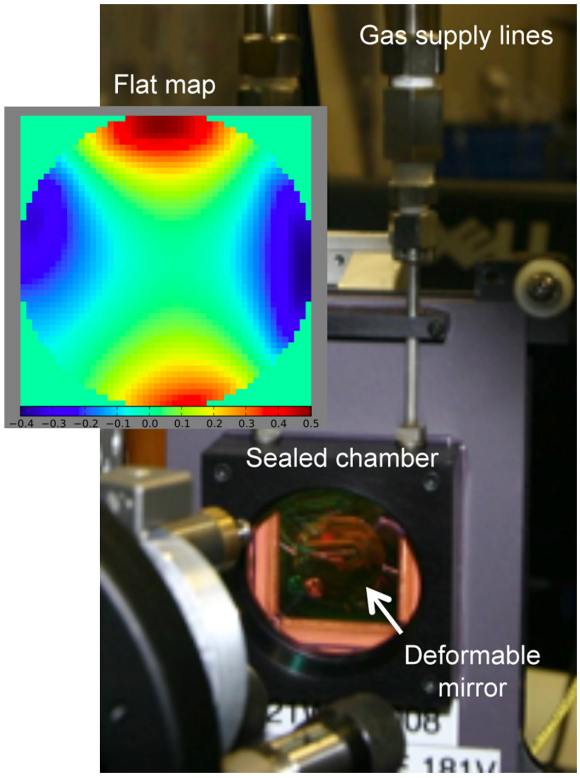


FIG. 4.—2k DM mounted in the SCExAO instrument. The gas supply lines to the sealed chamber can be seen along with the gold surface of the DM itself. Inset: The map applied to the DM in order to compensate for the DM surface figure. The map shows the magnitude of the modulation of the actuators in microns. The window is 50 actuators in diameter corresponding to the functioning region of the DM. The map required to flatten the DM surface is dominated by the Zernike which represents astigmatism. See the electronic edition of the *PASP* for a color version of this figure.

components which can be probed/controlled by the DM and hence the region of control in the focal plane. For our case, the 45 actuator pupil diameter means that the fastest modulation that can be Nyquist sampled by the DM consists of 22.5 cycles across the diameter. This means that spatial frequencies out to $22.5\lambda/D$ from the PSF can be addressed, which defines the radius of the control region in the focal plane (0.9″ in H-band, highlighted in Fig. 5). Figure 5 shows the surface of the DM, where the five dead actuators (actuators that cannot be modulated) are clearly visible. Fortunately, it was possible to position three of the actuators behind the secondary or outside the pupil and one is partially blocked by the spiders of the telescope, leaving only one dead segment in the illuminated pupil. Dead segments compromise the maximum contrast achievable post-coronagraph as they diffract light and hence the 1–2 illuminated for SCExAO is close to ideal. Note that most coronagraph Lyot stops currently installed in SCExAO do not mask out the actuators, but there are plans to include such masks in the future. In addition, the resolution of the DM can be qualitatively examined in Figure 5 where a bitmap image of the PI, Guyon's face, and the bat symbol have been imprinted in phase. The camera used to capture the image is not conjugated to the plane of the DM so that phase information gets recorded as amplitude information on the camera. The resolution of the images generated by the DM demonstrates the high level of sophistication MEMS technology has reached.

The surface of the unpowered DM is not flat. The inset to Figure 4 shows the voltage map that needs to be applied to obtain a flat DM in units of microns of stroke, known as a flatmap. The distortion of the surface of the DM is clearly dominated by astigmatism. Images of the PSF of SCExAO taken without and with the flat map applied are shown in the top left panels of Figure 5. The image without the flatmap is consistent with the presence of strong astigmatism. The PSF post flatmap is diffraction-limited. It is clear that the maximum stroke required for correction near the edges approaches 0.5 μm. This means that 25% of the 2 μm stroke of the actuators is used up to flatten the DM.

3.1.2. Wavefront Sensing

Wavefront correction within the SCExAO instrument comes in two stages: low spatial and temporal frequencies are partially corrected by AO188 prior to being injected into SCExAO where a final correction including higher order modes is implemented. In good seeing conditions, AO188 can offer 30–40% Strehl ratios in the H-band. The high-order wavefront correction, which is the focus of this section, is facilitated by a pyramid wavefront sensor (PyWFS). The PyWFS is chosen because of its large dynamic range and its sensitivity properties (Guyon 2005). In its standard implementation, a tip/tilt mirror modulates the location of the PSF in a circular trajectory which is centered on the apex of the pyramid. This implementation has been used to correct seeing-limited light to Strehl ratios as high as 90% in very good seeing on LBTAO (Esposito et al. 2011) and MagAO (Close et al. 2013). By removing modulation, however, the range over which the sensor responds linearly to aberrations is greatly reduced. Hence, to utilize a PyWFS without modulation, the wavefronts must be partially corrected by another sensor initially. With AO188 providing such a correction upstream, this implementation is possible with the PyWFS on SCExAO. SCExAO incorporates a piezo-driven mirror mount (shown in Fig. 3) to provide the modulation functionality, with the driver carefully synchronized to the frame rate of the camera. The implementation allows for continuous changes of the modulation radius (see Fig. 6). Such an architecture enables the possibility to start with a modulated PyWFS that has a larger range of linearity and slowly transition to a nonmodulated sensor for maximum sensitivity and the highest Strehl ratio, as wavefront errors are gradually reduced.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

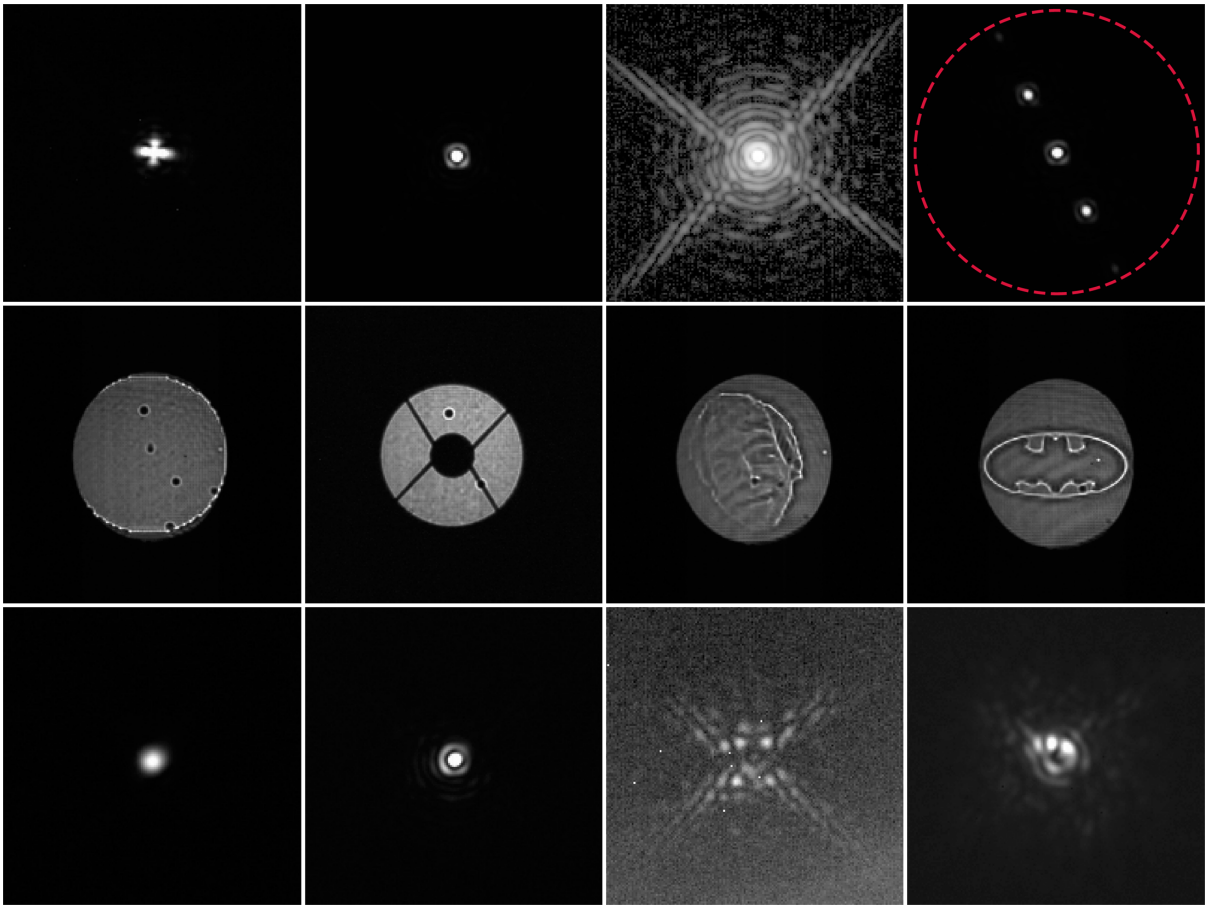


FIG. 5.—Top row: PSF of SCExAO at 1550 nm with the unpowered DM surface (left). Strong astigmatism is clearly visible. An image of the instrument PSF taken with the optimum flat map applied with linear (middle-left) and logarithmic (middle-right) scaling. The image is diffraction-limited and numerous Airy rings can be seen. Image with several artificial speckles applied (right). The *dashed ring* designates the edge of the control region of the DM corresponding to a radius of $22.5\lambda/D$ or 900 mas at 1550 nm. Second row: Pupil images showing the unmasked surface of the DM with five dead actuators (left), the internal spider mask in place masking several actuators (middle-left), an image of the PI Guyon (middle-right), and the bat symbol (right). These two images demonstrate the resolution of the DM. Bottom row: Image of the PSF taken in the laboratory with a laser at 1550 nm, after the conventional PIAA lenses (left), with the IPIAA lenses (middle-left) as well as with the achromatic focal plane mask (AFPM) ($1.9\lambda/D$ IWA) (*middle-right*). Right: An image with the PIAA, IPIAA, and AFPM is shown taken on-sky on the night of July 25, 2013, with the full H-band. See the electronic edition of the *PASP* for a color version of this figure.

The PyWFS has undergone laboratory and initial testing on-sky. In its initial format, it exploited a microlens array instead of a pyramid optic as it was easier to obtain (pyramid optics need to be custom made). However, it was determined that although small microlenses have good interlens quality, which keeps diffraction effects at a minimum, they have a limited field-of-view which limits their use with modulation. On the other hand, larger microlenses remove this limitation but the interlens quality is poor and results in strong diffraction. Hence, it was not possible from the two microlens arrays tested to simultaneously obtain a large field-of-view and low diffraction (there may indeed be microlenses on the market that can achieve both) (Clergeon 2014). For this reason, a dedicated pyramid optic was obtained. The pyramid optic presented here is a double pyramid, as shown in Figure 6, and is a replica of the one used on MagAO (Close et al. 2013; Esposito et al. 2010). The pyramid optic segments the PSF and generates four images of the pupil on the camera (see Fig. 6). The OCAM^{2}k from FirstLight imaging (see full specs in Table 1) is used as the detector and is capable of photon counting, with low read noise, high frame rate, and low latency (<1 frame) which enables the correction of high temporal frequencies and allows operation on faint

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

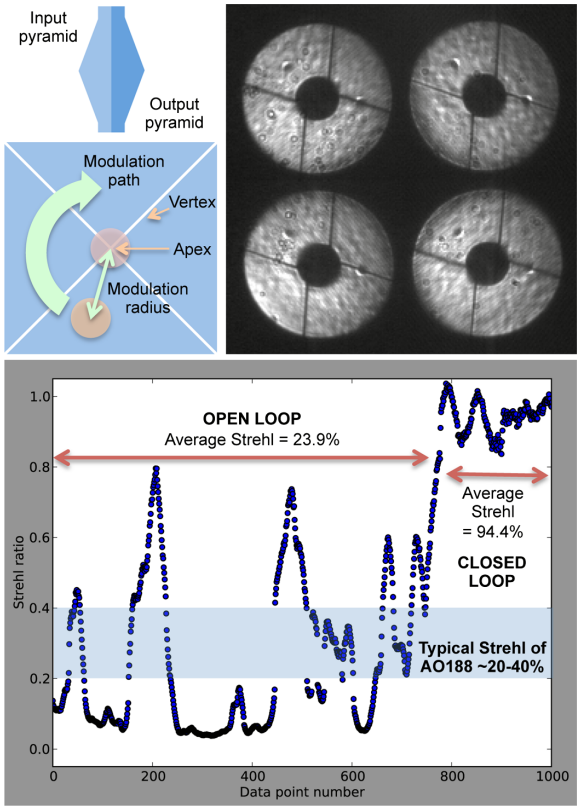


FIG. 6.—(Top left) A side view of the double pyramid. A very shallow angle (~1°) is required which is hard to produce so two steeper pyramids are used such that the cumulative effect of refraction between them is equivalent to a single shallower sloped pyramid optic (Esposito et al. 2010). Below this a front view of the pyramid is shown. The circles represent the PSF position without modulation (red spot) and with modulation (orange spot). The green arrow shows the path of the PSF across the front face of the pyramid optic when modulation is applied. (Top right) An image of the four pupils generated by the pyramid optic taken with modulation applied. (Bottom image) Strehl measured for the internal source with 300 nm RMS wavefront error applied to the turbulence simulator and the PyWFS loop open and then closed. NB, the Strehl ratio calculation procedure has a limited accuracy of about 5–7% when the Strehl ratio is >90%. Hence, values >1 are possible in this regime as shown in the above chart and should simply be interpreted as regions of high Strehl, with no emphasis put on the exact value. See the electronic edition of the *PASP* for a color version of this figure.

guide stars. To facilitate full-speed PyWFS loop operation (limited by camera frame rate), fast computations are required and for this a bank of GPUs is utilized. Details of the control loop architecture for the PyWFS are beyond the scope of this work. Here we focus on the recent performance of the loop.

In laboratory testing, the PyWFS successfully closed the loop with modulation (amplitude of $1.7\lambda/D$) on up to 830 modes. Our modal basis consisted of linearly independent modes obtained by singular value decomposition of an input basis consisting of five Zernike coefficients (tip, tilt, focus, astigmatisms) and the remainder Fourier modes (sine waves) up to a given spatial frequency of the DM. Nonorthogonality between input modes was addressed by rejecting (not controlling) low-eigenvalue modes. This basis set was chosen purely for convenience. This was achieved with the turbulence simulator generating 300 nm RMS wavefront error maps with a wind speed of 5 m/s and the low-order spatial frequencies scaled to 30% of the Kolmogorov power spectrum value to simulate upstream low-order wavefront correction. Light with wavelengths between 800 and 940 nm was used for the tests and the loop was run at 1 kHz. The Strehl of the images captured by the internal NIR camera was measured in the open- and closed-loop regime and the results depicted in Figure 6. The average Strehl in the open-loop regime was 23% which is consistent with the value predicted from Marechal's approximation (Hardy 1998) (22.7%) given 300 nm of wavefront error at 1550 nm. When the loop is closed on 830 modes, the Strehl improved to >90% on average confirming that given realistic post-AO correction, the PyWFS can indeed achieve extreme AO performance as required.

The nonmodulated mode was also tested with the turbulence simulator in the laboratory. However, due to the smaller linear range of the sensor without modulation the amplitude of wavefront errors was reduced. For the purposes of the test, it was set to 60 nm (windspeed 5 m/s and low-order Kolmogorov frequencies set to 30%). The loop was successfully closed at 1.5 kHz on the turbulence simulator on 1030 modes (which include five Zernikes). In this regime, it was clear that the speckles in the halo surrounding the PSF were held very static. The loop speed is constantly being improved by optimizing the code and removing delays. An operational speed as high as 3.5 kHz with minimal delay should be possible in the near future. Since the WFS detector operates in photon-counting mode, the loop will be able to run at full speed (3.5 kHz) even on faint targets (I-mag ~10), albeit at lower loop gain. Fainter targets still will require lower loop speeds. An optimal modal gain integrator will also be implemented soon.

Thus far the PyWFS has undergone some initial on-sky testing and has performed well on up to 130 modes (with modulation, 1.7 λ/D radius). However, as the speed of the loop was the same as AO188 (i.e., 1 kHz) and there were fewer modes than are corrected by AO188, a negligible improvement in Strehl was observed. Most of the gain came in the form of reduced tip/tilt jitter which was clearly visible in long integration time images. With further improvements in the AO loop code, the PyWFS should perform as demonstrated in the laboratory more recently, on-sky.

Noncommon path and chromatic low-order errors between the visible PyWFS and the IR coronagraphs are measured with the LLOWFS on the IR channel. The LLOWFS utilizes the light diffracted by the focal plane masks of the coronagraphs (discussed in detail in § 3.2), which is otherwise thrown away.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

A reflective Lyot stop is used to direct the diffracted light to the LOWFS camera (Singh et al. 2014). In this way, a reimaged PSF formed on the camera can be used to drive low-order, including tip/tilt, corrections by looking at the presence of asymmetries in the image. It has been tested thoroughly both in the laboratory and on-sky. Indeed, it was used on-sky in conjunction with the vector vortex coronagraph on Vega, on the nights of April 14 and 15, 2014, and produced residual RMS tip/tilt wavefront errors of $0.01\lambda/D$ (Singh et al. 2015).

In addition to the PyWFS and LLOWFS, we are testing other wavefront sensing techniques. One such technique is known as focal plane wavefront sensing which exploits eigenphase imaging techniques (Martinache 2013). The focal plane wavefront sensor relies on establishing a relationship between the phase of the wavefront in the pupil plane and the phase in the Fourier plane of the image. Although it has a limited range of linearity ($\sim\pi$ radians), which means that the wavefront must first be corrected to the 40% Strehl ratio level before this sensor can be utilized, it can boost the Strehl ratio to $>95\%$ in the H-band by correcting low-order modes. In addition, it operates just as effectively in the photon noise regime and is extremely powerful as noncommon path errors are eliminated. This wavefront sensor is currently under development and has been successfully tested on both the internal calibration source, in which case the aberrations due to the internal SCExAO optics were corrected as well as on-sky, where the static aberrations due to the telescope, AO188 optics, and SCExAO were all corrected. Some additional detail of this work can be found in Martinache et al. (2014a).

3.1.3. Coherent Speckle Modulation and Control

As a booster stage to the primary wavefront control loops, SCExAO makes use of coherent speckle modulation and control to both measure and attenuate residual starlight in the instrument's postcoronagraph focal plane. The 2k DM actuators are used to remove starlight from a predefined region, referred to as the dark hole (Malbet et al. 1995). Active modulation, induced by the 2k DM, creates coherent interferences between residual speckles of unknown complex amplitude and light added by modifying the DM's shape (this component's complex amplitude is known from a model of the DM response and the coronagraph optics). By iterating cycles of measurement and correction, starlight speckles that are sufficiently slow to last multiple cycles are removed from the dark hole area. This approach, developed and perfected in the last 20 yrs (Borde & Traub 2006; Codana & Kenworthy 2013; Give'on et al. 2007; Guyon et al. 2010; Malbet et al. 1995), is well suited to high-contrast imaging as it effectively targets slow speckles, which are the dominant source of confusion with exoplanets. It also allows coherent differential imaging (CDI), a powerful postprocessing diagnostic allowing true sources (incoherent with the central starlight) to be separated from residual starlight (Guyon et al. 2010). Compared to passive calibration techniques, such as angular differential imaging, CDI offers more flexibility, and achieves faster averaging of speckle noise. This is especially relevant at small angular separations, where ADI would require very long observation time to achieve the required speckle diversity. An example of a pair of speckles being generated by a periodic corrugation applied to the DM and used for starlight suppression is shown in the top right inset of Figure 5.

In SCExAO, coherent speckle control is implemented as discrete speckle nulling: The brightest speckles are identified in the dark hole region, and simultaneously modulated by the 2k DM, revealing their complex amplitudes. The 2k DM nominal shape is then updated to remove these speckles, and successive iterations of this loop gradually remove slow and static speckles. While discrete speckle nulling is not as efficient as more optimal global electric field inversion algorithms, it is far easier to implement and tune, and thus more robust for ground-based systems which have much larger wavefront errors than laboratory testbeds or space systems. This approach has been validated on SCExAO both in the laboratory (Martinache et al. 2012) and on-sky (Martinache et al. 2014b) and is a means of carving out a dark hole on one side of the PSF to boost contrast in that region. A recently taken image demonstrating the successful implementation of speckle nulling without a coronagraph on RX Boo is shown in the lower panel of Figure 7. The region where the nulling process was performed is outlined by the dashed white line and spans from 5–$22.5\lambda/D$. An image without the nulling applied is shown in the top panel of Figure 7 for comparison. This result was obtained on June 2, 2014, in favorable seeing conditions (seeing $<0.7''$). The nulling process reduced the average flux over the entire controlled area by 30% and by 58% in the region between 5–$12\lambda/D$, where the nulling was most effective. With better wavefront correction and the use of a coronagraph, the improvement in the contrast will grow.

The current limitations to achieving high-quality speckle nulling on-sky are wavefront correction, readout noise, and loop speed. As high-sensitivity cameras in the NIR are currently limited in regards to maximum frame rate (the Axiom cameras used are among the fastest commercially available at the time of writing), it is not possible for the active speckle nulling algorithm to pursue atmospherically induced speckles as they change from frame-to-frame. Hence, the current implementation of speckle nulling on SCExAO aims at removing the static and quasi-static speckles induced by diffraction from the secondary and spiders as well as optical aberrations. For this reason, it is important to have a high level of atmospheric wavefront correction on-sky so that the persistent speckles due to static and quasi-static aberrations can be easily identified. As the speckles are $\sim$1000$\times$ fainter than the PSF core, a magnitude limit for speckle nulling of 3–4 in the H-band is imposed by the current cameras used due to the high readout noise ($114\ e^-$). This places a severe limitation on potential targets of scientific interest. To alleviate these issues, SCExAO is acquiring a Microwave Kinetic Inductance Detector (MKID) which is a photon counting, energy

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

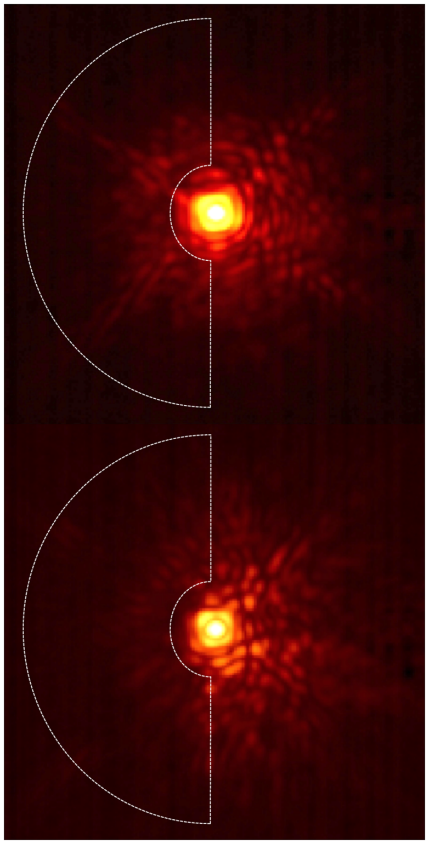


FIG. 7.—Top: RX Boo with no speckle nulling applied. Bottom: RX Boo with speckle nulling performed on the region enclosed by the *white dashed line*. Each image is a composite of 5000, 50 μs frames which have been shifted and added together. Each panel has a square root stretch applied to it and the maximum and minimum values are clipped for optimum viewing contrast. See the electronic edition of the *PASP* for a color version of this figure.

discriminating NIR array (Mazin et al. 2012). The MKID array will offer almost no readout noise or dark current and is capable of high frame rates (>1 kHz). This enables speckle nulling to be performed on fainter, more scientifically relevant targets, and for noncommon speckles due to chromatic dispersion in the atmosphere to be corrected for the first time, allowing for a significant improvement in detectability of faint companions. As the developmental time for the MKID array is several years, speckle control is being tested with a SAPHIRA (SELEX) array of avalanche photodiodes in the interim (Atkinson et al. 2014).

In contrast to speckle suppression, the addition of artificial speckles to the focal plane image can be utilized for precision astrometry when the on-axis starlight has been suppressed post-coronagraph. Further, by modulating the phase of the speckles during an exposure, they can be made incoherent with the speckles in the halo offering superior astrometric performance. In addition, by carefully calibrating the flux ratio between the PSF core and speckles, it is also possible to use these speckles for photometry and hence retrieving the contrast of companions as well (Martinache et al. 2014b). As opposed to diffractive grids utilized by other high-contrast imagers (Wang et al. 2014), the adaptive nature of the DM allows the speckle position, and brightness to be carefully tailored to each science case.

3.2. Coronagraphs

The advanced wavefront control techniques utilized on SCExAO build the foundation for high-contrast (10^{-5}–10^{-6}) imaging of faint companions with the onboard coronagraphs. The coronagraphs available in SCExAO are listed in Table 8. The performance of the coronagraphs used in SCExAO is limited by the level of wavefront control achieved. The PIAA/PIAACMC and the Vortex offer the lowest IWA and highest throughput but are more sensitive to wavefront error. On the other hand, the shaped pupil has a larger IWA and lower throughput but is less sensitive to residual wavefront error. Hence, the coronagraphs available are designed to span a large range of residual wavefront error and should be chosen accordingly.

The two key coronagraphs are the phase-induced amplitude apodization (PIAA) and the vector vortex types. PIAA refers to the act of remapping a flat-top pupil to a soft-edged pupil in order to remove the diffraction features associated with a hard-edged aperture (i.e., Airy rings) (Guyon 2003). These diffraction features make it difficult to suppress all of the light via a coronagraphic mask in the focal plane without blocking a close, faint companion. A combination of aspheric lenses are used to achieve the remapping in SCExAO and are referred to as PIAA lenses. SCExAO offers several types of remapping lenses. The first type is referred to as the conventional PIAA design and was presented in Lozi et al. (2009). Conventional PIAA lenses offer the most aggressive remapping, eliminating the secondary and converting the post-PIAA pupil into a prolate spheroid (a near-Gaussian which is finite in extent). An image depicting the remapping process between the two PIAA lenses in the laboratory is shown in Figure 8, while a radial profile of the apodization function is shown in Figure 9.

TABLE 8
DETAILS OF SCEXAO CORONAGRAPHS

Coronagraph type	Inner working angle (λ/D)	Waveband (s)
PIAA	1.5	y-K
PIAACMC	0.8	y-K
Vortex	2	H
MPIAA + vortex	1	H
MPIAA + 8-octant	2	H
4-Quadrant	2	H
Shaped pupil	3	y-K

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

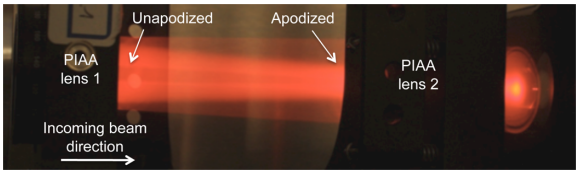


FIG. 8.—Picture of a visible beam being apodized by the conventional PIAA lenses in the laboratory (side view). The image is taken between the two lenses. The beam enters from the left of the image where the intensity across the beam is uniform except for the faint part in the middle of the beam (behind the secondary) and at the right hand side of the image, the beam is concentrated in the middle (i.e., apodized). See the electronic edition of the *PASP* for a color version of this figure.

To complete the softening of the edges of the beam, a binary mask is used which has a radially variant attenuation profile. Note that the binary mask is used to reduce the demand on the curvature changes across the aspheric surfaces of the PIAA lenses, and it is possible to eliminate it at the expense of increased manufacturing complexity of the PIAA lenses. As outlined in Guyon (2003), once the on-axis star has been suppressed with a focal plane mask, it is important to reformat the pupil to its original state in order to preserve the field-of-view. This can be done by using another set of PIAA lenses in reverse and the ensemble of lenses are referred to as the inverse PIAA lenses (IPIAA). The position of the PIAA and inverse PIAA lenses can be seen in Figure 3. Due to the low material dispersion of CaF_2, the conventional PIAA lens design used in SCExAO is achromatic across the NIR (y-K bands). However, an appropriate focal plane mask must be chosen to achieve this. If an ordinary opaque mask is used, then the size of the mask is wavelength dependent, and so is the IWA. To circumvent this issue, SCExAO uses focal plane masks that consist of a central cone surrounded by a ring of pits periodically positioned around the cone, made from a transmissive material on a substrate which refracts rather than reflects the on-axis star light (Newman et al. 2015). In this way, a variable focal plane mask which is achromatic across H-band can be achieved. In addition, since the light is strongly diffracted outwards by the focal plane masks, it can be redirected toward the LLOWFS via a reflective Lyot mask (Singh et al. 2014).

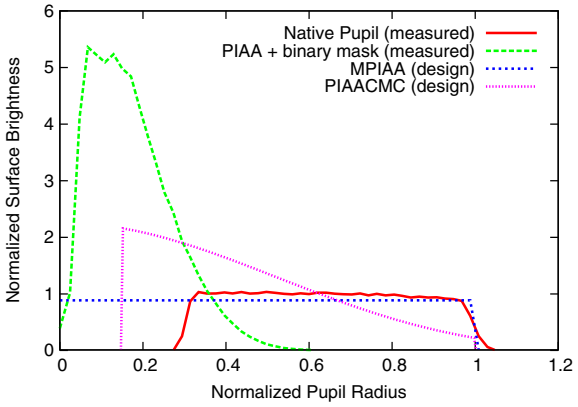


FIG. 9.—The figure shows a comparison of the radial apodization profiles of the various PIAA lenses as compared to the Subaru telescope pupil. The PIAA (*green line*) pulls the light inwards most aggressively, almost entirely removing the secondary and softening the edges of the beam. The PIAACMC (*pink line*) pulls some of the light inwards but does not entirely remove the presence of the secondary or soften the edges completely. The MPIAA (*blue line*) removes the presence of the secondary but does not soften the edges at all. See the electronic edition of the *PASP* for a color version of this figure.

Despite the fact that the conventional PIAA offered in SCExAO has an IWA of $1.5\lambda/D$, due to the aggressive remapping which causes an abrupt phase step in the central part of the beam post-PIAA, the contrast at $1.5\lambda/D$ is limited to 1×10^{-5} and is very sensitive to tip/tilt. To alleviate these issues, a modified version of the PIAA coronagraph can be used. It is referred to as the PIAA complex mask coronagraph (PIAACMC) and is outlined in greater detail in Guyon et al. (2010). The major difference is that the PIAA lenses used for the PIAACMC are less aggressive, which means the remapped pupil has soft edges but the secondary is still present as shown in Figure 9. The lenses themselves are in the same mounts as those for the PIAA so they can be replaced on the fly. The focal plane mask is now replaced with a partially transmissive, phase shifting mask which is manufactured via electron beam etching. The IWA of the coronagraph can be tuned by varying the opacity of the focal plane mask and in the limit when the mask is fully transmissive, the IWA is minimized at the expense of sensitivity to tip/tilt. Nonetheless, the PIAACMC offers a contrast of 1×10^{-6} at $1\lambda/D$, is less sensitive to tip/tilt than the PIAA, and is fully achromatic from y-K band. It is scheduled to be installed and commissioned in the near future.

A third and final type of PIAA is used to remap the pupil into a flat top without a central obstruction for an 8-octant coronagraph which is discussed in the following section (Oshiyama et al. 2014). The lenses are referred to as MPIAA lenses and reside in the same mounts as the other two. Despite the remapping, these optics are not apodizers as the pupil retains its hard edge post remapping. A comparison of the various apodization schemes is shown in Figure 9.

Other coronagraphs include the vortex (Mawet et al. 2009, 2010), 4-quadrant (Rouan et al. 2000), 8-octant (Murakami et al. 2010), and shaped pupil (Carlotti et al. 2012) versions. The vortex, 4-quadrant, and 8-octant coronagraphs are phase-mask coronagraphs as opposed to occulting coronagraphs and consist of two primary elements: a focal plane and Lyot stop mask. All focal plane masks are situated in a wheel in the focal plane while the Lyot stop masks are located in the Lyot mask wheel and the positions of both are shown in Figure 3.

The vortex coronagraph in SCExAO uses a H-band optimized, topographic charge two focal plane mask. This mask is constructed from a birefringent liquid crystal polymer material,

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

i.e., a waveplate where the optical axis (fast axis) orientation is spatially dependent and in this case a function of the azimuthal coordinate (Mawet et al. 2009). Although an IWA of $0.9\lambda/D$ is achievable with an unobstructed pupil with nonmanufacturing defects, it is limited to $2.0\lambda/D$ with the pupil geometry at the Subaru Telescope. However, the vector vortex on SCExAO could be used in conjunction with the MPIAA lenses to circumvent this issue and regain the inherent IWA. The vector vortex mask is more achromatic than a scalar mask and hence can operate across the full H-band (Mawet et al. 2009). As the nulling process is based on interference of light from different parts of the mask, best performance is achieved with higher Strehl ratios and stable centering of the PSF on the mask (5 mas tip/tilt error or below in this case).

The 8-octant coronagraph focal plane mask employed on SCExAO is based on photonic crystal technology (Murakami et al. 2010). It consists of eight triangular segments that comprise half-wave plates where the optical axes of a given segment are always orthogonal to its two nearest neighbors. This creates a π phase shift between adjacent segments for the transmitted beam which destructively interferes in the reimaged focal plane to null out the on-axis star. The 8-octant itself is not achromatic but broadband operation can be realized by placing a polarizer and analyzer before and after the mask respectively (Murakami et al. 2008) (note this is also true for the vortex coronagraph). This coronagraph exploits the pupil-reformatting MPIAA lenses described above to achieve an IWA of $\sim 2\lambda/D$ and offers very high-contrasts at these angular scales. Similar to the vector vortex it is also sensitive to tip/tilt and hence active control is the preferred mode of operation.

The 4-quadrant focal plane mask is a scalar mask which consists of segments that phase shift the light by π with respect to the neighboring segments. Although a perfectly manufactured 4-quadrant mask could offer an IWA of as low as $1\lambda/D$ if used in conjunction with the MPIAA lenses mentioned above, the mask in SCExAO has manufacturing defects and so cannot achieve such performance. The 4-quadrant in SCExAO is a prototype which serves its purpose for internal testing only.

The Lyot stop masks for the vector vortex, 4-quadrant, and 8-octant coronagraphs are designed to reflect rejected light toward the LOWFS camera for fine tip/tilt guiding which is discussed in the subsequent section. The vector vortex and 4-quadrant Lyot stop masks consist of a replica mask to the Subaru Telescope pupil geometry with slight modifications. Both masks have a slightly oversized secondary and spiders for better rejection; however, the 4-quadrant has a square secondary instead of a circular one. On the other hand, as the secondary is eliminated thanks to the MPIAA lenses, then the Lyot stop for the 8-octant is simply a slightly undersized circular aperture.

Finally, shaped pupil coronagraphs can also be tested on SCExAO. These coronagraphs are located in the pupil plane mask wheel and any focal or Lyot plane masks required are placed in the appropriate wheel. For further details, please see Carlotti et al. (2012).

3.3. Lucky Fourier Imaging

An important element of all adaptive optics systems is a real-time PSF monitoring camera. This is depicted as the Lucky imaging camera in Figure 3, the specifications of which are listed in Table 1. Currently, a narrowband of light ($\sim$30–50 nm) is steered toward this camera from the pupil plane masks of VAMPIRES and the PSF imaged. The camera runs at a high frame rate, subframed, and collects images rapidly which are primarily used for monitoring the PSF. The frames can subsequently also be used for traditional lucky imaging. However, a more advanced version of this technique named Lucky Fourier Imaging is commonly utilized (Garrel et al. 2012). The technique relies on looking for the strongest Fourier components of each image, and then synthesizing a single image with the extracted Fourier information. In this way, diffraction-limited images at 680 nm of targets like Vega (bottom image in Fig. 10) and Betelgeuse have been synthesized in 2″ seeing (Garrel 2012). This is clearly an extremely powerful tool and we propose to advance this imaging capability by adding multiple spectral channels.

3.4. Fiber Injection Unit

In addition to direct imaging, long baseline interferometry and high-precision radial velocity both stand to gain significantly from a stable and 90% Strehl PSF on an 8-m class telescope. For example, long baseline interferometers like the Optical Hawaiian Array for Nanoradian Astronomy (OHANA) combine beams from multiple telescopes once it has been transported to the combination room via single-mode optical fibers (Woillez et al. 2004). However, coupling efficiently into single-mode fibers is no mean feat, but with access to a stable PSF with 90% Strehl, it can be achieved. Indeed, this is already being exploited for the purposes of nulling interferometry on P1640 (Serabyn et al. 2010). Once the light has been coupled into a single-mode fiber, it could be used as an alternative feed for a conventional multimode fiber-fed spectrograph. The nontemporally and spatially varying PSF provided by a single-mode fiber can be used in precision radial velocity measurements to eliminate modal-noise, a limiting factor in achieving high precisions. For these reasons, we are developing a single-mode injection unit on the SCExAO platform.

To inject light into the fiber, it is tapped off with a retractable dichroic on its way to focus after OAP2 (see Fig. 3). A dichroic which reflects y, J, and H-short bands is currently used for this. An achromatic lens ($f = 10$ mm) is used to adjust the $f/\#$ of the beam before it is injected into the fiber which sits atop a stage. The 5-axis stage allows for XYZ translation via precise stepper motor actuators and course alignment of tip/tilt. The stage can be scanned through focus to maximize coupling into the fiber.

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

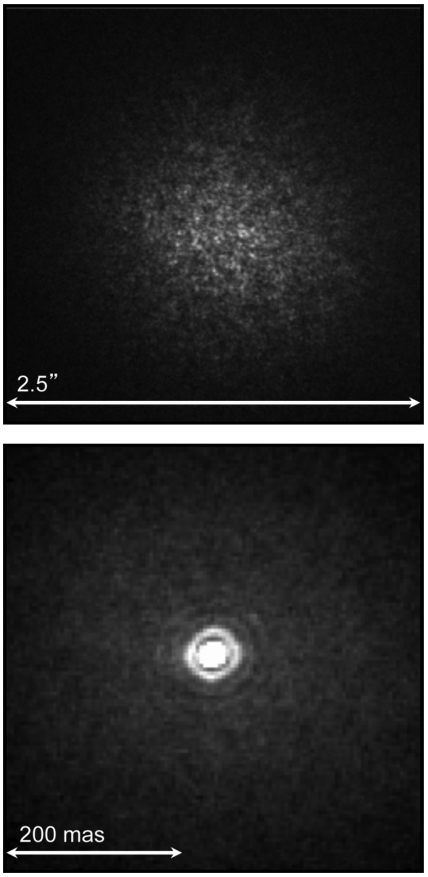


FIG. 10.—Top image: Vega in 2″ seeing at 680 nm. Bottom image: Synthesized image of Vega at 680 nm in the presence 2″ seeing. Image was reconstructed from a 1% selection of Fourier components across the 10^4 frames collected. A diffraction-limited PSF with a FWHM of 17 mas is obtained post-reconstruction (note: the scale of the bottom image differs from the top one). The data was acquired on 2012 February 5 and 6.

A further advantage of implementing such a module on SCExAO is that we can use the conventional PIAA lenses to more closely match the intensity distribution of the collection fiber and hence boost the coupling to a theoretical value of 100%. To make this useful on-sky, this injection system relies on the PyWFS delivering a high Strehl beam. The fine tip/tilt control is provided by the LLOWFS by using the transmitted H-band light. This unit is currently operational and may be utilized by instruments such as the high-precision spectrograph IRD for the survey it will undertake (Tamura et al. 2012). In addition, by developing such a unit, it becomes possible to exploit numerous other photonics technologies on-sky (Cvetojevic et al. 2012; Marien et al. 2012). The injection unit is described in more detail in Jovanovic et al. (2014b).

TABLE 9

COMMISSIONING STATUS OF SCEXAO MODES OF OPERATION AND MODULES

Module/mode	Commissioning status (completion date)
Wavefront control	
LLOWFS	Complete
SN	Complete
PyWFS	Partially complete (2015 fall)
PyWFS + LLOWFS	Partially complete (2015 fall)
SN + PyWFS + LLOWFS	Incomplete (late 2015)
Coronagraphs	
PIAA	Complete
Vortex	Complete
4-Quadrant	Complete
Shaped pupil	Complete
PIAACMC	Incomplete (2015 fall)
MPIAA + Vortex	Incomplete (late 2015)
MPIAA + 8-octant	Incomplete (2015 fall)
Visible imagers	
VAMPIRES	Complete
Lucky imaging	Partially complete (late 2015)
FIRST	Incomplete (late 2015)
NIR fiber injection	Complete
NIR science imagers	
CHARIS	Expected delivery early 2016 (mid-2016)
MKID	Expected delivery late 2016 (2017)

NOTES. —+ signifies that these modes/modules are operating in conjunction. SN-speckle nulling, PIAA will be replaced by the PIAACMC.

4. COMMISSIONING STATUS AND FUTURE EXTENSIONS

SCExAO is clearly a complex instrument with modules at various stages of commissioning. Table 9 summarizes the commissioning status/plan of the various modes of operation and modules of SCExAO.

Table 9 shows that the integral field spectrograph known as CHARIS will replace the HiCIAO imager from mid 2016. This instrument segments the focal plane with an array of microlenses, before dispersing each PSF and then reimaging onto a detector (see Table 1) (Peters et al. 2012). This allows for spatially resolved spectral information, albeit at low resolving powers. Such an instrument has three key advantages. First, background stars in a given image can quickly be identified. Second, owing to the fact that the instrument operates over a very broadband (J-K bands), the presence of a planet can be inferred by detecting the fixed speckle within a spectral data cube. Finally, low-resolution spectra of gas giants can be taken, enabling the atmospheres of these planets to be constrained and better understood (Brandt et al. 2014; McElwain et al. 2012). Indeed, presently operating integral field spectrographs such as OSIRIS at the Keck Telescope, the units in P1640 and GPI have been used to characterize the atmospheres of known planetary systems like HR8799 (Barman et al. 2011; Ingraham et al. 2014; Oppenheimer et al. 2013).

Although not included in the table, it has been proposed to outfit the IR arm of SCExAO with a polarimetric mode of

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

operation to study scattered dust in circumstellar disks. This mode, known as polarization differential imaging (PDI), has been hugely successful for the HiCIAO imager (Grady et al. 2013), and we aim to preserve this capability while offering a superior IWA. This IR polarimetric mode will complement one of the VAMPIRES offered in the visible, albeit on different spatial scales.

5. SUMMARY

The SCExAO instrument is a versatile high-contrast imaging platform which hosts advanced wavefront control systems, IR coronagraphs, and visible interferometers that are ideal for imaging at $<3\lambda/D$ (solar-system scales). The extreme adaptive optics system delivers the necessary wavefront correction to be able to push detection limits for ground-based observations at small angular separations and interferometer precisions. Such instruments will be critical to understanding the inner structure of circumstellar disks and planetary systems and how they form. In addition, they will provide the appropriate avenue to collecting spectra from planetary candidates and determining their physical properties for the first time. Further, the SCExAO platform is an ideal testbed for demonstrating and prototyping technologies for future ELTs and space-missions. SCExAO is the only high-contrast imager of its kind and will be uniquely positioned to contribute to exoplanetary science.

We are grateful to B. Elms for his contributions to the fabrication of parts for the SCExAO rebuild. The SCExAO team thanks the Subaru directorate for funding various grants to realize and develop the instrument.

APPENDIX

A.1. Off-Axis Parabolic Mirrors

All OAPs in the SCExAO instrument were manufactured via diamond turning of aluminum and overcoating with gold, are 50 mm in diameter, and were designed to work at a nominal off-axis angle of 17°. OAPs 1, 3, 4, and 5 have a $f = 255$ mm while OAP2 has a $f = 519$ mm. Although there is no data on the wavefront error of these optics, each optic was initially used to form an image in the visible and it was determined from this that the RMS wavefront error was $<\lambda/20$ at 630 nm over a 20 mm beam size corresponding to that which is used in SCExAO.

A.2. Internal NIR Camera Lenses

The focusing lenses for the science camera include a $f = 150$ and 50 mm converging, achromatic doublet which are AR-coated for the NIR region (1–1.65 μm). The distance between the lenses is set to be just larger than the sum of their focal lengths so that a slow beam is formed ($f/65$).

A.3. Deformable Mirror Environmental Controls

An interlock system which monitors the environmental conditions in the DM chamber (pressure, humidity) was put in place to prolong the life of the DM as they are known to age rapidly in high-humidity environments (Shea et al. 2004). A low-pressure regulator (Fairchild-M4100) was used to offer fine control of the injected dry air pressure to the DM chamber at the <1 psi level (with respect to ambient). The pressure is set to 0.4 psi above ambient when operating and monitored by a precise pressure sensor (FESTO–SDE1). Such low-pressure differentials are used so that the chamber does not deform significantly and hence induce any extra errors to the wavefront. A 1 psi pressure differential relief valve is used as a hard limit in case of over pressure in the circuit. The humidity in the circuit is measured with a moisture probe (Edgetech–HT120). The alarms for both the pressure and humidity sensors are used to control the power supplied to the DM electronics. When the humidity is below 15% and the pressure between 0.2 and 0.8 psi, the DM electronics will be powered and the actuators can be driven. However, if the humidity rises above 15% and/or the pressure goes above 0.8 psi or below 0.2 psi, then the alarms on the sensors will trigger a relay switch, to which the DM electronics are connected, to trip and cut the power to the DM. As a final level of reassurance, a low flow rate (250 mL/ min) flow controller is connected to the end of the line to insure that there is a very slow flow over the DM membrane and no turbulence in the chamber. A rapid flow could tear the thin silicon membrane (3 μm) and/or cause turbulence in the chamber which would be equivalent to seeing. The window to the chamber is 50 mm in diameter, made from a 12 mm thick piece of IR-fused silica which is AR-coated across the operating range.

REFERENCES

Artigau, E., Biller, B. A., Wahhaj, Z., Hartung, M., Hayward, T. L., Close, L. M., Chun, M. R., Li, M. C., et al. 2008, Proc. SPIE, 70141 Z
Atkinson, D., Hall, D., Baranec, C., Baker, I., Jacobson, S., & Riddle, R. 2014, Proc. SPIE, 915419 A
Barman, T. S., Macintosh, B., Konopacky, Q. M., & Marois, C. 2011, ApJ, 733, 65 B
Beuzit, J.-L., Feldt, M., Dohlen, K., Mouillet, D., Pugeta, P., Wildi, F., Abee, L., Antichif, J., et al. 2008, Proc. SPIE, 701418
Borde, P. J., & Traub, W. A. 2006, ApJ, 638, 488 B

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

Borucki, B., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., et al. 2010, Science, 327, 977
Brandt, T., McElwain, M., Janson, M., Knapp, G., Mede, K., Limbach, M., Groff, T., Burrows, A., et al. 2014, Proc. SPIE, 9148-49
Buton, C., et al. 2012, A&A, 549A, 5 B
Carlotti, A., Kasdin, N. J., Martinache, F., Vanderbei, R. J., Young, E. J., Che, G., Groff, T. D., & Guyon, O. 2012, Proc. SPIE, 84463 C
Charbonneau, D., Brown, T. M., Noyes, R. W., & Gillilna, R. L. 2001, ApJ, 568, 377
Clergeon, C. 2014, Ph.D. dissertation
Close, L. M., Males, J. R., Morzinski, K., Kopon, D., Follette, K., Rodigas, T. J., Hinz, P., Wu, Y.-L., et al. 2013, ApJ, 774, 94 C
Codana, J. L., & Kenworthy, M. 2013, ApJ, 767, 100 C
Cvetojevic, N., Jovanovic, N., Betters, C., Lawrence, J. S., Ellis, S., Robertson, G., & Bland-Hawthorn, J. 2012, A&A, 544, L 1
Dekany, R., Roberts, J., Burruss, R., Antonin, T., Baranec, C., Guiwits, S., Hale, D., Angione, J., et al. 2013, ApJ, 776, 130 D
Esposito, S., Riccardi, A., Fini, L., Puglisi, A. T., Pinna, E., Xompero, M., Briguglio, R., Quirs-Pacheco, F., et al. 2010, Proc. SPIE, 7736-09 F
Esposito, S., Riccardi, A., Pinna, E., Puglisi, A., Quirs-Pacheco, F., Arcidiacono, C., Xompero, M., Briguglio, R., et al. 2011, Proc. SPIE, 8149-02 E
Garrel, V., Guyon, O., & Baudoz, P. 2012, PASP, 124, 861
Garrel, V. 2012, PhD dissertation
Give'on, A., Kern, B., Shaklan, S., Moody, D. C., & Pueyo, L. 2007, Proc. SPIE, 66910 A
Gozdziewski, K., & Migaszewski, C. 2014, MNRAS, 440, 3140
Grady, C. A., Muto, T., Hashimoto, J., Fukagawa, M., Currie, T., Biller, B., Thalmann, C., Sitko, M. L., et al. 2013, ApJ, 762, 48
Guyon, O. 2003, A&A, 404, 379
———. 2005, ApJ, 629, 592
Guyon, O., Martinache, F., Belikov, R., & Soummer, R. 2010, ApJS, 190, 220
Guyon, O., Pluzhnik, E., Martinache, F., Totems, J., Tanaka, S., Matsuo, T., Blain, C., & Belikov, R. 2010, PASP, 122, 71 G
Han, E., Wang, S., Wright, J. T., Feng, Y. K., Zhao, M., Fakhouri, O., Brown, J. I., & Hancock, C. 2014, PASP, 126, 827
Hardy, J. W. 1998, Adaptive Optics for Astronomical Telescopes (New York: Oxford University Press)
Helmbrecht, M. A., Min, H., Carl, K., & Marc, B. 2011, Proc. SPIE, 793108
Huby, E., Duchene, G., Marchis, F., Lacour, S., Perrin, G., Kotani, T., Choquet, E., Gates, E. L., et al. 2013, A&A, 560, A 113
Huby, E., Perrin, G., Marchis, F., Lacour, S., Kotani, T., Duchene, G., Choquet, E., Gates, E. L., et al. 2012, A&A, 541A, A 55
Ingraham, P., et al. 2014, ApJ, 794L, 15 I
Jovanovic, N., Guyon, O., Martinache, F., Clergeon, C., Singh, G., Kudo, T., Newman, K., Kuhn, J., et al. 2014a, Proc. SPIE, 9147-1 Q
Jovanovic, N., Guyon, O., Martinache, F., Schwab, C., & Cvetojevic, N. 2014b, Proc. SPIE, 9147, 9147-287
Kraus, A., & Ireland, M. J. 2012, ApJ, 745, 1
Lafreniere, D., Marois, C., Doyon, R., & Barman, T. 2009, ApJ, 694L, 148 L
Lagrange, A.-M., Gratadour, D., Chauvin, G., Fusco, T., Ehrenreich, D., Mouillet, D., Rousset, G., Rouan, D., et al. 2009, A&A, 493, 21 L
Lenzen, R., Hartung, M., Brandner, W., Finger, G., Hubin, N., Lacomber, F., Lagrange, A.-M., Lehnert, M., et al. 2003, Proc. SPIE, 4841-944 L
Leon-Saval, S. G., Argyros, A., & Bland-Hawthorn, J. 2013, Nanophotonics, 2, 429
Lord, S. D. 1992, NASA Technical Memorandum 103957
Lozi, J., Marinache, F., & Guyon, O. 2009, PASP, 121, 1232
Macintosh, B., et al. 2014, Proc. Natl. Acad. Sci., 111, 12661
Malbet, F., Yu, J. W., & Shao, M. 1995, PASP, 107, 386 M
Mariën, G., Jovanović, N., Cvetojević, N., Williams, R., Haynes, R., Lawrence, J., Parker, Q., & Withford, M. J. 2012, MNRAS, 421, 3641
Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., Lafrenière, D., & Doyon, R. 2008, Science, 322, 1348 M
Martinache, F. 2013, PASP, 125, 422 M
Martinache, F., Guyon, O., Clergeon, C., & Blain, C. 2012, PASP, 124, 1288
Martinache, F., Guyon, O., Jovanovic, N., Clergeon, C., Singh, G., Kudo, T., Currie, T., Thalmann, C., et al. 2014a, Proc. SPIE, 9148, 914870
Martinache, F., Guyon, O., Jovanovic, N., Clergeon, C., Singh, G., Kudo, T., Currie, T., Thalmann, C., et al. 2014b, PASP, 126, 565
Mawet, D., Serabyn, E., Liewer, K., Burruss, R., Hickey, J., & Shemo, D. 2010, ApJ, 709, 53
Mawet, D., Serabyn, E., Liewer, K., Hanot, Ch., McEldowney, S., Shemo, D., & O'Brien, N. 2009, Opt. Exp., 17, 1902
Mayor, M., & Queloz, D. 1995, Nature, 378, 355
Mazin, B. A., Bumble, B., Meeker, S. R., O'Brien, K., McHugh, S., & Langman, E. 2012, Opt. Exp., 20, 1503
McElwain, M. W., Brandt, T. D., Janson, M., Knapp, G. R., Peters, M. A., Burrows, A. S., Carlotti, A., Carr, M. A., et al. 2012, Proc. SPIE, 84469 C
Minowa, Y., Hayano, Y., Oya, S., Watanabe, M., Hattori, M., Guyon, O., Egner, S., Saito, Y., et al. 2010, Proc. SPIE, 77363 N
Murakami, N., Nishikawa, J., Yokochi, K., Tamura, M., Baba, N., & Abe, L. 2010, ApJ, 714, 772
Murakami, N., Uemura, R., Baba, N., Nishikawa, J., Tamura, M., Hashimoto, N., & Abe, L. 2008, PASP, 120, 1112
Muto, T., Grady, C. A., Hashimoto, J., Fukagawa, M., Hornbeck, J. B., Sitko, M., Russell, R., Werren, C., et al. 2012, ApJ Lett., 748, L 22
Newman, K., Guyon, O., Balasubramanian, K., Belikov, R., Jovanovic, N., Martinache, F., & Wilson, D. 2015, PASP, 127, 437
Norris, B., Schworer, G., Tuthill, P., Jovanovic, N., Guyon, O., Stewart, P., & Martinache, F. 2015, MNRAS, 447, 2894 N
Norris, B., Tuthill, P. G., Ireland, M. J., Lacour, S., Zijlstra, A. A., Lykou, F., Evans, T. M., Stewart, P., et al. 2012a, Proc. SPIE, 844503 N
Norris, B., Tuthill, P. G., Ireland, M. J., Lacour, S., Zijlstra, A. A., Lykou, F., Evans, T. M., Stewart, P., et al. 2012b, Nature, 484, 220
Oppenheimer, B. R., Baranec, C., Beichman, C., Brenner, D., Burruss, R., Cady, E., Crepp, J. R., Dekany, R., et al. 2013, ApJ, 768, 24 O
Oshiyama, F., Murakami, N., Guyon, O., Martinache, F., Baba, N., Matsuo, T., Nishikawa, J., & Tamura, M. 2014, PASP, 126, 270
Pepe, F., Ehrenreich, D., & Meyer, M. R. 2014, Nature, 513, 358
Perrin, G., Lacour, S., Woillez, J., & Thiebaut, E. 2006, MNRAS, 373, 747 P

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

Peters, M. A., Groff, T., Kasdin, N. J., McElwain, M. W., Galvin, M., Carr, M. A., Lupton, R., Gunn, J. E., et al.2012, Proc. SPIE, 8446E, 7UP
Poyneer, L. A., De Rosa, R. J., Macintosh, B., Palmer, D. W., Perrin, M. D., Sadakuni, N., Savransky, D., Bauman, B., et al. 2014, Proc. SPIE, 9148E, 0KP
Rouan, D., Riaud, P., Boccaletti, A., Clenet, Y., & Labeyrie, A. 2000, PASP, 112, 1479
Serabyn, E., Mawet, D., & Burruss, R. 2010a, Nature, 464, 1018
Serabyn, E., Mennesson, B., Martin, S., Liewer, K., Mawet, D., Hanot, C., Loya, F., Colavita, M. M., et al. 2010b, Proc. SPIE, 77341 E
Shea, H. R., Gasparyan, A., Chan, H. B., Arney, S., Frahm, R. E., Lopez, D., Jin, S., & McConnell, P. 2004, Trans. on Dev. & Mat. Rel., 4, 198
Singh, G., Martinache, F., Baudoz, P., Guyon, O., Matsuo, T., Jovanovic, N., & Clergeon, C. 2014, PASP, 126, 586
Singh, G., Lozi, J., Guyon, O., Jovanovic, N., Baudoz, P., Martinache, F., Kudo, T., Serabyn, E., et al. 2015, PASP, 127, 857
Tamura, M. 2009, Proc. AIPC, 1158, 11 T
Tamura, M., Suto, H., Nishikawa, J., Kotani, T., Sato, B., Aoki, W., Usuda, T., Kurokawa, T., et al. 2012, Proc. SPIE, 84461 T
Tuthill, P., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555
Tuthill, P., Lacour, S., Amico, P., Ireland, M., Norris, B., Stewart, P., Evans, T., Kraus, A., et al. 2010, Proc. SPIE, 7735, 1O–T
Wang, J. J., Rajan, A., Graham, J. R., Savransky, D., Ingraham, P. J., Ward-Duong, K., Patience, J., De Rosa, R. J., et al. 2014, Proc. SPIE, 9147, 15
Woillez, J. M., Perrin, G., Guerin, J., Lai, O., Reynaud, F., Wizinowich, P. L., Neyman, C. R., Le Mignant, D., et al. 2004, Proc. SPIE, 5491, 1425

This content downloaded from 192.54.176.192 on Wed, 20 Jan 2016 14:15:21 UTC
All use subject to JSTOR Terms and Conditions

Aberrations de bas ordre dans le coronographe

S'il est nécessaire, pour augmenter le potentiel de détection d'un système d'imagerie haut contraste, d'employer un coronographe avec un inner working angle (IWA) aggressif (entre un et deux λ/D dans le cas de SCExAO), ce même coronographe devient particulièrement sensible aux erreurs de pointage et aux aberrations de bas ordre comme des erreurs de focus ou d'astigmatisme. Cela veut cependant aussi dire qu'un tel coronographe peut également être utilisé comme un outil très sensible de mesure de ces aberrations de bas ordre. La question est de savoir comment exploiter l'information contenue dans une des voies de sorties possibles de la lumière de ce coronographe.

Alors que l'image coronographique est difficilement exploitable sans introduire d'altération qui auront un impact sur la performance du système, la lumière rejetée par le coronographe peut assez facilement être utilisée à des fins métrologiques. Ce n'est pas une idée nouvelle : elle a été mise en oeuvre dans le cadre du projet Lyot[42] pour stabiliser l'alignement de la source sur le coronographe. L'unité de calibration de l'instrument GPI utilise également la lumière occultée pour en déduire les aberrations de bas ordre, après reformation d'une pupille (ne conservant qu'un champ de quelques λ/D) dans un analyseur de type Shack-Hartmann[43] de bas-ordre. La stabilité d'un tel contrôle de pointage atteint environ 2 mas dans des conditions normales d'observation depuis Gemini Sud.

42. A. P. Digby, S. Hinkley, B. R. Oppenheimer, A. Sivaramakrishnan, J. P. Lloyd, M. D. Perrin, L. C. Roberts, Jr., R. Soummer, D. Brenner, R. B. Makidon, M. Shara, J. Kuhn, J. Graham, P. Kalas, and L. Newburgh. The Challenges of Coronagraphic Astrometry. **ApJ**, 650:484–496, Oct. 2006

43. J. K. Wallace, R. S. Burruss, R. D. Bartos, T. Q. Trinh, L. A. Pueyo, S. F. Fregoso, J. R. Angione, and J. C. Shelton. The Gemini Planet Imager calibration wavefront sensor instrument. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2010

La sensibilité à la mesure du pointage est cependant maximale lorsqu'on laisse interférer la lumière en provenance des bords opposés de la pupille, ce qui se produit naturellement dans le plan focal. La solution offerte par un Shack-Hartmann, découpe la pupille en sous-pupilles dont la lumière ne peut plus interférer. Bien que robuste, cette solution est donc loin d'être optimisée pour la mesure du pointage. L'alternative décrite ici se place dans le plan focal coronographique : c'est le concept de "coronagraphic low order wavefront sensor" (CLOWFS)[44]. Cette implémentation utilise à cet effet, une caméra qui réimage la lumière réfléchie par un masque occulteur modifié de façon à augmenter la sensibilité de la mesure. La modification du masque occulteur est présentée en figure 32 : il est rendu réfléchissant dans une région annulaire comprise entre deux rayons critiques r_1 et r_2 dont les dimensions exactes dépendent du type de coronographe et de l'IWA pour lequel il est spécifié. L'intérêt de cette région réfléchissante annulaire plutôt que simplement circulaire est discutée dans l'article de 2009 : disons pour aller vite, que supprimer la contribution de la région centrale du masque retire une quantité importante de lumière qui apporte très peu d'information. La partie centrale du masque retire du signal du CLOWFS un offset qui se traduit par une meilleure sensibilité.

44. O. Guyon, T. Matsuo, and R. Angel. Coronagraphic Low-Order Wave-Front Sensor: Principle and Application to a Phase-Induced Amplitude Coronagraph. **ApJ**, 693:75–84, Mar. 2009

J'ai en parallèle de mon travail sur SCExAO, contribué à

l'implémentation et à l'exploitation d'un tel masque sur une expérience de haut contraste sur le "High Contrast Imaging Testbed" (HCIT) hébergé au JPL : pour ce système particulier, utilisé avec un PIAA, les deux rayons sont respectivement de 0.8 et 1.6 λ/D.

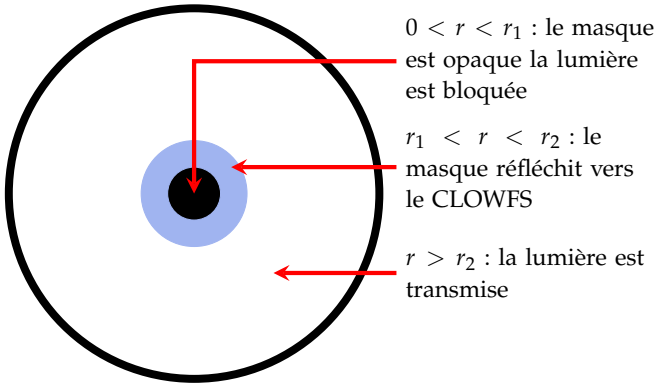


FIGURE 32: Masque focal occulteur modifié. Constitué de trois zones : une région opaque au centre, une région annulaire réfléchissante et une grande région transmissive. La lumière réfléchie par la section annulaire peut être utilisé dans une mesure des aberrations de bas ordre : c'est le concept de CLOWFS.

La lumière réfléchie par cette partie annulaire du plan focal est réimagée par une simple lentille sur une caméra délibérément défocalisée, permettant de lever une dégénérescence sur le signe des aberrations d'ordre pair, comme en diversité de phase. Avec un tel design, de très faibles variations dans l'alignement ou des erreurs de bas ordre se traduisent dans l'image réfléchie par des variations macroscopiques de la structure de l'image produite par cette caméra. On peut enregistrer la réponse de cet analyseur bas ordre à des modulations contrôlées de pointage et utiliser cette réponse pour fermer une boucle de contrôle qui stabilise l'image coronographique sur une position de référence, choisie au préalable.

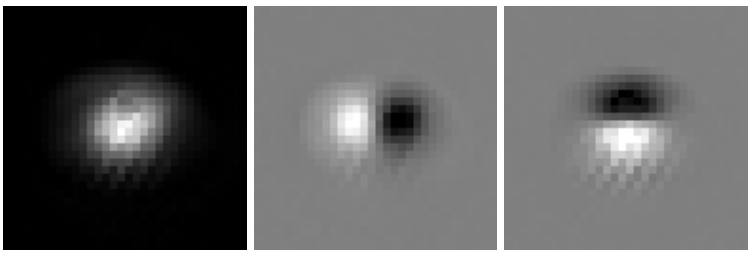


FIGURE 33: Exemple de données de calibration du CLOWFS mis en place sur le banc PIAA du HCIT. Les trois panneaux présentent de gauche à droite : l'image de référence, acquise lorsque la source est parfaitement alignée sur le coronographe ; la différence entre l'image instantanée et l'image de référence lorsque l'alignement est perturbé selon l'axe horizontal ; la même chose mais selon l'axe vertical.

POUR DE PETITES EXCURSIONS DE TIP-TILT ET AUTRES ABERRATIONS DE BAS ORDRE, la réponse de ce CLOWFS est linéaire : la différence entre l'image instantanée I et l'image de référence I_0 peut être projetée sur une base de modes acquis pendant une phase d'étalonnage (voir images de la figure 33). Dans l'expérience montée au HCIT, cette différence est simplement projetée sur les deux modes de tip et tilt, qui sont les plus importants à contrôler dans la perspective d'une mission d'imagerie haut contraste spatiale :

$$I - I_0 = \alpha R_x + \beta R_y + \epsilon + n, \tag{8}$$

où R_x et R_y sont les vecteurs de la matrice de réponse enregistrés pendant l'étalonnage, α et β, les coefficients à identifier, ϵ, un ré-

sidu associé à des modes non-contrôlés et n, un terme représentant le bruit de la mesure. Au HCIT (la description de ces résultats a fait l'objet d'un "Technology Milestone Report" joint à ce mémoire), nous avons pu montrer être capable de stabiliser, sur une échelle de temps de l'ordre de l'heure, la stabilité du pointage de la source sur le masque plan focal avec une erreur RMS estimée à $10^{-4}\ \lambda/D$ seulement, avec un biais estimé à $10^{-5}\ \lambda/D$.

CETTE INCROYABLE PRÉCISION, combinée à une excellente sensibilité intrinsèque, fait du CLOWFS une solution à mon sens incontournable pour un instrument haute performance dédié à l'imagerie haut contraste. CLOWFS fait donc évidemment partie des fonctionnalités offertes sur SCExAO : son implémentation actuelle, différente de celle décrite jusque là, est visible dans le panneau décrivant le banc IR de la figure 31. La R&D concernant ce sous-système sur SCExAO a pris deux directions : une utilisation pour le traitement *a posteriori* des images coronographiques qui a fait l'objet d'un travail en laboratoire par Frédéric Vogt, un des nombreux stagiaires de SCExAO[45] ; et une modification matérielle, rendant CLOWFS compatible avec les coronographes à masque de phase, qui a fait l'objet de la thèse de Garima Singh[46].

45. F. P. A. Vogt, F. Martinache, O. Guyon, T. Yoshikawa, K. Yokochi, V. Garrel, and T. Matsuo. Coronagraphic Low-Order Wavefront Sensor: Postprocessing Sensitivity Enhancer for High-Performance Coronagraphs. **PASP**, 123:1434–1441, Dec. 2011

46. G. Singh, F. Martinache, P. Baudoz, O. Guyon, T. Matsuo, N. Jovanovic, and C. Clergeon. Lyot-based Low Order Wavefront Sensor for Phase-mask Coronagraphs: Principle, Simulations and Laboratory Experiments. **PASP**, 126:586, June 2014







JPL Document D-74994

TECHNOLOGY DEVELOPMENT FOR EXOPLANET MISSIONS

Technology Milestone Report

Phase-Induced Amplitude Apodization (PIAA) Technology Development, Milestone #2

Instrument Tip-Tilt Control Demonstration at Sub-Milliarcsecond Levels

Olivier Guyon (PI)

University of Arizona

Brian Kern (Co-I), Jet Propulsion Laboratory

Frantz Martinache (Co-I), Subaru Telescope

25 Jun 2012

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

© 2012. All rights reserved

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

APPROVALS

Released by

Olivier Guyon, Principal Investigator
University of Arizona — JULY 17, 2012

Approved by

Peter R. Lawson
Exoplanet Exploration Program Chief Technologist, JPL — 7/17/2012

Marie Levine
Exoplanet Exploration Program Technology Manager, JPL — 7/17/2012

Peter R. Lawson
Exoplanet Exploration Program Manager (Acting), JPL — 7/17/2012

Douglas Hudgins
Exoplanet Exploration Program Scientist, NASA HQ — 7/30/2012

Lia LaPiana
Exoplanet Exploration Program Executive, NASA HQ — 7/23/2012

LIST OF ACRONYMS

CCD	charge-coupled device
CLOWFS	Coronagraphic Low-Order WaveFront Sensor
DM	deformable mirror
HCIT	High Contrast Imaging Testbed
IWA	inner working angle
JPL	Jet Propulsion Laboratory
MAM	Micro-Arcsecond Metrology
OAP	off-axis parabola
PIAA	phase-induced amplitude apodization
PSD	power spectral density (see Appendix 1)
PSF	point spread function
RMS	root mean square
sCMOS	Scientific CMOS (Complementary Metal-Oxide Semiconductor) image sensor

LIST OF ALGEBRAIC SYMBOLS

λ	wavelength of light (see also λ/D_{sky})
λ/D_{sky}	image plane distance, referenced to angles at input (Sec. 3.2)
δ_x and δ_y and δ	rms corrected residual x, y, and quadrature sum of x and y values (Sec. 3.5.7)
σ_x and σ_y and σ	mean corrected residual x, y, and quadrature sum of x and y values (Sec. 3.5.7)
ε_x and ε_y and ε	error in sinusoidal calibration meas., x, y, and quadrature sum of x and y (3.5.7)
Δ_x and Δ_y	actuator displacements used in CLOWFS calibration (Sec. 3.5.5)
C_x and C_y	$(I_x - I_{ref})/\Delta_x$ and $(I_y - I_{ref})/\Delta_y$ calibration of x and y actuator displacements (3.5.5)
C	$[C_x\ C_y]$ matrix of C_x and C_y column vectors (Sec. 3.5.5)
D	diameter of aperture or optical stop (see also λ/D_{sky}, Sec. 3.2), also
D	$I - I_{ref}$ column vector (Sec. 3.5.5)
f	focal length of optics (used only in definition of λ/D_{sky}, Sec 3.2), also
f	temporal frequency
i and j and k	indices to PSD calculations (Appendix 1)
I_{ref}	reference CLOWFS image (Sec. 3.5.5)
I_x and I_y	CLOWFS image after x and y actuator displacements (Sec. 3.5.5)
I	CLOWFS image (non-calibration)
m	mask of CLOWFS image pixels to be used (Secs. 3.5.1, 3.5.5), also
m	arbitrary measurement as input to PSD (Appendix 1)
M	PIAA off-axis magnification (Sec. 3.2)
P_x and P_y	PSD of x and y measurements
P	quadrature sum of P_x and P_y, implied radial (*i.e.*, total) motion PSD
S_x, S_y, $S_{1,x}$, $S_{1,y}$, S_2, S_3	Scaling ratios relating actuator voltages x_V, y_V to centroids at occulter x,y (3.5)
W	spectral window weighting coefficients (Appendix 1)
ξ and η	star centroid measured at science camera, in science camera pixels (Sec. 3.5.2)
x and y	star centroid x and y measurement at occulter, calibrated to λ/D_{sky} (Sec. 3.5.5)
x_V and y_V	x and y actuator voltages producing star centroid offsets, in volts (Sec. 3.5.5)
X	$[x_V\ y_V]^T$ 2-element column vector (Sec. 3.5.5)

TABLE OF CONTENTS

1. ABSTRACT 1
2. INTRODUCTION 1
2.1. Significance of milestone 1
2.2. Technical approach 2
3. EXPERIMENT CONFIGURATION 2
3.1. Optical layout 2
3.2. Off-axis distances 3
3.3. Occulter design 4
3.4. CLOWFS camera operating modes 5
3.5. Calibration and image analysis 6
3.5.1. CLOWFS image calibration 6
3.5.2. Science camera calibration of volts to pixels 7
3.5.3. Relation of science camera pixels to microns at the occulter 8
3.5.4. Relation of microns at the occulter to λ/D_{sky} 9
3.5.5. Image analysis 10
3.5.6. Control loop 11
3.5.7. Milestone statistics 11
3.6. Environment 11
3.6.1. In-air 11
3.6.2. Under vacuum 12
3.6.3. Artificial disturbances 14
4. MILESTONE PROCEDURES 15
5. NARRATIVE REPORT 15
5.1. Correction requirement 16
5.2. Calibration requirement 19
6. CONCLUSIONS 21
Appendix 1. Power Spectral Densities 22
Appendix 2. Error Estimates 23
A2.1. Random errors 23
A2.2. Systematic errors 23
A2.2.1. Errors in $S_{1,x}$ and $S_{1,y}$ 23
A2.2.2. Errors in S_2 24
A2.2.3. Errors in S_3 24
A2.2.4. Total systematic errors 24
REFERENCES 25

1. ABSTRACT

This report describes the completion of Milestone 2 of the Phase-Induced Amplitude Apodization (PIAA) Technology Development experiment, funded under the NASA 2009 Technology Development for Exoplanet Missions grants. PIAA Milestone 2 is a demonstration of instrument tip/tilt control in a high-contrast coronagraph at levels relevant for a space-based exoplanet imaging mission. The milestone is a requirement to "Demonstrate 0.01 λ/D RMS pointing jitter stability and 0.003 λ/D pointing calibration at the coronagraph focal plane mask over a continuous 1-hr time period."

This document describes the scientific relevance of tip-tilt control, the hardware used to create a Coronagraphic Low-Order WaveFront Sensor (CLOWFS) and tip-tilt control on the Jet Propulsion Lab (JPL) High Contrast Imaging Testbed (HCIT) PIAA testbed, the data analysis procedures, the relevant environmental concerns, and the results of the three Milestone 2 data runs satisfying the requirements. The milestone was met, with measured stability of approximately 0.001 λ/D and calibration to 0.0001 λ/D, near 10× and 30× better than the requirements.

2. INTRODUCTION

This milestone demonstrates the sub-milli-arcsecond tip-tilt control required for achieving and maintaining high contrast (10^{-9} raw contrast) at small inner working angles ($< 2\,\lambda/D$). While the milestone is part of an effort to advance the technology readiness of the PIAA coronagraph [1-10], both the significance of the milestone and the technical solution employed are applicable to all small inner working angle coronagraphs.

2.1. Significance of milestone

The first goal of this milestone is to demonstrate that pointing jitter at the coronagraph focal plane mask can be actively controlled to 0.01 λ/D RMS. On a 2 m diameter telescope, the 0.01 λ/D RMS pointing jitter corresponds to 0.5 milli-arcsecond (mas) at optical wavelengths, which is the radius of a Sun-like star at 10 pc distance. Achieving the RMS pointing jitter goal specified for this milestone therefore ensures that the coronagraph leak due to pointing errors is no more than the light leak due to the finite stellar angular size for telescope diameters of 2 m or more. Small inner working angle (IWA) coronagraphs, such as the PIAA coronagraph, are fundamentally limited in performance by the stellar angular size, while coronagraphs with larger IWA are insensitive to stellar angular size and can tolerate larger pointing errors. By setting the RMS pointing jitter goal at 0.01 λ/D, completion of this milestone thus validates that pointing jitter can be controlled to a level such that it is not a major contributor to coronagraph light leakage, regardless of the coronagraph design employed.

The second goal of the milestone is to demonstrate that residual tip-tilt errors in the coronagraph can be calibrated to 0.003 λ/D in order to allow calibration of coronagraphic leaks due to pointing errors to a 10^{-11} contrast level on a small IWA coronagraph. While the first goal (RMS jitter) ensures that the contribution of pointing errors is below the 10^{-9} contrast level and is therefore not a significant source of photon noise, this second goal ensures that pointing errors cannot produce planet detection false positives.

The rationale for these goals is discussed in more detail in Sec. 2.2 of the Milestone 2 White Paper.

2.2. Technical approach

The approach for accomplishing this milestone is to implement and operate a dedicated sensor, the Coronagraphic Low-Order Wave-Front Sensor (CLOWFS) [11], which uses starlight otherwise rejected by the coronagraph. Using the light that falls on the central (within the coronagraph IWA) part of the focal plane mask offers two fundamental advantages over schemes relying on analysis of coronagraphic science images for pointing control:

(1) a large number of photons is available for the measurement, allowing fast and accurate tip-tilt estimation

(2) pointing errors can be measured before they start producing coronagraphic leaks in the science image

The milestone demonstration was performed on the existing PIAA coronagraph table which is currently in the vacuum Micro-Arcsecond Metrology (MAM) chamber at JPL. The CLOWFS implementation for this milestone closely follows the design described in the original CLOWFS publication [11]. The only significant difference is that only one defocused image is acquired by the CLOWFS camera, while the original CLOWFS assumes two images are acquired. This simplification has no impact on the CLOWFS tip-tilt functionality or performance.

3. EXPERIMENT CONFIGURATION

3.1. Optical layout

The HCIT PIAA testbed was assembled in 2009 and has been conducting high-contrast imaging experiments since [12-13]. It uses the first generation of PIAA mirrors fabricated by Axsys [7]. The implementation of a CLOWFS system required a new occulter, relay optics to deliver an image to the CLOWFS camera, and piezo actuators to perform the correction. The current testbed layout, including the components installed as part of this experiment, is shown in Fig. 1. All of the optical components are mounted on a 5'×8' optical table, which is installed in a vacuum chamber whose pressure typically reaches 10^{-6} Torr. Light is fed from a laser source outside the vacuum chamber through a single-mode fiber.

The optical path relevant to the CLOWFS system begins with an 808 nm frequency-stabilized laser diode, pigtailed to a single-mode fiber that passes via a feedthrough into the vacuum chamber. The fiber tip then illuminates a pinhole at the source location marked in Fig. 1. The source is followed by a fold mirror with piezoelectric actuators controlling its tip-tilt. The light then passes through the two PIAA mirrors, M1 and M2. The output focus of M2 is then relayed 1:1 to the occulter, where part of the PSF is reflected to the CLOWFS camera and part is transmitted to the science camera. A pair of lenses relays the light reflected off of the occulter to an image near the CLOWFS camera with a 2.5× magnification. This image is intentionally defocused before landing on the CLOWFS camera.

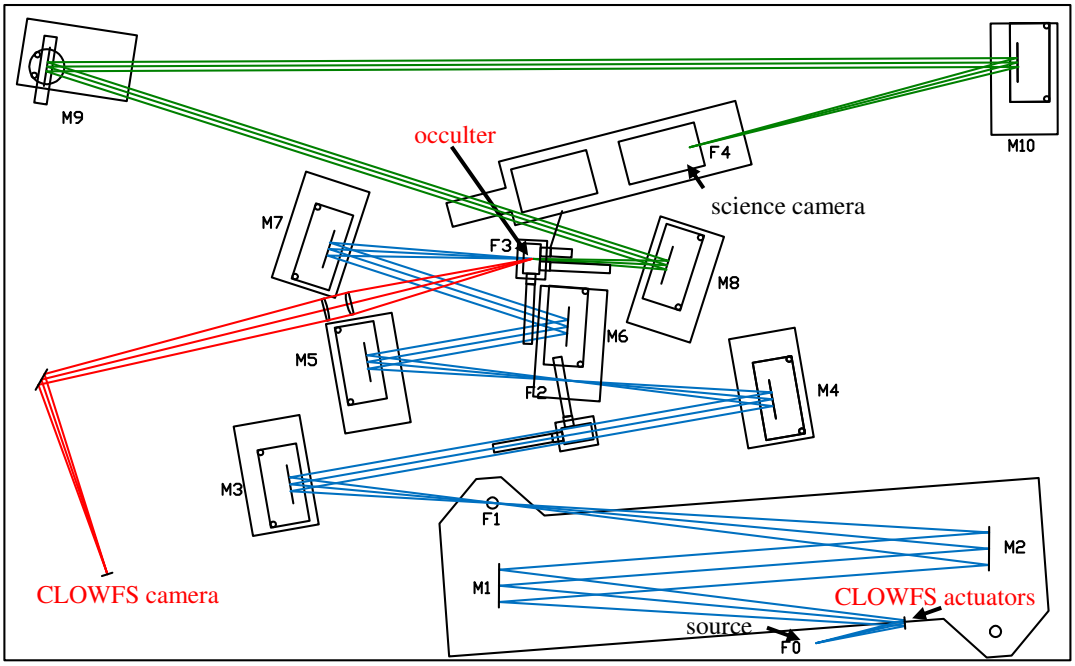


Fig. 1. Optical layout of HCIT PIAA table. The source is at the bottom-right, followed by a piezo-driven tip-tilt actuated fold mirror ("CLOWFS actuators"), then by PIAA mirrors M1 and M2, which form a source image at F1. M3 is an off-axis parabola (OAP) that collimates the light from F1 and forms an image of PIAA M2 at the postapodizer (shown here as an unmarked translation stage). OAPs M4 and M5 image and collimate the light from F1, and relay an image of the postapozider onto the deformable mirror (DM) at M6. The OAP M7 forms a source image at the occulter (F3), which reflects part of the beam into the CLOWFS path (in red) and transmits part to the science camera path (in green, through OAP M8, flat M9, and OAP M10). The plane of the table is the x,z plane (variously oriented by defining z along the chief ray), and y is perpendicular to the table (y is parallel to gravity).

3.2. Off-axis distances

For the purposes of calibrating and controlling tip-tilt in this system, the only aspect of the PIAA mirrors that is relevant is that the images of off-axis sources receive a magnification, as measured by the location of the centroid of an off-axis image relative to the off-axis position of the source, that is not simply an effect of changing focal lengths and aperture diameters, *i.e.*, it is not a result of scaling by $f\lambda/D$. The imaging scales of the system can be understood to be defined by (f/D) at each image plane, as would typically apply to OAPs (or any powered optics which have conic surfaces), while the "additional" effects of the PIAA mirrors are encapsulated in a single multiplicative parameter M. This off-axis magnification due to the PIAA mirrors depends on the PIAA mirror prescription, as well as on the limiting stop. When this system is used for coronagraphic measurements at the science camera, a Lyot stop is inserted between the occulter and the science camera. This Lyot stop determines the off-axis magnification M for these PIAA mirrors to be $M = 2.5$, as seen by the science camera with the coronagraph fully assembled. For the

purposes of this tip-tilt experiment, all off-axis distances will be quoted as if in units of λ/D_{sky}, which is defined by:

$$\lambda/D_{sky} = \begin{cases} f\lambda/D, & \text{image planes upstream of PIAA,} \\ Mf\lambda/D, & \text{image planes downstream of PIAA,} \end{cases}$$

where $M = 2.5$, $\lambda = 808$ nm, and $(f/D) = 16$ (approximately) at the source and at the occulter. While the expression λ/D_{sky} is dimensionless as written, which is appropriate for a telescope measuring angles on the sky, in this experiment it always has dimensions of length, as defined above. Specifically, at the source, $\lambda/D_{sky} = 13$ μm, and at the occulter, $\lambda/D_{sky} = 33$ μm.

This off-axis magnification, as a multiplicative factor, is essentially constant as a function of off-axis distance for the purposes of the current experiment, varying by < 1% for off-axis distances well beyond 100 λ/D_{sky}.

3.3. Occulter design

The occulter used in conjunction with the CLOWFS system must prevent light within the coronagraph IWA from reaching the science focal plane, and direct starlight to the CLOWFS camera for tip-tilt measurement. As shown in Fig. 2, the occulter, when seen in transmission, blocks all light falling within 1.6 λ/D_{sky} of the optical axis. A reflective annulus directs some of the starlight onto the CLOWFS camera, while the central part of the mask is absorptive. Of the total PSF, 9% of the intensity falls on the reflective annulus, while 6% passes through the blank substrate.

As detailed in the original CLOWFS publication [11], the role of the opaque central core is to prevent most of the light in the PSF core from reaching the CLOWFS. This light is not very sensitive to tip-tilt aberrations, while the light in the steepest part of the PSF profile (falling on the reflective annulus) is most sensitive to small tip-tilt errors. The presence of this opaque core thus offers two fundamental benefits:

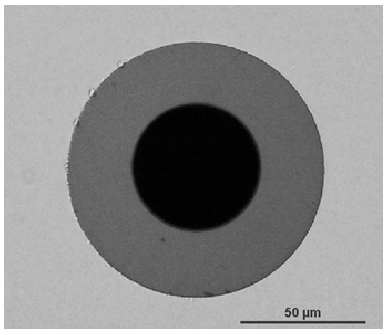


Fig. 2. Microscope image (50×) of CLOWFS occulter, on a glass substrate. The center portion is absorptive, the surrounding annulus is reflective, and the outer region is transmitting. The reflected portion of the beam goes to the CLOWFS camera, while the transmitted portion goes through an optional Lyot stop to the science camera. The absorptive core has a radius of 0.8 λ/D_{sky}, and the reflective annulus has an outer radius of 1.6 λ/D_{sky}.

(1) A small tip-tilt error produces a macroscopic change in the light reflected to the CLOWFS camera.

(2) The measurement is largely immune to non-common path tip-tilt errors (for example, motion of the CLOWFS camera), as the measurement is referenced to the location of the opaque spot on the focal plane mask.

The occulter shown in Fig. 2 was manufactured by lithographic techniques on a glass substrate. Several such masks were manufactured with slightly different outer radii. The mask adopted for the milestone measurements offers a 2 λ/D_{sky} coronagraph IWA.

3.4. CLOWFS camera operating modes

The CLOWFS camera currently in use is an E2V CCD39 with JPL-built supporting camera electronics. This is a split frame transfer, four-quadrant readout CCD, with 80×80 active 24 μm square pixels. The HCIT PIAA table can be operated either in air or under vacuum, and the operating environment imposes a number of specific constraints in the CLOWFS camera operation. The CLOWFS camera has a fixed frame rate, as currently implemented, of 10 Hz, but has latencies and delays that range from 0.5 s to 7 s, depending on operating mode and number of frames being collected as a batch. This frame rate is much lower than what the CCD is capable of (1000 Hz), but is limited by the software and hardware currently implemented. Rather than improving the current electronics to make use of available bandwidth, an entirely new "Scientific CMOS image sensor" (sCMOS) camera is being prepared for use as a replacement, which delivers 100 Hz (potentially up to 1.7 kHz) frame rates but has not yet been interfaced with the HCIT PIAA table.

In air, there are two CLOWFS camera modes most frequently used: a batch mode acquiring several frames at a time, and single frame mode. In batch mode, typically 100 contiguous frames are acquired at 10 Hz (taking 10 s total), followed by a 5 s latency for software overhead. In single frame mode, a single 0.1 s frame is collected and is available within milliseconds of being read out, but due to software overheads, 0.5 s pass before another frame can be stored, resulting in a latency of ms but a minimum cycle time of 0.6 s. To perform high-bandwidth corrections, the single-frame mode can be used in air, with a 1.7 Hz correction frequency.

Under vacuum, the camera fails from overheating if left powered on continuously, but can operate in a quasi-steady state indefinitely with power duty cycles below 1/4. The vacuum operating mode is then to power on the camera, typically acquire 16 frames at 10 Hz, then power down the camera, repeating the cycle once every 7 s. The images are available within approximately 2 s of readout. Of the 16 frames acquired, the first 4 are corrupted by the power-on sequence of the camera, and the remaining 12 frames (or fewer if there are read errors) are averaged to produce a single mean frame.

The CLOWFS camera is operated uncooled, and its read noise and dark current performance do not affect the work presented here in any meaningful way (an upper limit to this noise is identified in Sec. 3.6.2). To keep light levels appropriate to the fixed 10 ms exposure time, a 10^{-4} neutral density filter is inserted before the CLOWFS camera. Every frame taken with the CLOWFS camera is bias-subtracted, gain-normalized between the four amplifiers, and dark-subtracted before any additional data reduction described in this report.

3.5. Calibration and image analysis

The calibration steps of the CLOWFS system can be divided into two categories: calibration of the CLOWFS camera and actuators to allow closed-loop correction in terms of actuator voltages, and determination of a scale factor relating volts to $\lambda/D_{\rm sky}$ measures. The CLOWFS tip-tilt corrections do not rely on the scaling of voltages to $\lambda/D_{\rm sky}$, and so require only the first category to operate. A schematic diagram of the different calibration steps is shown in Fig. 3.

The scaling relationships are quantified as two constants, S_x and S_y. These constants relate the actuator voltages $x_{\rm V}$ and $y_{\rm V}$ to centroids at the occulter, x and y (measured in $\lambda/D_{\rm sky}$), via

$$x = S_x\, x_{\rm V}$$
$$y = S_y\, y_{\rm V}$$

The scaling constants themselves are broken into individually determined factors,

$$S_x = S_{1,x}\, S_2\, S_3$$
$$S_y = S_{1,y}\, S_2\, S_3$$

where $S_{1,x}$, $S_{1,y}$, S_2 and S_3 are defined in the sections that follow.

3.5.1. CLOWFS image calibration

The CLOWFS camera and actuator calibration involves only empirical correlations between actuator motions and observed changes in CLOWFS images, with no consideration given to *a priori* calculations of the optical properties of the system. The procedure is to acquire, in principle, three images: a reference image, an image taken after moving only the x actuator, and an image taken after moving only the y

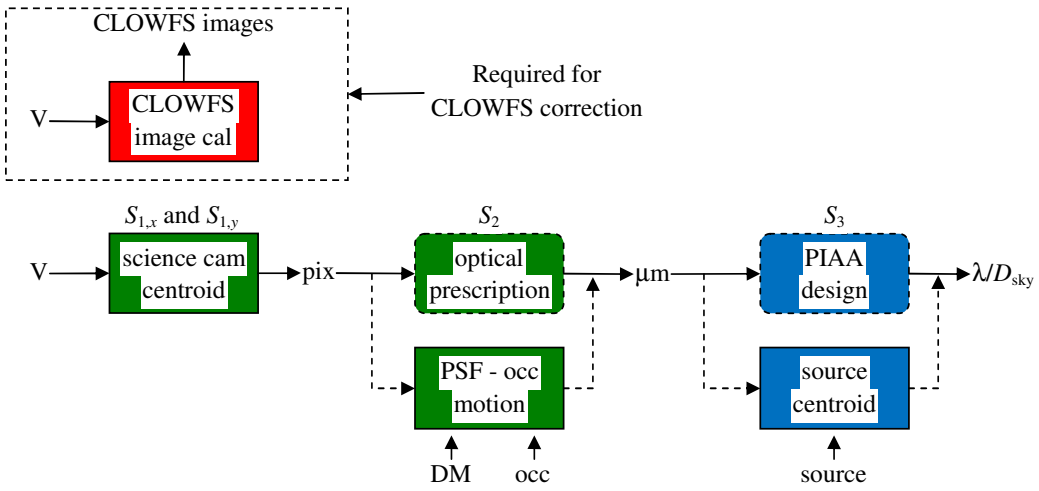


Fig. 3. Schematic showing relationships between calibration steps. The "CLOWFS image cal" step (top-left, in red) is the only step that is relevant for tip-tilt correction; the remaining steps are required to scale residual offsets to $\lambda/D_{\rm sky}$. The rounded boxes with dashed borders represent non-measured quantities, and the dashed flow-lines denote verification steps which are used to establish error estimates but are not included in the direct calibration. The steps and inputs involving the focal plane occulter use the abbreviation "occ." The colors relate to the regions of Fig. 1 where the relationships are established.

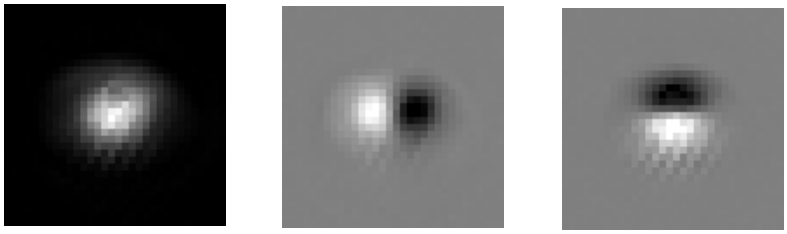


Fig. 4. (LEFT) CLOWFS reference image, (CENTER) CLOWFS *x*-calibration difference image, (RIGHT) CLOWFS *y*-calibration difference image, both calibrations using approximately $50\times10^{-3}\ \lambda/D_{\rm sky}$ offsets. The intensity scale on the center and right panels is ± 0.04× the full scale on the left panel. These images were the calibration for Milestone run #1.

actuator. Every image, during calibration and during correction, is normalized by total counts in a photometric aperture 12 pixels in radius. The normalized reference image is subtracted from every normalized image taken during calibration and during correction; all data is considered to be differential with respect to the reference image. The CLOWFS *x*- and *y*-calibration difference images, examples of which are shown in Fig. 4, form a two-element basis used to reduce all subsequent difference images into *x* and *y* offsets (scalars), which are the error signals for closed-loop correction. Because the relevant region of the image is small compared to the entire CCD field, a mask is applied, with radius 12 pixels, to exclude the "zero" portions of the image from further analysis. This mask is identical to the photometric aperture used for normalization.

Using the image analysis described in Section 3.5.5 below, all subsequent images used in normal operation of the CLOWFS camera (*i.e.*, non-calibration images) are reduced to a pair of *x*-actuator and *y*-actuator voltages, which can be considered an error signal relative to the configuration of the reference image. This calibration alone is enough to perform closed-loop corrections, in terms of actuator voltages. This CLOWFS image calibration is repeated every milestone run.

3.5.2. Science camera calibration of volts to pixels

The science camera is used to establish a correlation between piezo actuator voltages and centroid motion of the PSF at the occulter, measured in science camera pixels. This is quantified by the scaling constants $S_{1,x}$ and $S_{1,y}$. The occulter is removed, so that the science camera is directly imaging the occulting plane. By applying sinusoidal voltage patterns to the *x* and *y* actuators, simultaneously at two different temporal frequencies, and measuring the centroid motion at the science camera, a relationship between volts and science camera pixels is established. These centroid measurements are shown in Fig. 5, with ξ and η denoting the science camera pixel axes (*x* approximately aligned with ξ, and *y* with η). The relationship between volts and pixels is different for the *x* actuator than for the *y* actuator, because the piezo actuators themselves have different gains. There was no accommodation given to cross-terms in the analysis of the centroid data, as would be expected by a misalignment of the actuator axes and the science camera pixel axes (certainly present at some level). This is readily apparent in the ξ residuals in Fig. 5, which show the temporal frequency of the *y* actuator sinusoid, with about 2.7% of the η amplitude. The magnitude of this term does not affect the calibration enough to merit a more complicated treatment (residuals in Fig. 5 have a different vertical scale than the original measurements).

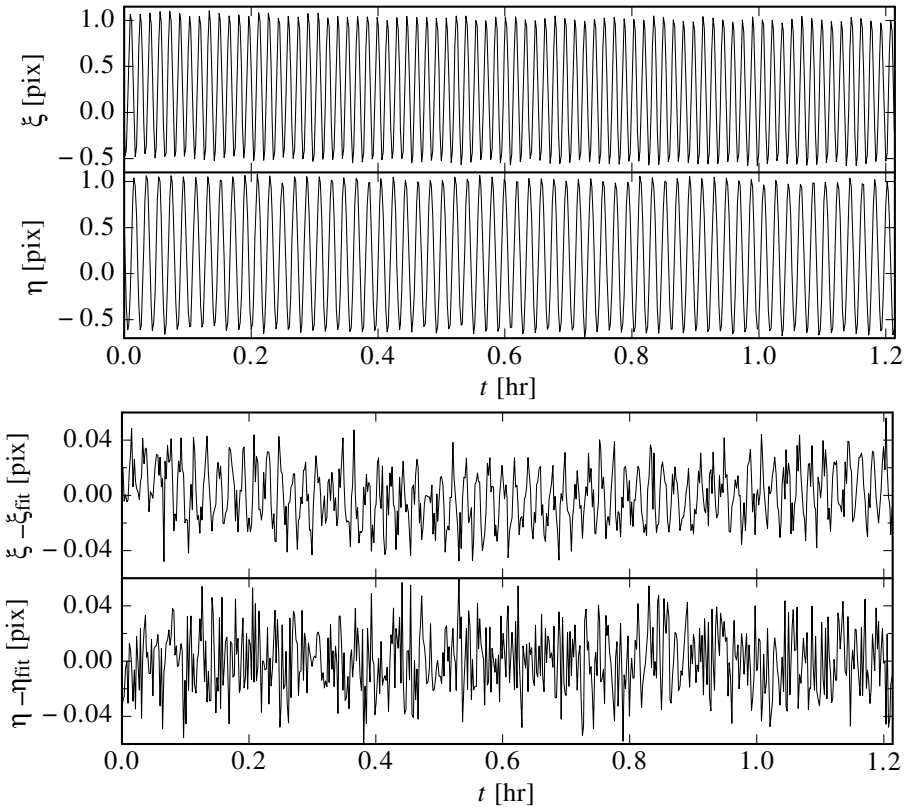


Fig. 5. (TOP) Science camera centroid x- and y-calibration of CLOWFS actuator, and (BOTTOM) residuals from fit. The data are ξ- and η-centroids of unocculted images taken with the science camera, while applying 2 V sinusoids at two different temporal frequencies to the x and y CLOWFS actuators. The data are fitted to the sum of a first-order polynomial and a sinusoid of unknown amplitude and phase (frequency is not fit). The residual plots are displayed with a 15× finer vertical scale than the centroid data. The zero-points of ξ and η in the top panel are nominal pixel boundaries, which have arbitrary offsets with respect to the initial position of the star centroid.

3.5.3. Relation of science camera pixels to microns at the occulter

There are two techniques used to determine the relationship between pixels at the science camera and microns at the occulter, the first of which determines S_2 and the second of which estimates the errors on S_2 (an upper limit to the uncertainty). The first technique uses the *a priori* defined specification of the CCD and the optical prescription of the PIAA coronagraph. The second technique is a dynamic empirical technique involving measurements with the science camera.

The first technique to measure S_2, labeled "optical prescription" in Fig. 3, involves just two numbers. The two OAPs that image the occulter onto the science camera have a nominal magnification of 2.5, and the science camera CCD pixels are nominally 13 µm square, implying S_2 = 5.2 µm/pix at the occulter. This number is used in the analysis that follows.

The second technique also measures S_2, but is used only to estimate errors on S_2. This is labeled "PSF – occ motion" in Fig. 3. The technique is to take a science camera image with the occulter in, to move both the PSF and the occulter to recreate the same science camera image shifted by some number of pixels. The occulter motion is controlled by an encoded stepper motor, so the encoder can then be used to determine the distance the occulter moved, corresponding to a number of pixels of motion at the science camera.

This second technique for S_2 works best if the occulted PSF is morphologically unchanged except for location on the science camera. The CLOWFS actuators are not appropriate to offset the PSF in this way, as the PIAA optics create a distorted off-axis PSF when the actuators move by substantial amounts. Rather, a set of piezo actuators that control tip-tilt of the DM mount are used for this calibration; these actuators are not used in any other part of this experiment, but they do allow for simple translation of the PSF at the occulter, with negligible change to the PSF morphology.

The scaling from the first technique for S_2, using a priori knowledge, is used for all image analysis, while the results from the second technique appear in the error analysis in Appendix 2.

3.5.4. Relation of microns at the occulter to λ/D_{sky}

There are two techniques used to determine S_3, the relationship between microns at the occulter and λ/D_{sky}. The first S_3 technique is to use the nominal optical prescription, and the PIAA off-axis magnification, to calculate $Mf\lambda/D$. The second S_3 technique uses a dynamic empirical technique to measure relevant quantities at the science camera. Just as in the previous section, the *a priori* knowledge (first technique) determines S_3 as used in the analysis, while the second technique provides an error estimate.

The first S_3 technique is labeled "PIAA prescription" in Fig. 3. The optical prescription gives $f\lambda/D$, and the PIAA off-axis magnification (see Sec. 3.2) is $M = 2.5$, giving $\lambda/D_{sky} = 33$ μm at the occulter. This number is used in the image analysis to follow.

The second S_3 technique, labeled "source centroid" in Fig. 3, begins by offsetting the source by a known distance, using an encoded motor to translate the source (not the piezo-controlled tip-tilt of a fold mirror), and measuring the centroid motion at the science camera, with the Lyot stop in. This gives a relationship between microns of motion at the source and pixels of motion at the science camera, which can be converted (using the result of 3.5.3) to relate microns of motion at the source to microns of motion at the occulter, which is a measurement of $(Mf\lambda/D)_{occ}$ / $(f\lambda/D)_{src}$. Nominally, $(f/D)_{occ} = (f/D)_{src}$, and of course λ is the same everywhere, so this measurement yields a measurement of M. f/D is measured separately by moving the camera longitudinally between a source image plane and a pupil image plane, using the encoder on the camera motion stage to know the distance, and measuring the pupil image diameter in pixels at the science camera, with the Lyot stop out. Using the known pixel size, this establishes f/D at the camera via empirical measurements. The magnification between the camera and the occulter was established as 2.5 in Sec. 3.5.3 and A2.2.2, so $(f/D)_{cam} = 2.5\ (f/D)_{occ}$. Using the specified laser wavelength (measured by a spectrometer to be accurate to < 0.1%), this yields a measurement of $Mf\lambda/D$ at the occulter, having assumed the science camera pixel size and $(f/D)_{occ} = (f/D)_{src}$.

The details of the use of the results from the second technique's error estimates appear in Appendix 2.

3.5.5. Image analysis

The calibration difference images are $C_x = (I_x - I_{ref})/\Delta_x$, and $C_y = (I_y - I_{ref})/\Delta_y$, where Δ_x and Δ_y are the x and y actuator motion, in V, used to construct the calibration difference images, as described in Section 3.5.1. C_x and C_y have dimensions of normalized intensity per volt. A linear model of actuator motion would then predict, for an arbitrary choice of actuator commands x_V and y_V, a difference image, $D = I - I_{ref}$, according to

$$D = x_V C_x + y_V C_y,$$

which is rewritten in matrix notation as

$$D = CX$$

where D is a vector of individual normalized intensities per pixel in the difference image, C is a matrix with two columns, a column of values per pixel from C_x, and a column of values per pixel from C_y, and X is a two-element vector containing x_V and y_V. The two-dimensional nature of each image (D, C_x, and C_y) is not relevant beyond ensuring that the indexing of pixels is the same between images, *i.e.*, any linear operations operate on the same two-dimensional locations in each image. Consistent pixel registration is relevant, while identifying separate pixels as being neighbors or not, for example, is not relevant.

The mask m describing the region of interest (as described in Section 3.5.1) is a set of pixels, in this case, 437 pixels. Restricting the calculations to these pixels implies that D is a 437-element column vector, C is a 437×2 matrix, and X is a 2-element column vector.

For each subsequent image, using its associated difference image $D = I - I_{ref}$, we solve $D = CX$ for X in a linear least-squares sense, using

$$X = (C^T C)^{-1} C^T D$$

By construction, C_x and C_y are nearly orthogonal, so $C^T C$ is well conditioned and the inversion is straightforward. This treatment implies identical weighting for each pixel in the image plane. While it would be preferable to weight them inversely to the expected measurement noise per pixel, in practice, the measurement noise is small compared to any relevant signals in question (see Section 3.6.2), and the complication of nonuniform weighting was not necessary.

One benefit of all image analysis being differential is that this linear analysis will be unaffected by errors in bias subtraction, dark subtraction, and flat-fielding.

Once each difference image is reduced to x_V and y_V as above, these error signals are translated into λ/D_{sky} units, as established in sections 3.5.2-3.5.4, and denoted simply x and y. In addition to interpreting x and y as the actuator motions that would translate the stellar image from the reference image to the observed image, x and y can also be interpreted as the x and y star centroids at the occulter, measured in λ/D_{sky}, assuming the applicability of linearity in the measurements. This linearity assumption is tested, in part, by the milestone calibration verification requirement, described in Sec. 5.2, where the agreement is found to be excellent.

3.5.6. Control loop

An integrator control law was used, with an integrator gain equal to 1. In the simplest terms, new voltage = old voltage − 1× measured offset. Under vacuum (see Secs. 3.4 and 3.6.2), sensing was done once every 7 s, using (at most) 12 frames of 0.1 s each, all averaged for a single measurement. The duty cycle of exposure time to sensing period is then 1.2 / 7 = 0.17. In a single sensing cycle, the time lag between the end of the exposure of the last frame and the application of a correction was approximately 0.5 s (this number was not well characterized). The temporal delay between sensing and measurement is thus much shorter than the time span between measurements.

The CLOWFS response is only linear if the tip-tilt error is within ~ $300\times10^{-3}\ \lambda/D_{\mathrm{sky}}$ (the exact level depends on the degree of nonlinearity tolerated), as described in Ref. 11. The results described in this report were all acquired within this range, but the conclusions from these results should not be extrapolated beyond the linearity domain. Linearity impacts the CLOWFS performance on two levels:

(1) it may be difficult to close the CLOWFS control loop if the initial tip-tilt value is well outside the linearity range. Laboratory and on-sky experience at the Subaru Telescope has shown that the CLOWFS loop converges with an initial error as large as $0.5\ \lambda/D_{\mathrm{sky}}$.

(2) non-linearities may affect the interpretation of the closed loop data. In this report, the tip-tilt residuals after correction are too small (~ $10^{-3}\ \lambda/D_{\mathrm{sky}}$) for such non-linearity to be noticeable, but care must be taken when extrapolating these results and the inferred error transfer function to large input error levels.

3.5.7. Milestone statistics

The three statistics that are used to satisfy the milestone requirements are an rms value, to measure the size of the tip-tilt residuals (*i.e.*, jitter), a mean, to verify that the average offset is acceptably small, and the measured error in a sinusoidal calibration signal. The rms values are $\sigma_x = \langle x^2 \rangle^{1/2}$, $\sigma_y = \langle y^2 \rangle^{1/2}$, and $\sigma = \langle x^2 + y^2 \rangle^{1/2}$, where $\langle \ldots \rangle$ is a mean over an entire sequence. Note that the rms values are not standard deviations, in that they do not subtract the mean x and y values. The mean values are $\delta_x = \langle x \rangle$, $\delta_y = \langle y \rangle$, and $\delta = (\langle x \rangle^2 + \langle y \rangle^2)^{1/2}$. The measurement of x and y sinusoidal calibration signals, of known amplitudes, will have errors (difference between known and measured amplitudes) ε_x and ε_y, and a total error $\varepsilon = (\varepsilon_x^2 + \varepsilon_y^2)^{1/2}$. Both the δ and ε statistics are considered calibration residuals.

3.6. Environment

3.6.1. In-air

When operating in air, the uncorrected x and y power spectral densities were measured (the definition of power spectral density, along with the window used in analysis and smoothing of the plots, used here is described in Appendix 1), and their sum is shown in Fig. 6. It is adequately represented as a broken power law, $P \propto f^n$, with indices $n = 0$ for $f < 0.03$ Hz, and $n = -3$ for $f > 0.03$ Hz. The spectrum shown in Fig. 6 is just one example, but on other occasions, while the low-frequency power level and knee frequency were similar, the high-frequency power law index was shallower, $n \sim -2$, implying more high-frequency power and correspondingly larger residuals after closed-loop correction. The reasons for the different in-air behaviors have not been investigated.

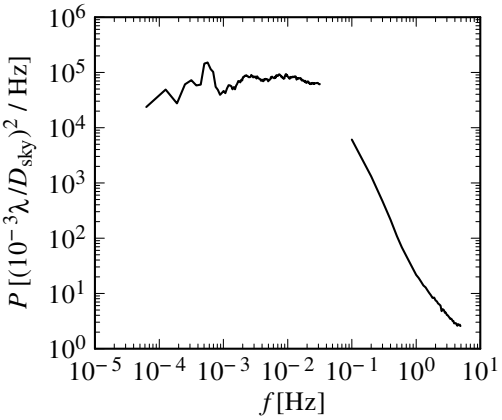


Fig. 6. Power spectral density of uncorrected CLOWFS measurements in air, with x and y powers summed. The data were acquired in sets of 100 images taken at 10 Hz, followed by a 7 s delay before acquiring the next set of 100 images. The PSD below 0.03 Hz is calculated between sets of images, while the PSD above 0.1 Hz is an average of PSDs from sets of 100 images. The break in the frequency coverage from 0.03 Hz to 0.1 Hz is indicative of the pause between sets of 100 images. The plot is smoothed as described in Appendix 1.

The uncorrected in-air rms motion (quadrature sum of x and y) shown in Fig. 6 over a ~ 4.5 hour time span is approximately $40\times10^{-3}\ \lambda/D_{\rm sky}$, larger than the milestone threshold of $10\times10^{-3}\ \lambda/D_{\rm sky}$ by a factor of 4. The expected residual level after closed-loop correction, with a correction every 0.6 s, would be in the neighborhood of $20\times10^{-3}\ \lambda/D_{\rm sky}$, which depends strongly on the high-frequency power-law index, and on the closed-loop gain at frequencies above the unity-gain frequency, $1/(2\pi\times\ 0.6\ {\rm s}) = 0.27$ Hz, the frequency beyond which existing power is amplified.

The data shown in Fig. 6 were acquired over a longer time interval than the milestone requirement (4.5 hours as compared to a 1 hour minimum requirement), and reducing the time interval would reduce the rms image motion. However, because the power spectral density is relatively flat at low frequencies, the rms does not decrease significantly until the time interval decreases to below 100 s (*i.e.*, $\int P\,{\rm d}f$ from 0 to 10^{-2} Hz is small compared to $\int P\,{\rm d}f$ from 10^{-2} Hz to ∞). For time intervals relevant to the milestone requirement, the uncorrected rms will always be around 40×10^{-3}.

In any case, the only road to achieve milestone levels in this in-air environment is to increase the correction bandwidth, which is limited by the camera readout rate. As stated in Section 3.4, this approach has been postponed until a new camera system is integrated into the coronagraph.

3.6.2. Under vacuum

The image motion under vacuum is far less than that in air. A 12-hour sample of uncorrected CLOWFS measurements, taken under vacuum with CLOWFS camera images sampled once every 7 s, is shown in Fig. 7. What is immediately apparent about the y measurements (y is vertical) is that they are dominated by a linear drift term. The origin of this drift is not known, and does not correlate with temperature. For the purposes of describing the power spectral density of the uncorrected CLOWFS measurements under

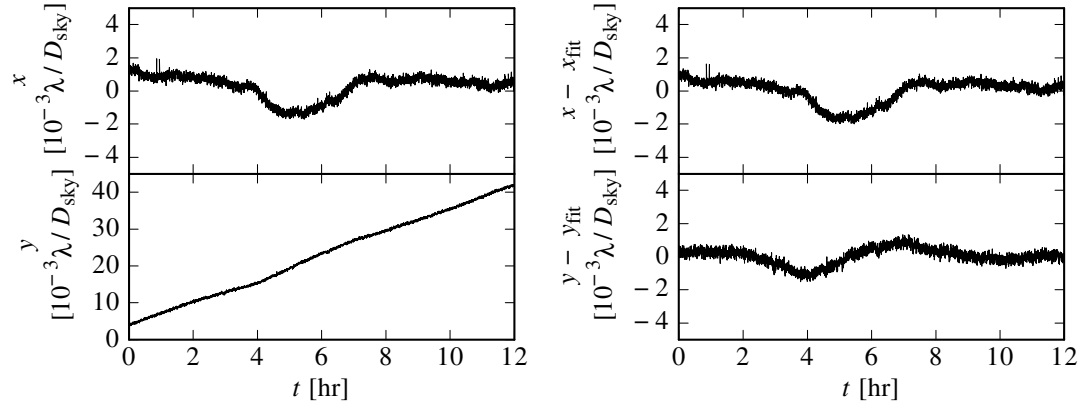


Fig. 7. (LEFT) Uncorrected CLOWFS measurements, (RIGHT) uncorrected CLOWFS drift-subtracted (first-order polynomial) residuals, taken under vacuum with no disturbances added.

vacuum, it is worthwhile to examine a drift-subtracted time series of measurements, shown in the right-hand panel of Fig. 7. A first-order polynomial is subtracted to remove drift. The drift-subtracted rms motion (quadrature sum of x and y) over the entire 12-hour span is $0.75\times10^{-3}\ \lambda/D_{\rm sky}$, dominated by low-frequency power. This is far below the milestone requirement of $10\times10^{-3}\ \lambda/D_{\rm sky}$, without applying any correction.

The power spectral density of the uncorrected drift-subtracted CLOWFS motion under vacuum is shown in Fig. 8. What is apparent, beyond the dominance of low-frequency power, is that the high-frequency power appears to have reached an asymptote, likely a white noise level due to measurement noise. This is a comfortably low noise level, at approximately $0.5\times(10^{-3}\ \lambda/D_{\rm sky})^2$/Hz. At this level, the rms noise contribution integrated over all frequencies up to the Nyquist frequency of (1/14) Hz, is only $0.2\times10^{-3}\ \lambda/D_{\rm sky}$.

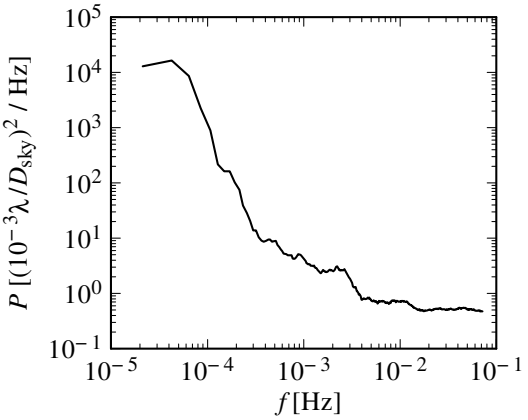


Fig. 8. Power spectral density of uncorrected drift-subtracted residuals in right-hand panel of Fig. 7, with x and y powers summed. The flat portion of the spectrum at the highest frequencies is indicative of a white measurement noise component, at a level of approximately $0.5\times(10^{-3}\ \lambda/D_{\rm sky})^2$/Hz, corresponding approximately to $0.2\times10^{-3}\ \lambda/D_{\rm sky}$ rms noise in the time sequence of Fig. 7. The plot is smoothed as described in Appendix 1.

3.6.3. Artificial disturbances

Given the low "background" CLOWFS motion power under vacuum, the approach taken here to demonstrate the Milestone 2 performance is to operate under vacuum and introduce disturbances through the same actuators used to implement the CLOWFS correction. The power spectral density of these disturbances is chosen to mimic the in-air motion broken power-law shape, with a knee frequency chosen to be (1/7200) Hz. This particular knee frequency was chosen for two reasons, first to be relevant to the time interval appropriate to the milestone demonstration (1250×7 s = 8750 s), and second to be higher than the observed in-air knee frequency (0.03 Hz) scaled by the expected future CLOWFS operating frequency (100 Hz) divided by the current operating frequency (1/7 Hz), which comes to (1/23000) Hz. The implication is that after upgrading to the new CLOWFS camera, the in-air corrected performance should be better than the milestone performance demonstrated here under vacuum with added disturbances.

The overall level of added disturbances was chosen to deliver around $90\times10^{-3}\ \lambda/D_{\mathrm{sky}}$ rms in a two-hour interval. This corresponds to a 5 mas pointing jitter for a 2 m telescope in visible light, comparable to the pointing jitter achieved in orbit on existing observatories (HST, Kepler and Spitzer). The level of added disturbance is therefore representative of what may be experienced on a space telescope which is otherwise not highly optimized for pointing jitter mitigation. We however note that the frequency distribution of the pointing error is very specific to the telescope and spacecraft design (reaction wheel speed and damping), and our added disturbance power spectrum (described below) may thus be quite different from what would be experienced by the telescope.

The power-law indices chosen for the disturbances were $n = 0$ and $n = -3$ for frequencies below and above the knee, similar to what was seen in Fig. 6. The disturbances were calculated and applied at intervals of 3.5 s, two samples for every one CLOWFS camera frame and correction iteration (once every 7 s). When no corrections are applied, the x and y measurements are measurements of the disturbances alone. Specifically, they are the sum of the ambient under-vacuum disturbances (Fig. 8) and the artificial disturbances, but in every case the ambient under-vacuum disturbances are negligible compared to the artificial disturbances.

Each pre-calculated sequence of disturbances can be applied to the system as many times as desired. Specifically, a sequence of disturbances can be applied while simultaneously applying corrections, and then the same sequence can be applied without also applying corrections. The same image analysis (Sec. 3.5.5) is performed whether corrections are applied or not, ensuring that the differences between measurements of corrected and uncorrected sequences using the same disturbances do represent the effect of the corrections, to the level of repeatability of the disturbances. The measured repeatability between two 2-hour sequences (separated by 6.5 hours) is $10^{-3}\ \lambda/D_{\mathrm{sky}}$ rms, compared to the $90\times10^{-3}\ \lambda/D_{\mathrm{sky}}$ rms overall disturbance level. This $10^{-3}\ \lambda/D_{\mathrm{sky}}$ level is smaller than that shown in Fig. 7 (taken under vacuum with no disturbances added) without removing drift ($2\times10^{-3}\ \lambda/D_{\mathrm{sky}}$ rms over 2 hours) but larger than that after removing drift ($0.3\times10^{-3}\ \lambda/D_{\mathrm{sky}}$ rms over 2 hours). If a disturbance sequence is repeated on a timescale over which the drift is correlated (at least 12 hours, see Fig. 7), the drift should largely vanish from to the $10^{-3}\ \lambda/D_{\mathrm{sky}}$ repeatability measurement, implying that the $10^{-3}\ \lambda/D_{\mathrm{sky}}$ repeatability has a small contribution from the underlying vacuum disturbances (which, of course, do not repeat).

4. MILESTONE PROCEDURES

The success criteria for the milestone are those defined in the white paper, repeated here, with only the section numbers changed to match the numbering in this document.

4.1. The duration of the experiment for each of the measurements should be at least 1 hr and correspond to at least 1000 CLOWFS samples. Both constraints must be satisfied: the measurement should therefore last more than 1 hr if the frame rate is insufficient to obtain 1000 samples within 1 hr, and the measurement will consist of more than 1000 samples if the frame rate is sufficiently fast to obtain the 1000 samples in less than 1 hr.

Rationale: *This ensures that a sufficiently large number of measurement points are used toward computing the jitter and calibration residuals.*

4.2. A residual jitter of 0.01 λ/D or smaller shall be achieved as defined in Sec. 3.5.7. This requirement is on the radial jitter value: the quadratic sum of the jitter in x and y must be 0.01 λ/D or smaller.

Rationale: *This provides evidence that the coronagraph leak due to tip-tilt errors is smaller than the contribution of stellar angular size.*

4.3. A calibration residual of 0.003 λ/D or smaller shall be achieved as defined in Sec. 3.5.7. This requirement is on the radial calibration residual value: the quadratic sum of the calibration residual in x and y must be 0.003 λ/D or smaller.

Rationale: *This ensures that calibration of tip-tilt errors is sufficient to recover planets 10 times fainter than the 1×10^{-9} contrast raw scattered light due to stellar angular size, zodiacal light and exozodiacal light.*

4.4. Elements 4.1 – 4.3 must be satisfied on three separate occasions with a reset of the CLOWFS (discarding calibration and re-calibrating the CLOWFS) between each demonstration.

Rationale: *This provides evidence of the repeatability of the contrast demonstration. The CLOWFS reset between data sets ensures that the three data sets can be considered as independent and do not represent an unusually good configuration that cannot be reproduced. There is no required interval between demonstrations; subsequent demonstrations can begin as soon as prior demonstrations have ended.*

5. NARRATIVE REPORT

The milestone data were recorded on three runs between March 22 and 23, 2012, all under vacuum. Each run comprised four parts: a calibration of x and y actuator motions relative to CLOWFS images, a 1250-iteration sequence with the correction loop operating and a predetermined sequence of disturbances applied, a 1250-iteration sequence with no correction and the same sequence of disturbances applied, and a 625-iteration sequence of uncorrected measurements of small-amplitude sinusoids to verify sensitivity thresholds. After the calibration, all CLOWFS data are acquired at a rate of 1 iteration per 7 s, so 1250 iterations lasts nearly 2.5 hours, and the entire set of 1250 + 1250 + 625 iterations lasted just over 6 hours per milestone run. A schematic representation of the activities in each milestone run, and the statistics produced, is shown in Fig. 9.

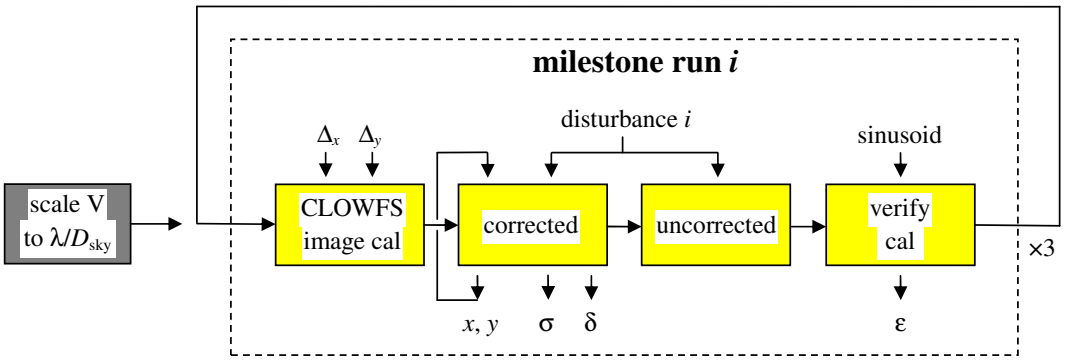


Fig. 9. Schematic showing sequence of calibration and measurement activities, and the production of milestone statistics. The first scaling step (in gray) happens only once. The milestone runs are repeated three times, producing a value of σ, δ, and ε for each run. Each milestone run has a different sequence of disturbances, but each sequence is repeated twice in a single milestone run, once applying corrections and once with no corrections applied.

The choice of 1250 iterations per corrected milestone sequence contains a margin in excess of the requirement of 1000 iterations, and at least 1 hour. Since the disturbance power is dominated by low temporal frequencies, longer time intervals should present more challenging conditions, so the choice of 1250 iterations rather than the minimum 1000 iterations does not bias the results to lower residuals.

The characteristics of the three milestone data runs are very similar. The disturbances added were different realizations of the same underlying temporal power spectrum, so while the specific time sequences of the disturbances differ, their power spectra are all, by construction, very similar.

5.1. Correction requirement

The corrected and uncorrected *x* and *y* measurements for the three milestone runs are shown in Fig. 10. Because the corrected measurements are so much smaller than the uncorrected measurements, Fig. 11 contains only the corrected measurements, with the vertical axis scaled appropriately. The time axes on each of these plots should be interpreted as time since the start of each disturbance sequence, and imply synchronicity in the applied disturbances between the uncorrected and corrected sequences, but the corrected and uncorrected data were not concurrent. To restate, one entire corrected sequence is obtained and plotted versus time, then the disturbances are repeated and the uncorrected sequence is obtained and overlaid on the same plot. At any given time *t* in one plot, the uncorrected and corrected data points at that time had the same disturbance applied, but the uncorrected data point was acquired 2.5 hours after the corrected data point, for example.

The correction requirement of the milestone states that the rms total motion (quadrature sum of *x* and *y* measurements) after correction should be less than $10\times10^{-3}\ \lambda/D_{sky}$. The uncorrected sequences show disturbances at approximately 9× this value, specifically $\sigma = 87,\ 87,$ and $95\times10^{-3}\ \lambda/D_{sky}$. For the corrected sequences, the rms total motion was $\sigma = 1.07,\ 1.10,$ and $1.05\times10^{-3}\ \lambda/D_{sky}$, more than a factor of 9x better than the milestone requirement.

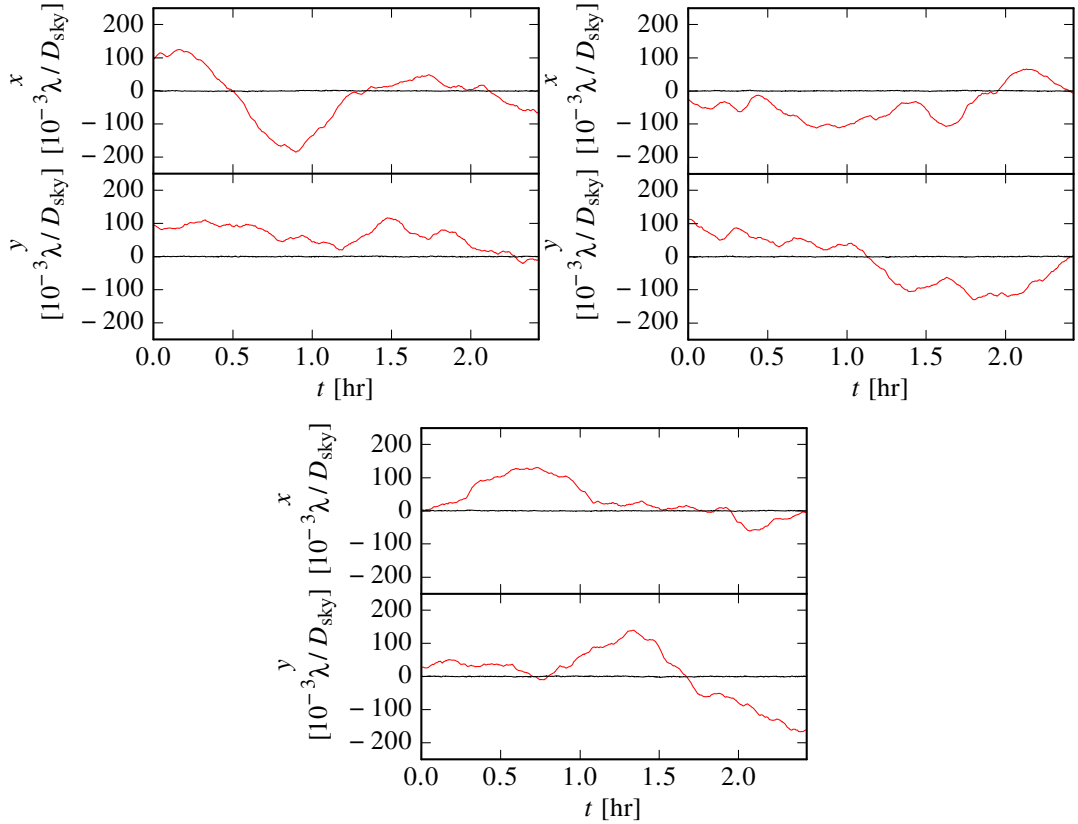


Fig. 10. (TOP LEFT) Milestone 2 CLOWFS data run #1, (TOP RIGHT) Milestone 2 CLOWFS data run #2, (BOTTOM) Milestone 2 CLOWFS data run #3. All runs are performed in vacuum, with injected disturbances. Each plot is an overlay of two sequences experiencing the same disturbances, with black lines representing corrected measurements, and red lines representing uncorrected measurements. Iterations are separated by 7 s, and each sequence comprises 1250 iterations (*i.e.*, 1250 corrected iterations and 1250 uncorrected iterations). The quadrature sums of x and y standard deviations are 87, 87, and 95 $\times 10^{-3}$ λ/D_{sky} for the uncorrected data in the three runs.

No particular attention was paid to the quality of the match between the disturbances introduced and the corresponding uncorrected measurements. The only point that is relevant for this demonstration is the power spectral density of the uncorrected measurements versus the corrected measurements, and the repeatability of the disturbances between the two sequences. For comparison, a representative 2.5-hour interval from the left panel of Fig. 7 (uncorrected measurements with no disturbances) would be dominated by the linear drift term, which would contribute approximately $8{\times}10^{-3}$ λ/D_{sky} to the uncorrected measurements. This number is small compared to a representative uncorrected peak-to-valley measurement of ~ $200{\times}10^{-3}$ λ/D_{sky}.

The power spectral densities of the uncorrected and corrected sequences are shown in Fig. 12. The differences in the power spectra between the three uncorrected runs are small relative to the differences

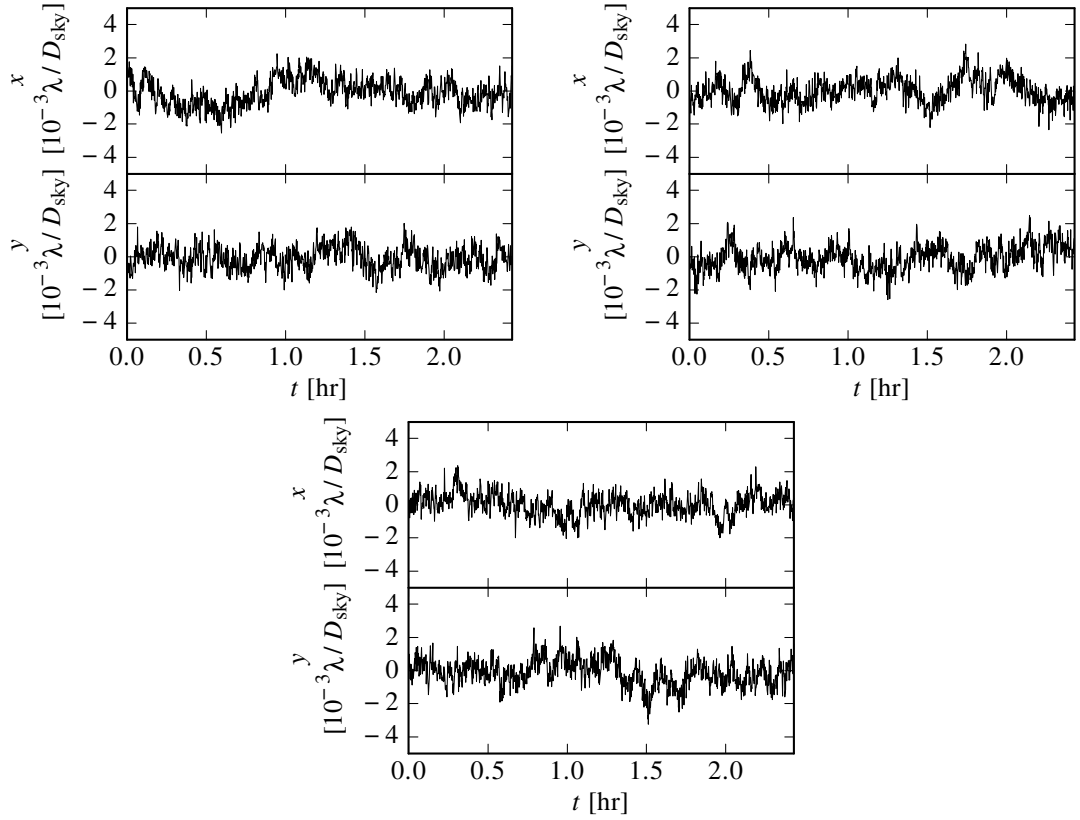


Fig. 11. Same data in each panel as in Fig. 10, but only corrected CLOWFS data are shown, with vertical scale adjusted by 50× relative to Fig. 10. The quadrature sums of x and y rms residuals are 1.07, 1.10, and 1.05 $\times 10^{-3}\ \lambda/D_{\rm sky}$ for the three runs, well under the Milestone 2 requirement of $10\times 10^{-3}\ \lambda/D_{\rm sky}$.

between uncorrected and corrected sequences. None of the measurements reaches the asymptotic white noise level seen in Fig. 8, at $0.5\times(10^{-3}\ \lambda/D_{\rm sky})^2$ / Hz, justifying the statement that measurement noise plays no significant role in the residual tip-tilt measurements.

For a given milestone run (*i.e.*, given identical disturbances), the uncorrected and corrected power spectral densities should be related by the square of the error transfer function of the CLOWFS system, as plotted in Fig. 13. The unity-gain frequency should be $f_1 = 1/(2\pi\times 7\ \text{s}) = 0.023$ Hz, and the square root of the ratio of corrected to uncorrected power should be approximately (f/f_1) for $f < f_1$. More accurately, the error transfer function at all frequencies is calculated numerically for a 1.2 s integration of images on a 7 s iteration period, and delays smaller than (7 s – 1.2 s =) 5.8 s. The match between the measured transfer functions and the expected transfer function (the colored lines compared to the black dashed line in Fig. 13) is quite good.

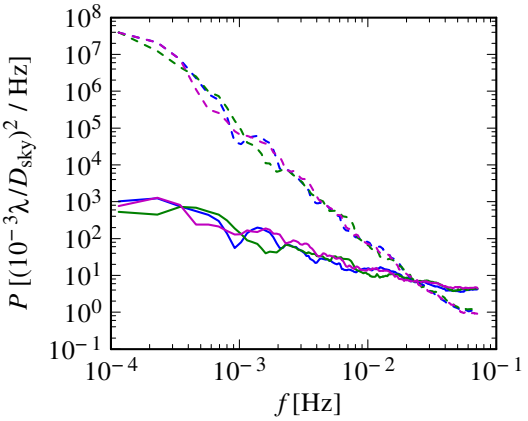


Fig. 12. Power spectral densities of Milestone 2 data runs, same data as in Figs. 10 and 11 with x and y powers summed. Solid lines are corrected and dashed lines are uncorrected. Blue is run #1, green is run #2, and magenta is run #3. The closed-loop unity-gain frequency of the correction with iterations separated by 7 s is expected *a priori* to be $1/(2\pi\times 7\ {\rm s}) = 0.023$ Hz, in excellent agreement with the measurements. The plot is smoothed as described in Appendix 1.

5.2. Calibration requirement

The calibration requirement takes two forms, first as a measurement of the mean residual after correction, and second as a measurement of uncorrected sinusoidal inputs.

The first requirement is that the magnitudes of mean x and mean y corrected measurements be less than

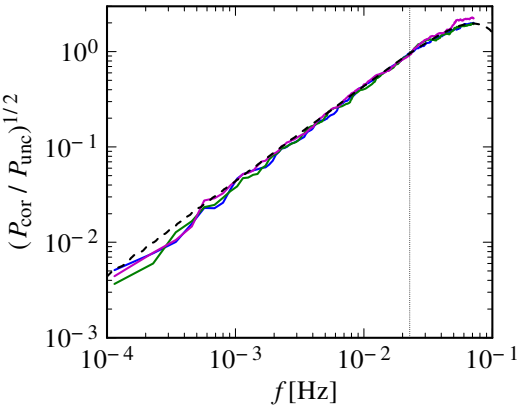


Fig. 13. Error transfer function. The three colored lines correspond to the square roots of the ratios of the colored lines in Fig. 12, the corrected power (solid lines in Fig. 12) divided by uncorrected power (dashed lines in Fig. 12), each smoothed as described in Appendix 1. The black vertical dotted line marks f_1, where f_1 is the closed-loop unity-gain frequency. The black dashed line is a numeric expectation of the transfer function, for a loop operating with 1.2 s integrations per 7 s iteration.

$3\times10^{-3}\ \lambda/D_{sky}$. The means of the corrected x and y measurements from the three runs (the data in Fig. 11) are $(\delta_x,\delta_y) = (-0.11,\ -0.09)$, $(0.01, -0.08)$, and $(0.00, -0.13)\times10^{-3}\ \lambda/D_{sky}$, giving rms quadrature sums of x and y means of $\delta = 0.14$, 0.08, and $0.13\times10^{-3}\ \lambda/D_{sky}$, all well below the $3\times10^{-3}\ \lambda/D_{sky}$ requirement. These are tabulated and summarized in the conclusions in Sec. 6.

The second requirement is that the amplitude of sinusoidal input signals be measured with errors below $3\times10^{-3}\ \lambda/D_{sky}$. This test is performed using the same technique as the uncorrected measurement of disturbances described above, except that the disturbances are not a sequence of random variates selected from a nominal power spectral density, but are simply sinusoids. The amplitudes of the sinusoids were set at levels comparable to the corrected residuals. The x and y amplitudes were chosen based on equal voltages (70 mV), which come to slightly different λ/D_{sky} amplitudes (because of different piezo gains), $1.81\times10^{-3}\ \lambda/D_{sky}$ for x and $1.91\times10^{-3}\ \lambda/D_{sky}$ for y. Converting these sinusoidal amplitudes to rms amplitudes and adding them in quadrature yields $1.86\times10^{-3}\ \lambda/D_{sky}$, about 75% larger than the rms measured corrected summed residuals σ of 1.07, 1.10, and $1.05\times10^{-3}\ \lambda/D_{sky}$. The two sinusoids are applied simultaneously, with a 56 s period in x and 70 s period in y.

The calibration test is analyzed through the power spectral densities of the uncorrected measurements. These power spectral densities are shown in Fig. 14. These PSDs are calculated without the Nuttall window used in the other PSD analyses in this milestone, but rather with a "natural" equal-weighted (rectangular) window, to retain high spectral resolution. The data are not drift-subtracted as was done for Fig. 7, so the continuum levels are dominated at low frequencies by this (directional) drift term, but this continuum level is negligible at the sinusoidal input frequencies. The peaks in the two PSDs are large enough compared to the background power levels ($>\times10^3$) that no subtraction is performed, the peak

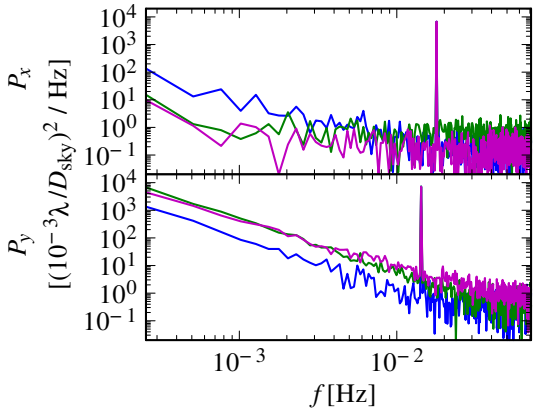


Fig. 14. Power spectral densities of Milestone 2 calibration verification runs. Blue is run #1, green is run #2, and magenta is run #3. The x and y actuators were driven with sinusoids of amplitudes $1.81\times10^{-3}\ \lambda/D_{sky}$ and $1.91\times10^{-3}\ \lambda/D_{sky}$, with periods of 56 and 70 s, respectively, in vacuum with no other disturbances added, and no corrections performed. Each peak has 10^3-$10^4\times$ the PSD of the continuum at nearby frequencies, so no attempt was made to subtract any "background" power. The rms difference between input and measured amplitudes is 1.7% for the 6 peak measurements (3 runs, 2 axes). This plot is not smoothed.

levels are simply integrated over one frequency bin to convert them into amplitudes. The x and y fractional errors in the measured amplitudes with respect to the amplitudes actually applied, for the 3 runs were (2.6%, 0.4%), (0.6%, 1.3%), and (2.6%, -0.7%). The rms quadrature sum of x and y for the three runs was 2.3%. The milestone calibration requirement is that the calibration error be below $3\times10^{-3}\ \lambda/D_{sky}$, while this experiment measures, for signals of comparable (slightly larger) amplitudes than on average, $\varepsilon = 0.04\times10^{-3}\ \lambda/D_{sky}$, nearly two orders of magnitude below the requirement, for input levels higher than the corrected residuals.

6. CONCLUSIONS

This document describes the completion of the PIAA Milestone #2 requirements, demonstrating tip-tilt control at levels relevant to flight missions. The requirements are that tip-tilt be controlled to $10\times10^{-3}\ \lambda/D_{sky}$ rms over 1000 iterations, and that the calibration errors be below $3\times10^{-3}\ \lambda/D_{sky}$. These requirements must be met on each of three separate runs.

The results, as described in detail above, can be summarized as follows:

	correction requirement			calibration requirement					
			Require < 10			Require < 3			Require < 3
Run	σ_x	σ_y	σ	δ_x	δ_y	δ	ε_x	ε_y	ε
1	0.82	0.71	1.08	−0.11	−0.09	0.14	0.05	0.01	0.05
2	0.77	0.78	1.10	0.01	−0.08	0.08	0.01	0.02	0.03
3	0.69	0.80	1.06	0.00	−0.13	0.13	0.05	−0.01	0.05

Table 1. Milestone requirements and observed performance. The units for all measurements are $10^{-3}\ \lambda/D_{sky}$. The σ measurements are rms residuals after correction, δ are mean residuals, and ε are errors in sinusoidal measurements.

These results perform better than the milestone requirements by a large margin, between 9× and 100×. For the correction and the mean residuals, this is in the presence of disturbances at rms levels near $90\times10^{-3}\ \lambda/D_{sky}$. The uncertainty in the scale factor determination, which is a systematic multiplication of every measured value presented in terms of λ/D_{sky}, is estimated at 3.6% (see Appendix 2), which introduces negligible changes to the interpretations of the margins by which these measurements exceeded the requirements. We claim that the tip-tilt milestone requirements have been met, with a substantial margin.

The CLOWFS, as implemented in this milestone report, is compatible with any coronagraph that uses an opaque focal plane mask to block starlight. This includes conventional Lyot type coronagraphs and band-limited Lyot coronagraphs, as well as most PIAA coronagraphs. The results described in this paper can be generalized to these coronagraphs.

A number of other coronagraphs use phase-shifting (non-absorbing) focal plane masks, such as the Vortex coronagraph, the 4 quadrant or 8 octant coronagraphs, and the PIAACMC. For these coronagraphs, the CLOWFS implementation is different: the Lyot mask after the focal plane mask is a mirror with a hole, reflecting the light that is outside the geometric pupil to the CLOWFS camera. The

coronagraph focal plane is re-imaged on the CLOWFS detector, which, for a perfect alignment, shows the same PSF as lands on the occulter, *i.e.*, the light rejected by the coronagraph is sent to the CLOWFS detector. The CLOWFS image is a linear function of tip and tilt, and the CLOWFS control loop can be identical to the control loop implemented for this milestone.

While this alternate "Lyot pickoff" implementation of the CLOWFS has been numerically simulated and is expected to yield the same precision as the implementation presented in this paper, it has never been tested in the laboratory. The results obtained for this milestone therefore cannot be considered to be applicable to these phase mask coronagraphs - experimental validation of the CLOWFS is thus needed for these coronagraphs.

The specific disturbances chosen to demonstrate the performance of the CLOWFS system do not have general significance, but the demonstrated error transfer function, as shown in Fig. 13, should be generally realizable for similarly constructed CLOWFS systems. This generality, at least for $f < f_1$, would hold provided that the system is adequately represented in a linear analysis. As described in Ref. 11 (e.g., Fig. 5 in Ref. 11), the level of nonlinearities in the CLOWFS response (due mostly to the annular occulter geometry) is expected to reach approximately 20% at displacements of $200\times10^{-3}\ \lambda/D_{\mathrm{sky}}$. The closed-loop performance demonstrated in this report is well below those displacements (by $> 10^2\times$), but these results should not be interpreted as validating performance in regimes where this constraint is violated.

Appendix 1. Power Spectral Densities

The calculations of power spectral densities in this experiment all begin with a temporal sequence of n real-valued measurements, taken at time intervals separated by Δt. For notation in this appendix, the measurements are denoted m_i with i ranging from 0 to $n-1$. By construction, this is a discrete sequence, and so discrete transforms will be used in the analysis.

When taking the Fourier Transform of the sequence m, one of two windows is applied. In cases where spectral resolution is less important than dynamic range (low sidelobes), a Nuttall 4-term continuous first derivative window [14] is applied,

$$w_i = a_0 + a_1\cos(2\pi i/n) + a_2\cos(4\pi i/n) + a_3\cos(6\pi i/n)$$
$$a_0=0.355768,\ a_1=-0.487396,\ a_2=0.144232,\ a_3=-0.012604$$

In cases where spectral resolution is more important than dynamic range, a rectangular (or top-hat) filter is applied,

$$w_i = 1$$

w_i is dimensionless in all cases.

In practice, all power spectral densities described in this document use the Nutall window, except for the measurement of sinusoidal amplitudes to satisfy the calibration milestone requirements, as used in Fig. 14, which uses the rectangular window.

The power spectral densities reported here are one-sided power (i.e., undefined for negative frequencies), and the zero-frequency result is not meaningful in any of these analyses, so for positive frequencies,

$$P_k = 2\Delta t\ |\Sigma_j\ m_j\ w_j\ e^{2\pi ijk/n}|^2\ /\ (n\ \Sigma_j\ w_j^2),$$

for any k from 1 to $n/2-1$. The sums $\Sigma_j\ldots$ are taken for $j=0$ to $j=n-1$, and P_k is interpreted as the power spectral density at frequency $f = k/(n\Delta t)$, and has dimensions of $[m]^2$ / Hz, where $[m]$ is the dimension of the measurements m_i.

All power spectral densities shown in this document are plotted on a log-log scale, and for every case except Fig. 14 (Fig. 14 contains the only PSD using a rectangular window, where spectral resolution is a priority), the plot shows the individual powers smoothed by a rolling average of fractional frequency width,

$$P_{k,\text{plot}} = \langle P_j \rangle \text{ for } 0.8k < j < 1.2k,$$

where $\langle\ldots\rangle$ denotes a mean. At high frequencies, this reduces the noise inherent in the power spectral density calculations.

Appendix 2. Error Estimates

The errors relevant to this experiment are readily categorized as random and systematic errors. The random noise is estimated from the under-vacuum PSD with no corrections and no disturbances added (Fig. 8), while the systematic errors are estimated by the chain of errors in the individual components of the scaling constants S_x and S_y.

A2.1. Random errors

An estimate of the random errors comes from the shape of the power spectral density of uncorrected, undisturbed under-vacuum measurements as shown in Fig. 8. The high-frequency end of the PSD appears to be approaching an asymptotic value, which would be true for a white-noise component, as would be expected for any errors that are uncorrelated between CLOWFS images. These errors would most likely be shot noise in the signal and read noise in the detector. This random noise floor, corresponds approximately to $0.2\times10^{-3}\ \lambda/D_{\text{sky}}$ rms, uncorrelated between measurements. This number is comfortably small so that further precision in the noise analysis is not warranted.

A2.2. Systematic errors

The dominant systematic noise source is in the determination of the scale factors S_x and S_y relating volts to λ/D_{sky}, as shown in Figs. 3 and 9, and described in detail in Secs. 3.5.2-3.5.4. This scale factor determination is performed once for the entire experiment, so errors in its determination will produce a uniform multiplicative error to every measurement that is later expressed in λ/D_{sky}.

A2.2.1. Errors in $S_{1,x}$ and $S_{1,y}$

The errors in $S_{1,x}$ and $S_{1,y}$ can be estimated by the rms residuals (bottom of Fig. 5), which are 0.02 pix rms (similar for each of x and y). A simplistic treatment (assuming uncorrelated errors) would say that the amplitude measurements (top of Fig. 5) of ~ 0.8 pix, over 625 samples, would see a contribution of $0.02\text{ pix}/(625)^{1/2} \sim 10^{-3}$ pix, for a relative error on the order of 0.1%. To be clear, this is the contribution of random errors, assumed uncorrelated, to the measurement of the centroid sinusoidal amplitude at the science camera. This relative error is much smaller than those of S_2 and S_3.

A2.2.2. Errors in S_2

The error in S_2, from camera pixels to microns at the occulter, is estimated from the results of moving the occulter and PSF together, and equating the pixel motion on the camera with the encoder motion at the occulter. The occulter encoder registered a 67.5 micron move, while the science camera image moved by 12.9 pixels, giving a measurement of 5.23 microns/pix. This differs by 0.6% from the nominal values in Sec. 3.5.3 of 5.2 microns/pix, which was used in the analysis. The accuracies of the two numbers involved were 0.25 microns at the occulter and 0.1 pix at the camera for each of the two locations that were differenced. An overly conservative treatment would say that an uncertainty of 0.5 microns (twice the individual error) out of 67.5, and 0.2 pix out of 12.9 together give a 1.7% relative uncertainty. The difference between the measured and assumed values is well within this (0.6%), but a conservative treatment would assume that 1.7% is an upper limit to the uncertainty in this step of the scaling.

A2.2.3. Errors in S_3

The estimation of errors in S_3 is broken into two different measurements. The measurement of f/D at the science camera is done by establishing the distance from the source image plane to the pupil image plane using the encoders on the camera stage, 228.5 mm, which itself needs no error bars because the pupil diameter measurement scales with errors in this number. The pupil diameter is 427 pix ± 4 pix (it has an irregular edge), for a relative uncertainty of 1%. Using the specified CCD pixel size of 13 μm, and the laser wavelength of 808 nm, $f\lambda/D$ = 33.2 μm at the camera. The magnification from the camera to the occulter is 2.5 (as implicitly verified in the second step of the scale factor determination, described in the previous paragraph), giving $f\lambda/D$ = 13.3 μm at the occulter, with an uncertainty of approximately 1%. The additional measurement needed to verify the relationship between microns at the occulter and λ/D_{sky} is the PIAA off-axis magnification M. One unresolved ambiguity in the available measurements is the magnification in the absence of the PIAA optics, between the source and the occulter. Without having independent measure of this, a mismatch in the expected 1:1 match in input and output focal lengths of the M1 and M2 PIAA mirrors would be interpreted here as a change in off-axis magnification. Without any reason to believe that there would be an anticorrelation between these errors, attributing all of the measured error to only the off-axis magnification should be a conservative treatment. Referencing all measurements to distances at the occulter (again using the prior scale factors), the relative mismatch between the expected and measured off-axis magnification is 2.7%. Coupled with the 1% uncertainty in the $f\lambda/D$ determination, the total uncertainty in the third step is 2.9%.

A2.2.4. Total systematic errors

Taking the systematic uncertainties of S_1, S_2, and S_3 (of which the errors in S_1 are negligible) together in quadrature, the total uncertainty in the scale factor S_x and S_y determination is 3.4% (approximately the same in both axes). A summary of these errors appears in Table 2.

	nominal value	measured value	uncertainty	origin of uncertainty
$S_{1,x}$ and $S_{1,y}$	use measured	(0.80, 0.84) pix / 5 V	0.1%	residuals to fit
S_2	5.21 μm/pix	5.23 μm/pix	1.7%	accuracy of meas. values
S_3	(λ/D_{sky}) / (33.3 μm)	(λ/D_{sky}) / (33.4 μm)	2.9%	accuracy of meas. values

Table 2. Scale factor values, both nominal and measured, and their estimated uncertainties.

REFERENCES

[1] Guyon, O. 2003, "Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging," A&A **404**, 379-387.

[2] Traub, W.A.; Vanderbei, R.J. 2003, "Two-Mirror Apodization for High-Contrast Imaging," ApJ, **599**, 695-701.

[3] Vanderbei, R.J.; Traub, W.A. 2005, "Pupil Mapping in Two Dimensions for High-Contrast Imaging," ApJ **626**, 1079-1090.

[4] Guyon, O., Pluzhnik, E.A., Galicher, R., Martinache, F., Ridgway, S.T., Woodruff, R.A. 2005, "Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronagraph. I. Principle," ApJ **622**, 744-758.

[5] Martinache, F., Guyon, O., Pluzhnik, E.A., Galicher, R., Ridgway, S.T. 2006, "Exoplanet Imaging with a Phase-Induced Amplitude Apodization Coronagraph. II. Performance," ApJ **639**, 1129-1137.

[6] Vanderbei, R.J. 2006, "Diffraction Analysis of Two-dimensional Pupil Mapping for High-Contrast Imaging," ApJ **636**, 528.

[7] Pluzhnik, E.A., Guyon, O., Ridgway, S.T., Martinache, F., Woodruff, R.A., Blain, C., Galicher, R. 2006, "Exoplanet Imaging with a Phase-Induced Amplitude Apodization Coronagraph. III. Diffraction Effects and Coronagraph Design," ApJ **644**, 1246-1257.

[8] Guyon, O., Pluzhnik, E.A., Kuchner, M.J., Collins, B., Ridgway, S.T. 2006, "Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs", ApJSS **167**, 81-99.

[9] Lozi, J., Martinache, F., Guyon, O. 2009, "Phase-Induced Amplitude Apodization on centrally obscured pupils: design and first laboratory demonstration for the Subaru Telescope pupil," PASP **121**, 1232-1244.

[10] Guyon, O., Martinache, F., Belikov, R., Soummer, R. 2010, "High Performance PIAA Coronagraphy with Complex Amplitude Focal Plane Masks," ApJ **192**, 220-232.

[11] Guyon, O., Matsuo, T., Angel, R. 2009, "Coronagraphic Low Order Wavefront Sensor: Principle and Application to a Phase-Induced Amplitude Coronagraph," ApJ **693**, 75-84.

[12] Kern, B., Give'on, A., Kuhnert, A., Levine-West, M.B., McMichael, I., Moody, D.C., Niessner, A.F., Pueyo, L., Shaklan, S.B., Traub, W.A., Trauger, J.T., Belikov, R., Guyon, O. 2009, "Phase-induced amplitude apodization (PIAA) coronagraph testing at the High Contrast Imaging Testbed," Proc. SPIE **7440**, 74400H.

[13] Kern, B., Give'on, A., Kuhnert, A., Niessner, A., Guyon, O. 2011, "Laboratory testing of a Phase-Induced Amplitude Apodization (PIAA) coronagraph," Proc. SPIE **8151**, 815104.

[14] Nuttall, A.H. 1981, "Some Windows with Very Good Sidelobe Behavior," IEEE Trans. Acoust., Speech, Signal Process., **ASSP-29**, 84-91.

CLOWFS pour le traitement a posteriori

L'INFORMATION PRODUITE PAR CLOWFS PENDANT LA PHASE D'ÉTALONNAGE peut, en effet, en plus du contrôle en temps réel, être utilisée pour améliorer la qualité de l'étalonnage des images produites par le coronographe, lors du traitement *a posteriori* des données. Une stratégie classique d'observation consiste, avant ou après acquisition d'images sur une source d'intérêt, à pointer une source pouvant jouer le rôle de calibrateur : un objet ponctuel, de type spectral et de magnitude semblable à la source d'intérêt, pas trop éloigné de cette dernière, de sorte que les conditions d'observation (*airmass*, flexions mécaniques du télescope, performance du système d'OA, illumination du détecteur scientifique) n'évoluent peu ou pas pendant le passage entre les deux objets. Cette pratique d'observation se retrouve également en interférométrie NRM (voir chapitre précédent) ou longue base, pour mieux estimer à un instant donné, l'impact de l'infrastructure sur la mesure des modules de visibilité.

SI DES IMAGES CORONOGRAPHIQUES ACQUISES SUR UN TEL CALIBRATEUR SONT DISPONIBLES, elles peuvent être soustraites de l'image acquise sur la source d'intérêt. Les biais systématiques présents dans l'image de départ, tels les speckles quasi-statiques et les résidus de diffraction, peuvent en théorie être éliminés. L'ensemble de ces biais peuvent être désignés par le terme général de "fuite coronographique". La procédure de soustraction de PSF coronographique, permet en théorie de révéler le contenu de l'image correspondant aux sources incohérentes avec l'étoile centrale comme les planètes ou les disques. Il y a cependant une limite à la fidélité des conditions d'observation entre celle sur la source d'intérêt et la source de calibration, qui limite le potentiel de cette procédure.

DES ACQUISITIONS SIMULTANÉES par la caméra scientifique et la caméra CLOWFS pendant l'observation de la source de calibration peuvent *booster* le potentiel de cette procédure. L'idée est simple : il s'agit de construire un dictionnaire ou table de correspondance entre l'image CLOWFS instantanée et la distribution des fuites coronographiques dans l'image focale collectée par l'observateur. La télémétrie des résidus en boucle ouverte ou en boucle fermée, produite par CLOWFS pendant l'observation sur la source d'intérêt, peut, une fois combinée avec les informations contenues dans le dictionnaire, alors prédire la distribution des fuites coronographiques à soustraire de l'image. La construction de cette image étalon est illustrée dans la figure 34.

La figure 35 illustre le potentiel de cette procédure, en s'intéressant à une série d'images produites par SCExAO en laboratoire, en 2011. Pour cette expérience, les acquisitions par les caméras focales et CLOWFS étaient synchronisées par un shutter commun placé en amont dans l'instrument. Les vibrations induites par une pompe

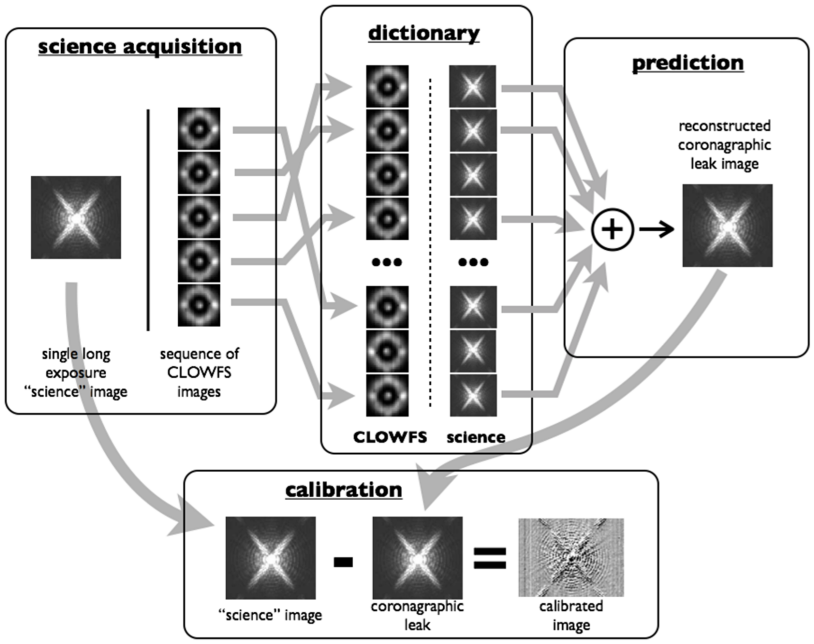


FIGURE 34: Traitement *a posteriori* des images coronographiques. Un dictionnaire d'images (CLOWFS / science) acquis pendant l'observation d'une source de référence est utilisé pour prédire le terme de fuite coronographique affectant une image acquise sur une source d'intérêt astrophysique.

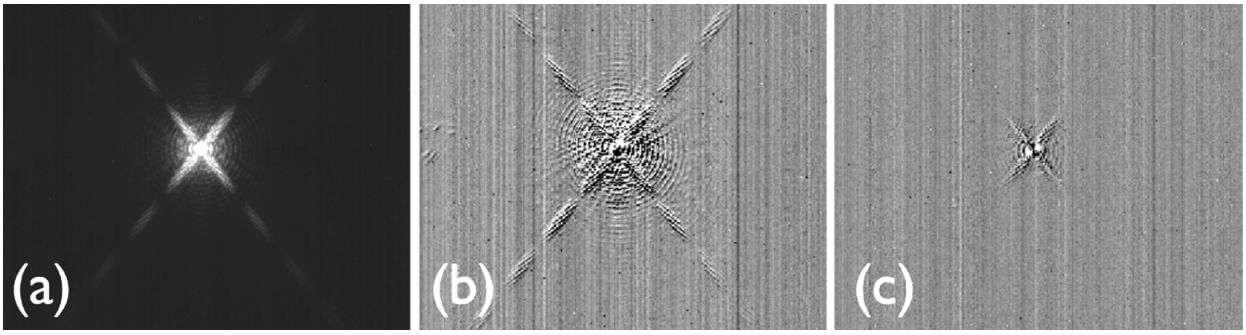


FIGURE 35: Performance de l'algorithme de traitement a posteriori utilisant le dictionnaire d'images introduit plus haut. De gauche à droite, on visualise : (a) l'image focale brute, (b) cette même image dont a été soustraite une PSF de référence acquise par une méthode classique et (c) cette même image, après soustraction d'une PSF construite grâce à la télémétrie CLOWFS.

d'aquarium, vissée à la table optique introduisaient majoritairement du tip-tilt, se traduisant par des fuites coronographiques.

Lors d'observations au télescope, il est habituel d'intercaler entre deux pointages sur des objets d'intérêt, des observations sur des objets dits de calibration, sans structure apparente et qu'on tente d'observer dans des conditions les plus proches de celles qui régnaient lors de l'acquisition sur l'objet d'intérêt. Dans des observations en optique adaptative, cette acquisition d'un objet de calibration permet d'estimer (même s'il s'agit d'un abus de language) ce qu'on va appeler "la PSF coronographique". Avec un étalonnage idéal, cette PSF coronographique après mise à l'échelle, peut être soustraite à l'image acquise sur l'objet d'intérêt pour révéler les structures faibles, initialement dominées par les résidus de diffraction.

La figure 35 compare la performance d'une telle calibration à celle décrite dans cette partie du manuscrit. La PSF coronographique est construite en choisissant aléatoirement dans la série complète d'images, un nombre suffisant d'images co-additionnées. Le résultat de cette procédure de calibration dite "classique" est visible dans le panneau (b) de la figure. Il est à comparer au résultat de la procédure exploitant la connaissance du dictionnnaire d'images, visible dans le panneau (c).

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **123**:1434–1441, 2011 December
© 2011. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

Coronagraphic Low-Order Wavefront Sensor: Postprocessing Sensitivity Enhancer for High-Performance Coronagraphs

FRÉDÉRIC P. A. VOGT,[1] FRANTZ MARTINACHE, OLIVIER GUYON, TAKASHI YOSHIKAWA, KAITO YOKOCHI, VINCENT GARREL, AND TARO MATSUO

National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720; frantz@naoj.org

Received 2011 March 10; accepted 2011 November 7; published 2011 December 2

ABSTRACT. Detection and characterization of exoplanets by direct imaging requires a coronagraph designed to deliver high-contrast at small angular separation. To achieve this, an accurate control of low-order aberrations, such as pointing and focus errors, is essential to optimize coronagraphic rejection and avoid the possible confusion between exoplanet light and coronagraphic leaks in the science image. Simulations and laboratory prototyping have shown that a coronagraphic low-order wavefront sensor (CLOWFS), using a single defocused image of a reflective focal-plane ring, can be used to control tip-tilt to an accuracy of $10^{-3}\ \lambda/D$. This article demonstrates that the data acquired by CLOWFS can also be used in postprocessing to calibrate residual coronagraphic leaks from the science image. Using both the CLOWFS camera and the science camera in the system, we quantify the accuracy of the method and its ability to successfully remove light due to low-order errors from the science image. We also report the implementation and performance of the CLOWFS on the Subaru Coronagraphic Extreme-AO (SCExAO) system and its expected on-sky performance. In the laboratory, with a level of disturbance similar to that encountered in a post-AO beam, CLOWFS postprocessing has achieved speckle calibration to $1/300$ of the raw speckle level. This is about 40 times better than could be done with an idealized PSF subtraction that does not rely on CLOWFS.

1. INTRODUCTION

Since the discovery of 51 Pegasi by Mayor & Queloz (1995), several hundred extrasolar planets have been uncovered, for the most part via indirect detection methods such as radial velocity measurements of the reflex motion of their host star and photometric transit. Direct imaging was recently able to produce the first high-contrast images of *solid* planetary candidates orbiting nearby main-sequence stars: Fomalhaut (Kalas et al. 2008), Beta Pictoris (Lagrange et al. 2009), and HR 8799 (Marois et al. 2008; Lafrenière et al. 2009). Following up on these direct detections, Janson et al. (2010) obtained the first direct extrasolar planet spectrum, opening the way to better characterization of planetary atmospheres. These direct detections are currently limited to planets at large orbital separations of several dozens of AU and only probe a small fraction of the distribution of known extrasolar planets, for which the median orbital separation is likely closer to 1 AU.

At a near-infrared wavelength, an 8 m class telescope provides sufficient angular resolution (40 mas at $\lambda = \sim 1.6\ \mu$m) to be able to detect companions in the habitable zone of nearby stars ($d < 30$ pc). High-contrast imaging near the diffraction limit, however, requires very good control and calibration of the wavefront aberrations in the optical system. From the ground, this task is complicated by the rapidly changing atmospheric wavefront creating speckled images. While adaptive optics (AO) offers a major improvement, performance is still limited, and the direct imaging of planets often requires postprocessing techniques such as angular differential imaging (ADI; Marois et al. 2006), which is most efficient at large ($\gtrapprox 10\ \lambda/D$) angular separations.

In the near future, high-contrast imaging systems employing new coronagraphic and wavefront control techniques will greatly improve our ability to image exoplanets. These extreme-AO systems include the Gemini Planet Imager (GPI; Macintosh et al. 2006), ESO's Spectro-Polarimetric High-Contrast Exoplanet Research (SPHERE; Beuzit et al. 2008), and Subaru's SCExAO (Martinache & Guyon 2009). They will use efficient coronagraphs and high-speed high-order wavefront corrections to reduce speckles in the coronagraphic image. The images of the HR 8799 planetary system obtained by Serabyn et al. (2010) on a well-corrected 1.5 m aperture using a high-performance coronagraph demonstrate the relevance of this approach.

Imaging companions close to the edge of the occulting mask in a coronagraph (which is ∼40 mas on SCExAO), however, remains an unachieved feat. The current state of the art for point-spread function (PSF) calibration is an optimized version of ADI called locally optimized combination of images (LOCI), introduced by Lafrenière et al. (2007). LOCI can calibrate high-order aberrations to a very high level of contrast (12 magnitudes

[1] Current address: Mount Stromlo and Siding Spring Observatories, Research School of Astronomy and Astrophysics, The Australian National University, Cotter Road, Weston Creek, ACT 2611, Australia.

and higher), but like any technique relying on ADI, only for angular separations greater than 10 λ/D. Detection at the edge of a ~1 λ/D occulter is extremely sensitive to low-order aberrations such as pointing and focus.

Of these aberrations, pointing is especially critical, since near the occulter, a tip-tilt excursion along a given direction mimics the signal of a true companion in a coronagraphic image. This issue, first identified for a high-contrast spaceborne coronagraph, has been addressed by Guyon et al. (2009) with the coronagraphic low-order wavefront sensor (CLOWFS), a scheme using the light occulted by a modified focal-plane mask as an accurate pointing tracker.

The idea of tracking using the light that is otherwise lost in the coronagraphic focal plane is not a novelty. For instance, it was successfully implemented on the Lyot project (Digby et al. 2006). The calibration unit of GPI also uses the light occulted by the focal-plane mask to measure low-order aberrations after reimaging the pupil in a Shack-Hartmann-type wavefront sensor (Wallace et al. 2010): pointing performance with this scheme reaches 2 mas for typical expected observing conditions.

Maximum sensitivity to pointing errors is reached when the light from opposite edges of the pupil is allowed to interfere, which naturally happens in the focal plane. In this respect, while robust to a wide range of errors, a Shack-Hartmann sensor does not appear to be optimal to measure pointing: because it splits the pupil into subpupils, this capability is lost, resulting in lesser performance than a focal-plane–based wavefront-sensing scheme (Guyon 2005).

The originality of the CLOWFS design resides in its dual-zone focal-plane mask, designed to suppress a strong offset to the signal, carrying most of the power but no information, in a manner reminiscent of strioscopy. The suppression of this offset turns the otherwise imperceptible changes due to small pointing errors into a macroscopic change of the CLOWFS image. Using this scheme, Guyon et al. (2009) were able to stabilize tip-tilt at the level of 10^{-3} λ/D in a closed-loop system for $\lambda = 633$ nm, in a laboratory experiment. The current implementation of CLOWFS on SCExAO exhibits pointing residuals of less than 0.2 mas, with a 50 Hz frame rate.

While this level of performance is quite remarkable, we demonstrate in this work that additional calibration can be achieved in postprocessing and can lead to an improved subtraction of coronagraphic leaks due to low-order aberrations in a long exposure. Using the SCExAO system as a test bed, we experimentally demonstrate a 40 times improvement of the detection limit over a classical calibration procedure at angular separation of a few λ/D.

This article is organized as follows:In § 2, we introduce how to use CLOWFS for postprocessing of coronagraphic images. In § 3, we describe the implementation of the concept on the SCExAO experiment of the Subaru telescope, and we present our results in § 4. In § 5, we summarize our results and discuss possible updates to the presented CLOWFS configuration.

2. CLOWFS POSTPROCESSING PRINCIPLE

A high-performance PSF calibration procedure such as LOCI (Lafrenière et al. 2007), is based on a direct analysis of the science data alone. Yet near the edge of the coronagraphic mask, it is impossible, from such data only, to distinguish between the signal of an actual faint companion and the one of a systematic tip-tilt excursion off the coronagraphic mask at a given azimuth, which would be, for instance, due to a vibration. Data acquired with CLOWFS during a long exposure on the science camera can, however, be used to discriminate the two situations in postprocessing.

The scheme proposed in this article is a form of adaptive optics PSF reconstruction, which uses measurements acquired in a wavefront sensor to estimate the long-exposure PSF in the science camera (Veran et al. 1997; Gendron et al. 2006). Adaptive optics PSF reconstruction has been implemented on several adaptive optics systems (Harder & Chelli 2000; Jolissaint et al. 2010) and relies on the fact that the wavefront sensor, by measuring residual wavefront errors at a fixed spatial sampling, can be used to estimate the inner part of the PSF in the science camera. This estimation can be done at the wavefront-sensing sampling speed (typically 100 Hz to 1 kHz) and is then averaged for the duration of the science exposure. In this article, we reconstruct the very inner part of the coronagraphic PSF using CLOWFS telemetry. Compared with previous implementations of adaptive optics PSF reconstruction, our scheme is better suited to high-contrast coronagraphic imaging, as it uses a sensor that is highly sensitive to low-order aberrations and is free of noncommon path errors. Our PSF reconstruction is limited to the very inner part of the coronagraphic PSF, however, and is most effective in the 1 λ/D-wide area immediately around the focal-plane mask. At larger angular separation, wavefront errors can create speckles without producing a signal in the CLOWFS camera, and other calibration approaches must be used to reconstruct the PSF: for example, telemetry from a higher-order wavefront sensor, differential spectral imaging, or the ADI/LOCI technique. We also propose to use a new empirical image-based PSF reconstruction algorithm, where CLOWFS images are matched to science camera images to reconstruct the PSF, as opposed to relying on a model of the adaptive optics system. This empirical approach is more robust, simple to implement, and made possible in our case by the small number of modes measured by the CLOWFS.

A thorough description of the theory and hardware implementation of CLOWFS was provided by Guyon et al. (2009). Here, it suffices to remember that it operates due to an optimized dual-zone focal-plane mask, absorbing at its center, and is reflective in an annulus whose outer edge defines the inner working angle of the system (1.5 λ/D for SCExAO). Figure 1 shows the actual implementation of CLOWFS on SCExAO. A lens reimages the reflective ring of the occulting mask on a detector deliberately placed out of focus, referred to as the CLOWFS camera.

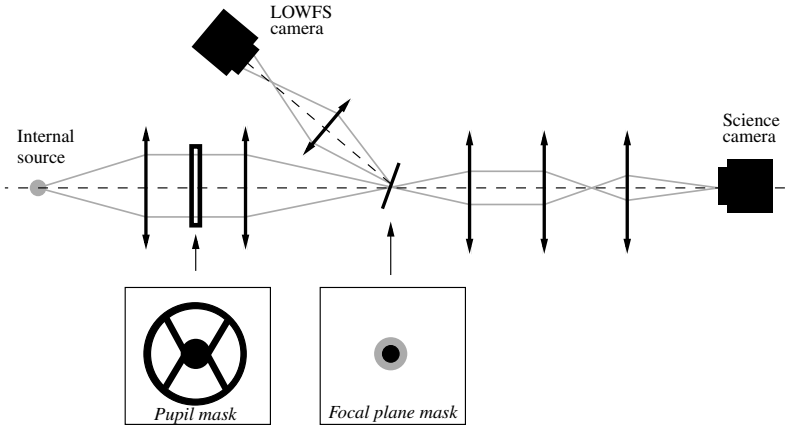


FIG. 1.—Optical layout of SCExAO used in Lyot-coronagraph mode. A mask lit by an IR ($\lambda = 1.55$ μm) laser diode emulates the Subaru Telescope pupil. The beam is focused on the dual-zone focal-plane occulting mask described in the text. The light reflected by the focal-plane mask feeds the (slightly defocused) CLOWFS detector, used to characterize pointing errors. The light that is not intercepted by the mask is then reimaged on the science detector. Examples of images obtained on both cameras are presented in Figs. 2 and 4.

During the (long) exposure on the science camera, the CLOWFS camera acquires a continuous stream of short (typically millisecond) exposures. An example of one such CLOWFS camera image is provided in Figure 2: variations of the distribution of intensity in this image are used to identify drifts in pointing, as well as changes in focus. Guyon et al. (2009) demonstrated that over a small range (~0.2 λ/D) of pointing errors, a linear model relates the changes in CLOWFS images to the actual pointing error, and they took advantage of this in a closed-loop system, stabilizing the pointing at the level of 10^{-3} λ/D over extended periods of time (~1 hr).

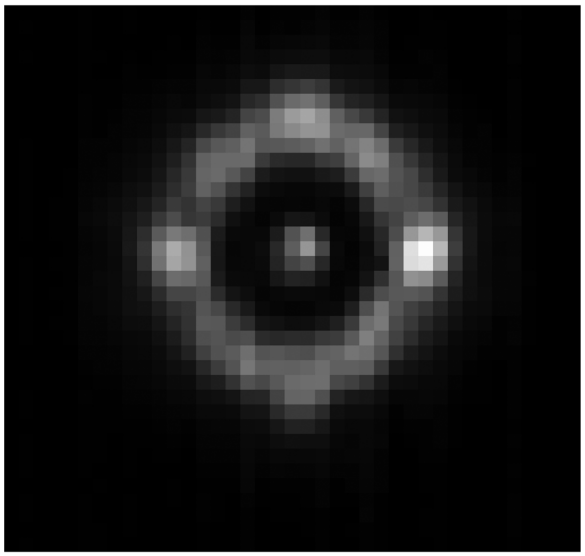


FIG. 2.—Example of CLOWFS camera image. Variations in the distribution of intensity in the image of the reflective annular region of the mask are used to track minute changes in the pointing.

Additional calibration of the coronagraphic leaks due to the residual tip-tilt error can be achieved in postprocessing. This calibration relies on the one-to-one correspondence that exists between images simultaneously acquired on the science and the CLOWFS cameras. In either closed or open loops, recording CLOWFS images at a high temporal rate during a long exposure on the science camera can help predict the level of coronagraphic leaks attributable to pointing errors. A synthetic pointing leak image can be built and subtracted from the science image, to calibrate the low-order aberration residuals, ultimately improving contrast limits.

In its simplest version (see § 5 for a discussion of the possible complements), the postacquisition calibration of pointing errors with CLOWFS is a three-step procedure, described in the following sections and illustrated in Figure 3.

Step 1.—Calibrate the static optical configuration of the system (i.e., fine variations in optical path and alignment), using simultaneous pairs of images acquired with the science and the CLOWFS camera. This bank of pairs of images will be referred to as the dictionary (see § 2.1).

Step 2.—Track low-order modes during a long science exposure, continuously acquiring images at a fast frame rate with the CLOWFS camera (see § 2.2).

Step 3.—Postprocess the data by identifying a match for each CLOWFS image acquired during the science exposure in the dictionary (see § 2.2).

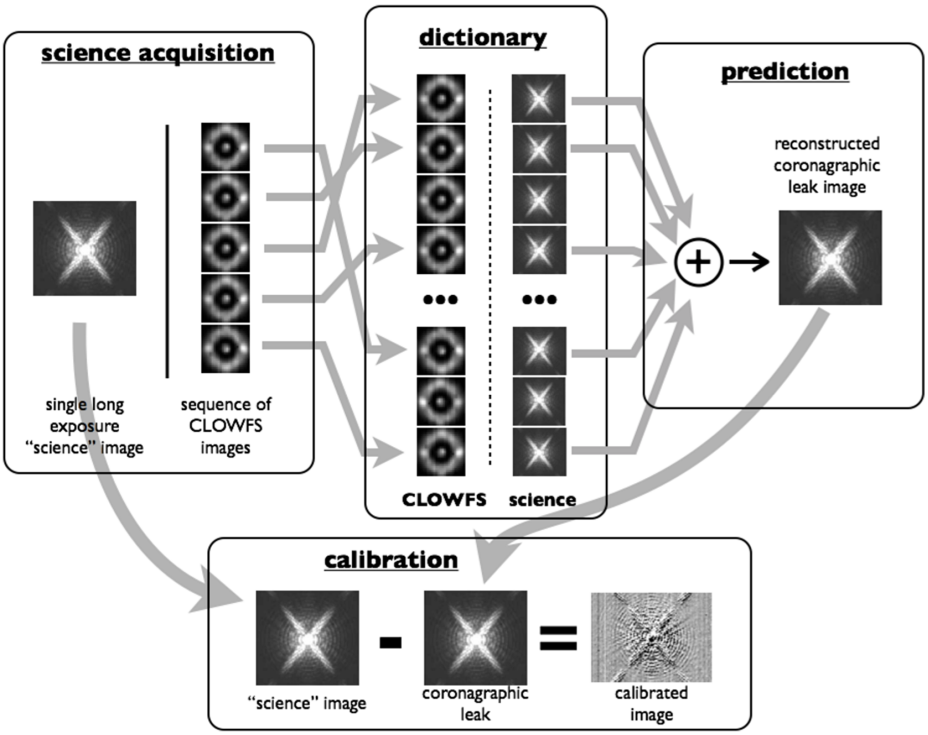


FIG. 3.—CLOWFS postacquisition pointing-error calibration procedure, defined as the MMA.

2.1. System Calibration: Building up the Dictionary

During step 1, pairs of simultaneous short-exposure images are acquired with the CLOWFS and science cameras on a calibration source (internal source or single star). Images for exposure k are, respectively, labeled DC_k and DS_k for the CLOWFS and science images. These images are stored into a database, called the dictionary.

After one such acquisition, as long as the optics remain stable, the CLOWFS image can be used as a key that points to a coronagraphic leak term (the corresponding science image). Later, any instant CLOWFS image can be compared with the entries in the dictionary: the best-matching entry in the dictionary allows the amount of coronagraphic leak to be predicted.

While slowly varying noncommon path errors can occur after the coronagraph, their effect can be kept small by minimizing the number of optics used after the CLOWFS focal-plane mask, as well as by regularly refreshing the content of the dictionary, to keep it up to date with the current status of the system.

A preliminary version of the dictionary can therefore be compiled in the laboratory prior to the actual observing, using a calibration source. However, to minimize systematic error terms, it must be completed with more up-to-date images acquired on a series of calibration (nonresolved) stars of spectral type and magnitude comparable with the science target, to get the best possible match. Ideally, during acquisition on the calibration star, one wants to cover a range of pointing errors that is larger than that experienced during the science exposure.

It is important to emphasize here that low-order aberrations errors are not explicitly calculated: instead, their consequences on the coronagraphic image are directly recorded, via the CLOWFS system. Being pragmatic, this approach eliminates the need for a high-fidelity (yet, most likely, imperfect) model of the coronagraph: ultimately, the ability to precisely characterize the coronagraphic leaks is determined by the coverage of the dictionary, which can be made arbitrarily large.

2.2. Computing and Subtracting Coronagraphic Leaks

During a long (i.e., typically greater than one second) exposure on the science camera, the CLOWFS camera acquires a sequence of short exposures (step 2), labeled C_i, that are saved along with the science camera image (see Fig. 3). To predict the coronagraphic leak in the science image, a matching algorithm (referred to as the match-maker algorithm [MMA]) identifies

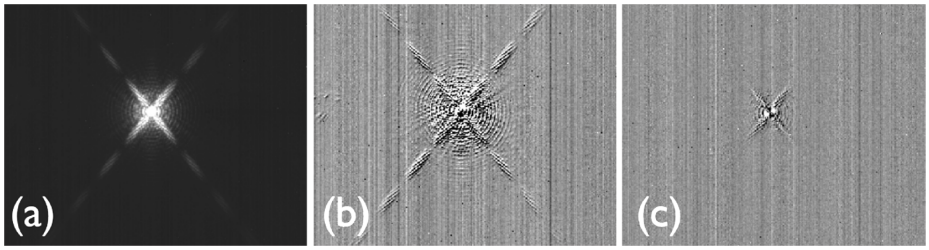


FIG. 4.—(*a*) Long-exposure "science camera" image, nonlinear intensity scale. Airy rings and diffraction spikes due to the peculiar Subaru Telescope pupil are clearly visible. (*b*) Image after standard subtraction of a PSF without using the proposed CLOWFS image selection procedure. (*c*) Image after CLOWFS postprocessing calibration; (*b*) and (*c*) use the same intensity scale.

each image of the CLOWFS sequence C_i to the CLOWFS image of the dictionary DC_k that best matches it. Figure 2 shows the typical structure of one CLOWFS image. The region of interest within the image is defined as a disk circumscribing the ring visible in the image. The criterion used to evaluate the match is simply the pixel-to-pixel deviation between the two images, averaged over this region of interest:

$$\sigma(C_i, DC_k) = \sqrt{\langle (C_i - DC_k)^2 \rangle}. \quad (1)$$

The operation is repeated for all images $(C_i)_{i=1}^{n}$ of the CLOWFS sequence. The matching short-exposure science camera images (DS_k) of the dictionary are then co-added to form an estimate of the residuals in the long-exposure coronagraphic image that can be attributed to varying low-order aberrations during the exposure on the science target. This final image is simply subtracted from the long-exposure to calibrate these pointing-error residuals.

3. LABORATORY DEMONSTRATION

The experimental demonstration of the proposed post-processing technique was performed using the SCExAO bench in the Subaru laboratory (Martinache & Guyon 2009). Among extreme-AO systems, SCExAO specializes in the high-contrast characterization of the innermost ($<0.2''$) surrounding of stars and, in that scope, implements a high-performance phase-induced amplitude apodization (PIAA)-based (Guyon 2003) coronagraph that takes full benefit of the angular resolution of the 8 m Subaru Telescope. For our test, the special optics described by Lozi et al. (2009) were taken out of the beam, and SCExAO was reduced to a conventional near-IR Lyot coronagraph without a Lyot stop (see Aime & Soummer [2003] for an introduction to coronagraphy with apodized pupils). SCExAO is operated at the IR wavelength (H band) and uses two identical uncooled fast-readout InGaAs cameras (Xenics model XS-1.7-320).

The light source is a $\lambda = 1.55$ μm laser diode, illuminating a 17 mm pupil mask emulating the Subaru Telescope pupil, producing the characteristic spikes in all the science images seen in this article. The focal-plane mask is 85 μm in radius and, given the f-number of the beam, gives a 4.75 λ/D inner working angle. Note that when the PIAA is inserted in the beam, the focal-plane scale is altered (Guyon et al. 2005) and the inner working angle becomes 1.5 λ/D.

For this experiment, a 25 minute sequence of 15,000 pairs of 1 ms exposure images was acquired with the CLOWFS and coronagraphic cameras on the SCExAO test bed at a ~10 Hz frame rate.[2] A series of 150 contiguous pairs of images in the middle of this sequence was isolated to simulate a long-exposure acquisition. The science images were dark-subtracted and co-added to simulate a single long-exposure coronagraphic image. The corresponding sequence of raw CLOWFS images was simply stored (they form the S_{c_k} list of images). The other 14,850 pairs of images, labeled D_{c_k} and D_{s_k}, were combined to form the dictionary.

The approach is illustrated by Figure 4. The simulated science frame (i.e., co-addition of the 150 selected images) is located in Figure 4a, and other panels compare the effect of two calibration procedures: standard PSF subtraction, as explained in § 4.2, in Figure 4b and the proposed CLOWFS post-processing procedure in Figure 4c.

4. RESULTS

4.1. Best Match for a CLOWFS Image

For small wavefront aberrations, Guyon et al. (2009) showed that the CLOWFS image is a linear function of the low-order modes to be measured: typically, (but not restricted to) pointing, focus, and astigmatism. The approach proposed here only uses the average pixel-to-pixel deviation between two CLOWFS images, reducing the contribution of all these modes to a

[2] During recent engineering observations with SCExAO, the CLOWFS loop was successfully closed, using 5–10 ms exposures at an improved 50 Hz frame rate.

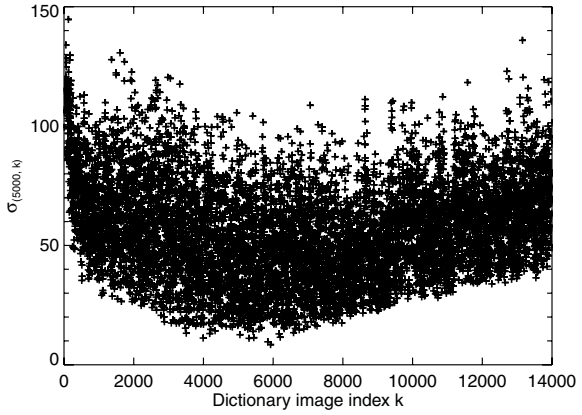


FIG. 5.—Evolution of the mean deviation $\sigma = \sqrt{\langle (DC_{5000} - DC_k)^2 \rangle}$ over the 25 minute acquisition sequence of the dictionary, using image index 5000 as reference. In addition to the slow trend that moves the average signal measured over the course of the sequence, the effect of the source of vibration used on the bench is obvious, as most of the recorded range of signal is explored in less than a minute.

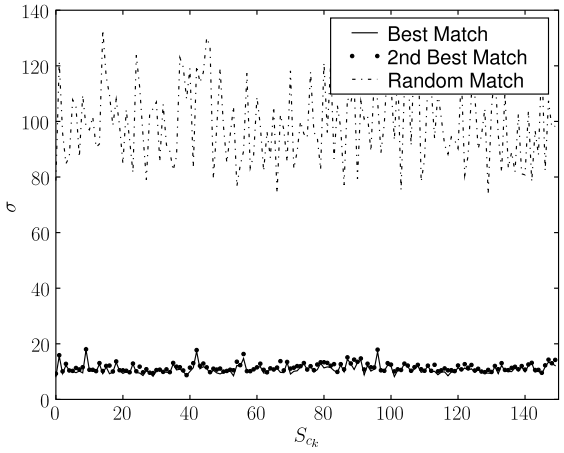


FIG. 6.—Mean deviation σ between each image S_{Ck} of the CLOWFS sequence and its respective best fit (*solid line*) and second-bit fit (*dots*). The dash-dotted line shows the same deviation, this time between S_{Ck} and a random CLOWFS image, averaged over 200 experiments.

single scalar, and does not attempt at separating the different contributions.

To assess the extent of the coverage of these low-order modes by the dictionary, one can arbitrarily pick one image as reference and calculate the deviation of each image of the sequence relative to this reference. Figure 5 illustrates one such experiment, using image 5000 as reference. The value of this reduced CLOWFS signal ranges from $\sigma \sim 10$ (the level of readout noise of the detector used for this experiment) to $\sigma \sim 140$, with a mean at ~ 55. Note, for reference, that in this configuration, a pure pointing error of $0.4\ \lambda/D$ induces a reduced CLOWFS signal ($\sigma \sim 110$).

The 25 minute sequence of signal was acquired with a strong source of vibration (an electric water pump) bolted to the bench, next to the focal plane. In addition to the low trend visible in Figure 5, the fast vibration ensures that CLOWFS quickly explores the full range of aberrations covered during this experiment. While building the dictionary, it is important to cover a sufficiently wide amplitude of low-order-mode signal, as in the postprocessing step, it improves the odds for the matching algorithm to find a dictionary image that matches any CLOWFS image.

Figure 6 (*solid line*) shows the value of the mean deviation σ between each image of the CLOWFS sequence and its best dictionary match. Hardly differentiable from the solid line is a dotted line, showing the deviation with the second-best match. For these two virtually identical cases, the deviation is dominated by the detector readout noise (~ 10 ADU pixel^{-1}). A randomly chosen dictionary image (*dash-dotted line*) exhibits a mean deviation that is, on average, 10 times as large.

4.2. Calibration of Coronagraphic Leaks

Figure 4 illustrates the significant improvement that the calibration of low-order-aberration–induced coronagraphic leaks with CLOWFS offers (Fig. 4c) over a more standard PSF subtraction (Fig. 4b).

The standard PSF subtraction shown in Figure 4b was derived from the data without knowledge of the CLOWFS frames. To simulate the standard acquisition of a PSF for calibration in a fair manner (in practice, this implies finding a single star of comparable brightness and spectral type, located at a comparable elevation), we randomly selected a number of images corresponding to the total science exposure duration from the science side of the dictionary and co-added them to form a second long-exposure frame simulation. Figure 4b shows the residuals after the PSF subtraction. While the outer part of the image is dominated by detector readout, the scientifically valuable central part of the field is dominated by residual speckle noise.

To produce the calibrated PSF shown in Figure 4c, the science images corresponding to indices selected by the MMA procedure are co-added to form what should be a very good PSF for the calibration of a coronagraphic image. The residual speckle level in this calibrated PSF is much lower than for the standard PSF subtraction.

While only one example is provided in Figure 4, a total of 200 different PSFs were also simulated, in order to provide a statistically relevant comparison between the CLOWFS and standard PSF subtraction residuals. Figure 7 shows all the radial profiles of raw standard-PSF-subtracted and CLOWFS-PSF-subtracted images. Each of the 200 simulated PSFs offers an improvement of the contrast over the entire field of view by

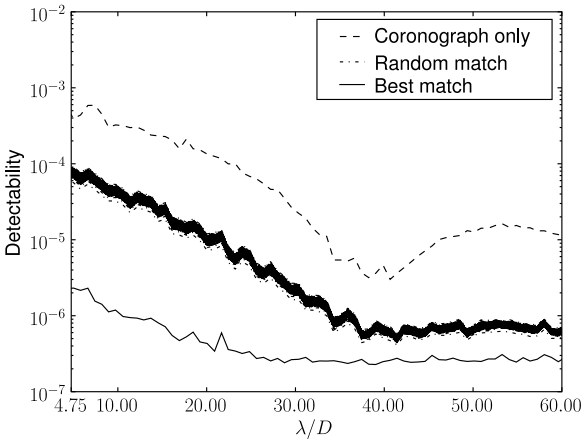


FIG. 7.—Standard deviation of the image profile as a function of angular separation. The figure compares three types of profiles: the raw science image (*dashed line*), the best-fit cleaned science image (*solid line*) using the proposed postprocessing procedure, and 200 random-match cleaned science images (*dot-dashed line*) simulating the standard PSF subtraction.

a factor of 10, in comparison with the raw image. In comparison, the CLOWFS PSF improves the calibration by almost 2 orders of magnitude between the inner working angle of the Lyot coronagraph ($4.75\ \lambda/D$ and ~$20\ \lambda/D$), where the contrast reaches the detector readout noise floor. In practice, a standard PSF subtraction would require a change of the pointing of the telescope to observe a calibration star. Such a change will change the structure of static aberrations, introducing a systematic error that will degrade the actual performance of the PSF subtraction, in comparison with what is presented here. On sky, CLOWFS-like calibrations are likely to prove to be even more advantageous.

5. CONCLUSION

The results presented in § 4 demonstrate that, used in postprocessing, CLOWFS improves the sensitivity of a coronagraph by a factor of 40, in comparison with more conventional—yet idealized–types of calibration. The method described here for this proof of concept is still quite rudimentary and could benefit from several improvements.

Instead of relying on an exhaustive search by a MMA-like algorithm on a massive database, necessarily containing redundant information, the dictionary could be simplified using approaches like principal component analysis, in which CLOWFS images would be projected onto a smaller subset of CLOWFS modes. In addition, an extended knowledge of several experimental parameters such as the telescope elevation, star color, or magnitude could be implemented in the dictionary to improve on the actual calibration.

The approach is nevertheless powerful, as it requires neither modeling of the coronagraph nor assumptions such as linearity of the response in order to work. The absence of such requirements makes it extremely robust and able to handle multiple situations and types of coronagraphic leaks.

The Subaru Coronagraphic Extreme-AO Project implements a CLOWFS used in a closed loop. The system architecture, however, allows for a simple implementation of the CLOWFS postprocessing, with no impact on the hardware, to calibrate residual coronagraphic leaks induced by low-order modes. CLOWFS on SCExAO currently generates a steady 4 Mbits s^{-1} data stream. While not prohibitively large for an experiment, the storage of entire nights of CLOWFS data will quickly prove to be impractical as SCExAO enters a more aggressive observing phase. Eventually, the calibration of coronagraphic leaks using CLOWFS will require the matching to be performed on-the-fly, so that only the final coronagraphic leak image is actually saved.

Because it makes no assumption on targets and the current state of the optical bench, CLOWFS used for both live and postprocessing is likely to play a significant part in the success of SCExAO, focused on the detection of planetary companions at small angular separation.

Frédéric Vogt thanks Brian Elms for sharing his experience and the machining of parts used for this project.

REFERENCES

Aime, C., & Soummer, R. 2003, in EAS Pub. Ser. 8, Astronomy with High Contrast Imaging (Les Ulis: EDP Sciences), 79
Beuzit, J., Feldt, M., & Dohlen, K., et al. 2008, Proc. SPIE, 7014, 70148
Digby, A. P., Hinkley, S., & Oppenheimer, B. R., et al. 2006, ApJ, 650, 484
Gendron, E., Clénet, Y., Fusco, T., & Rousset, G. 2006, A&A, 457, 359
Guyon, O. 2003, A&A, 404, 379
———. 2005, ApJ, 629, 592
Guyon, O., Matsuo, T., & Angel, R. 2009, ApJ, 693, 75
Guyon, O., Pluzhnik, E. A., & Galicher, R., et al. 2005, ApJ, 622, 744
Harder, S., & Chelli, A. 2000, A&AS, 142, 119
Janson, M., Bergfors, C., Goto, M., Brandner, W., & Lafrenière, D. 2010, ApJL, 710, 35
Jolissaint, L., Christou, J., Wizinowich, P., & Tolstoy, E. 2010, Proc. SPIE, 7736, 77361F
Kalas, P., Graham, J. R., & Chiang, E., et al. 2008, Science, 322, 1345
Lafrenière, D., Marois, C., Doyon, R., & Barman, T. 2009, ApJL, 694, 148
Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770
Lagrange, A., Gratadour, D., & Chauvin, G., et al. 2009, A&A, 493, L21

Lozi, J., Martinache, F., & Guyon, O. 2009, PASP, 121, 1232
Macintosh, B., Graham, J., & Palmer, D., et al. 2006, Proc SPIE, 6272, 62720L
Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556
Marois, C., Macintosh, B., & Barman, T., et al. 2008, Science, 322, 1348
Martinache, F., & Guyon, O. 2009, Proc. SPIE, 7440, 74400O
Mayor, M., & Queloz, D. 1995, Nature, 378, 355
Serabyn, G., Mawet, D., & Burruss, R. 2010, BAAS, 41, 587
Veran, J.-P., Rigaut, F., Maitre, H., & Rouan, D. 1997, J. Opt. Soc. Am. A, 14, 3057
Wallace, J. K., Burruss, R. S., & Bartos, R. D., et al. 2010, Proc. SPIE, 7736, 77365D

Contrôle actif de la diffraction dans le plan focal



FIGURE 36: "Give me a camera, sensitive and fast enough and a DM, and I shall beat the cr%p out of those bl**dy speckles !". Archimède, paraphrasé sans aucune honte lors de la conférence AO4ELT5.

La disponibilité d'un miroir déformable contrôlé par un algorithme utilisant l'image plan focal comme diagnostic complémentaire d'une optique adaptative permet d'aller au delà (voir figure 36) du genre de calibration rendu possible par l'interférométrie telle que je l'ai présentée dans la première partie de ce mémoire. Plutôt que d'utiliser le post-processing pour s'affranchir de des données corrigées par optique adaptative, ce qui reste une méthode palliative passive, une approche active est possible. Des modulations délibérées du miroir déformable, synchronisées avec les acquisitions par la caméra focale forment la base d'un test idéal de cohérence, qui permet de distinguer, dans une image à haute dynamique, la nature des speckles présents dans cette image.

Dans une description géométrique de l'optique adaptative, un miroir déformable peut être pensé comme une surface modulable, utilisée pour compenser les aberrations introduites par la turbulence atmosphérique : la surface d'onde originalement bosselée est aplatie ce qui se traduit par la production d'images de meilleure qualité : un exemple d'amélioration apportée par l'OA est présenté dans la figure 37.

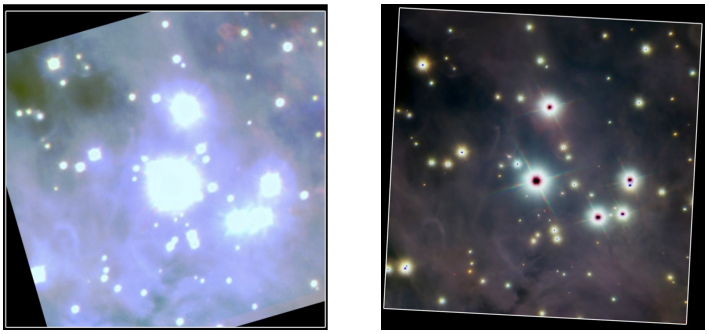


FIGURE 37: Images de la région centrale de la Nébuleuse d'Orion, dite le Trapèze, observée par le télescope Subaru, d'abord à gauche pendant des observations limitées par le seeing atmosphérique et à droite en présence d'optique adaptative. Les deux images ont été orientées de façon à faire correspondre les deux scènes astrophysiques. Le gain apporté par l'optique adaptative est évident.

Dans le cas d'une excellente correction adaptative, telle qu'elle est permise dans un système d'optique adaptative extrême comme SCExAO, il devient utile de décrire le miroir en terme de la diffraction qu'il génère dans le plan focal. Le miroir déformable de SCExAO, fait d'une grille périodique à trame carrée de 2000 actionneurs répartis dans une ouverture circulaire, permettent un contrôle très fin de la surface d'onde, comme montré dans la figure 38.

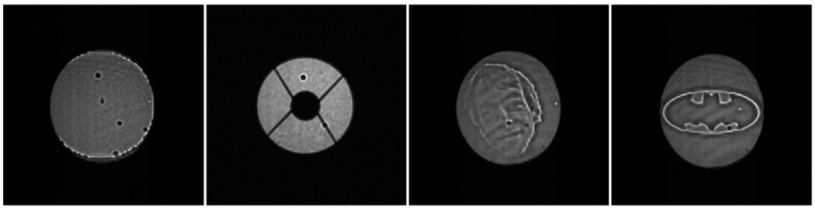


FIGURE 38: Image de la surface du miroir déformable de SCExAO. Une légère défocalisation permet (grâce à la diffraction de Fresnel) de faire apparaître la modulation de phase comme modulation d'amplitude. De gauche à droite : le miroir au repos avec les actionneurs morts visibles, le miroir à plat, avec un masque de pupille, et enfin, le miroir actionné de façon à faire apparaître Olivier Guyon ou le Batsignal.

La structure périodique du DM le rend particulièrement adapté à la génération de modulation sinusoïdales. Si une telle modulation

est appliquée, le DM se comporte comme un réseau de diffraction, déplaçant de la lumière normalement transmise sur l'axe à une distance inversement proportionnelle à la période de la modulation. Lorsque le miroir se trouve dans la pupille, c'est le nombre total de cycles de modulation[47] présents dans un diamètre de la pupille, qui va déterminer la distance séparant les ordres. Un exemple d'une telle modulation est présenté dans la figure 39 : une modulation de 15 cycles dans la pupille résulte en l'apparition d'ordres diffractés à une distance de 15 λ/D de l'axe optique.

47. pour un réseau blasé, on parlerait du nombre total de traits

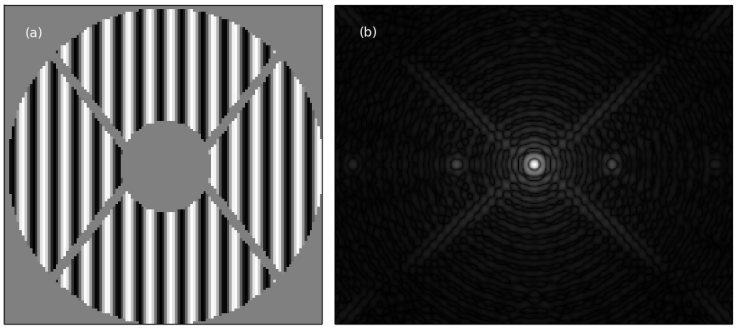


FIGURE 39: A gauche (a) représentation du front d'onde à l'intérieur de la pupille de l'instrument SCExAO. A droite (b) la figure de diffraction produite par ce front d'onde. En plus de l'image sur l'axe attendue, avec les fortes aigrettes de diffraction produites par l'ombre de la structure portant le miroir secondaire du télescope, la modulation sinusoidale (15 cycles dans le diamètre) produit des ordres diffractés à 15 λ/D. Les autres ordres diffractés visibles sur les bords de l'image sont crées par la grille discrète des 50 actionneurs présents dans un diamètre de pupille.

Cet ordre diffracté, c'est de la lumière de la source qui est observée sur l'axe. Si la source observée est un point source, cette lumière diffractée est parfaitement cohérente avec celle contenue dans les speckles déjà présents dans le plan focal. Les speckes et les ordres diffractés vont produire des interférences qui vont amplifier ou atténuer les speckles en place dans le plan focal : en ajustant l'amplitude et la phase de la modulation sinusoïdale, on peut trouver une solution qui va localement éteindre le speckle présent à un endroit précis du plan focal. En combinant ensemble plusieurs modulations sinusoïdales à des fréquences et des amplitudes différentes, il est possible de créer des régions plus contrastées dans le plan focal : c'est ce que l'on appelle le *speckle nulling*.

Des versions très sophistiquées d'algorithmes de contrôle global de front d'onde estimé depuis le plan focal ont été proposées pour créer une région de haut contraste, souvent désignée par l'expression anglais de *dark hole*[48]. La technique globale de *speckle nulling*[49] et ses variantes modernes comme la minimisation de stroke[50] sont utilisables pour un système optique censé rester stable entre l'étape de construction d'une matrice de réponse en amplitude complexe et l'étape où la boucle est fermée. Bien que l'ambition d'un système XAO comme SCExAO soit de rapprocher le système de telles conditions d'utilisation, par souci de robustesse, une approche itérative, basée sur une calibration minimaliste reste recommandée : c'est cette approche que je détaille ici.

48. F. Malbet, J. W. Yu, and M. Shao. High-Dynamic-Range Imaging Using a Deformable Mirror for Space Coronography. **PASP**, 107:386, Apr. 1995

49. P. J. Bordé and W. A. Traub. High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime. **ApJ**, 638:488–498, Feb. 2006

50. L. Pueyo, J. Kay, N. J. Kasdin, T. Groff, M. McElwain, A. Give'on, and R. Belikov. Optimal dark hole generation via two deformable mirrors with stroke minimization. **Appl. Opt.**, 48:6296, Nov. 2009

Amplitude de modulation et contraste dans l'image

Une modulation sinusoidale d'amplitude a (par exemple exprimée en microns) appliquée sur la surface du DM à pour un nombre de cycles donné k se traduit, pour une longueur d'onde λ donnée (exprimée dans la même unité que l'amplitude de la modulation), par une modulation de l'amplitude complexe A, dont le profil (à une dimension) s'écrit comme :

$$A(x) = \exp\left(i(4\pi a/\lambda)\sin 2\pi k x\right) \quad (9)$$

$$\approx 1 + i(4\pi a/\lambda)\sin 2\pi k x, \quad (10)$$

où la deuxième expression peut être utilisée lorsque l'amplitude a est petite devant la longueur d'onde λ.[51] Une relation de Fourier relie l'amplitude complexe dans la pupille à celle présente dans le plan focal. La transformée de Fourier de l'équation 10 est donc directement donnée par :

51. Notez que l'on a bien une amplitude de modulation de phase égale à quatre fois $\pi a/\lambda$ et pas deux, la réflexion sur le miroir doublant l'impact sur le front d'onde.

$$\hat{A}(u) = \delta(u) + i\frac{4\pi a}{\lambda}\left(\delta(u - 1/k) - \delta(u + 1/k)\right), \quad (11)$$

et la distribution d'intensité enregistrée par un détecteur situé dans ce plan focal est donné par la norme carrée de cette dernière expression : le contraste de luminosité entre les ordres diffractés et l'ordre transmis est directement donné (lorsque $a \ll \lambda$) par :

$$c = (4\pi a/\lambda)^2. \quad (12)$$

Dans une image haut contraste, à cause de la dynamique limitée de la caméra ou parce que l'on utilise un coronographe qui occulte l'ordre zéro, il n'est pas possible mesurer directement le contraste d'un speckle. On se contente de mesurer son intensité, dont on sait maintenant (voir equation 12) qu'elle est proportionnelle au carré de l'amplitude de la modulation :

$$I = \beta \times a^2. \quad (13)$$

La constante de proportionnalité C doit être étalonnée à chaque fois que les conditions d'observations sont modifiées : lorsqu'on change de cible ou lorsqu'on change le temps d'intégration de la caméra. Cet étalonnage se fait en introduisant délibérément une modulation d'amplitude connue, et en mesurant l'intensité effectivement enregistrée.

Interactions entre speckles

Les speckles introduits par le DM ne sont jamais isolés dans le plan focal : ils vont interférer avec les structures de diffraction déjà présentes dans le champ, avant la modulation. Selon la différence de phase entre cette modulation et les autres speckles déjà présents, l'interférence peut être constructive ou destructive. Pour produire une

interférence destructive qui résulte en une atténuation locale de l'intensité dans l'image, il faut que l'amplitude du speckle ajouté soit égale à celle du speckle en place, et que sa phase lui soit opposée. L'intensité I d'un speckle d'amplitude complexe inconnue ne dépend que du module de l'amplitude complexe, et pas de sa phase inconnue, notée Φ_0 :

$$I_0 = \beta \times \left|\left|a_0 e^{i\Phi_0}\right|\right|^2 = \beta \times \left|\left|a_0\right|\right|^2. \tag{14}$$

Le module de l'amplitude complexe a_0 peut être estimée à partie de cette intensité, à condition que la constante β ait été étalonnée au préalable. La phase de ce speckle ne peut par contre pas être devinée à partir de l'analyse d'une seule image. Pour la déterminer, on va suivre l'évolution de l'intensité à l'endroit du speckle original, lorsqu'il interfère avec une sonde d'amplitude complexe $ae^{i\Phi}$, dont on va faire varier la phase Φ. L'intensité de la somme cohérente de ces deux speckles est décrite par une équation classique d'interférences à deux ondes :

$$\begin{aligned} I &= \beta \times \left|\left|a_0 e^{i\Phi_0} + a e^{i\Phi}\right|\right|^2 && (15)\\ &= \beta \times \left(a_0^2 + a^2 + 2a_0\, a \cos\left(\Phi_0 - \Phi\right)\right). && (16) \end{aligned}$$

Cette fonction d'interférence est évaluée pour un nombre variable de sondes d'amplitude a constante mais de phase Φ uniformément échantillonnée entre 0 et 2π. Au moins trois valeurs distinctes de Φ sont nécessaires pour contraindre les valeurs de a_0 et Φ_0. En pratique, il est recommandé de moduler la phase Φ avec un pas plus fin, pour minimiser la sensibilité aux variations résiduelles de la phase instrumentale qui vient polluer la mesure.

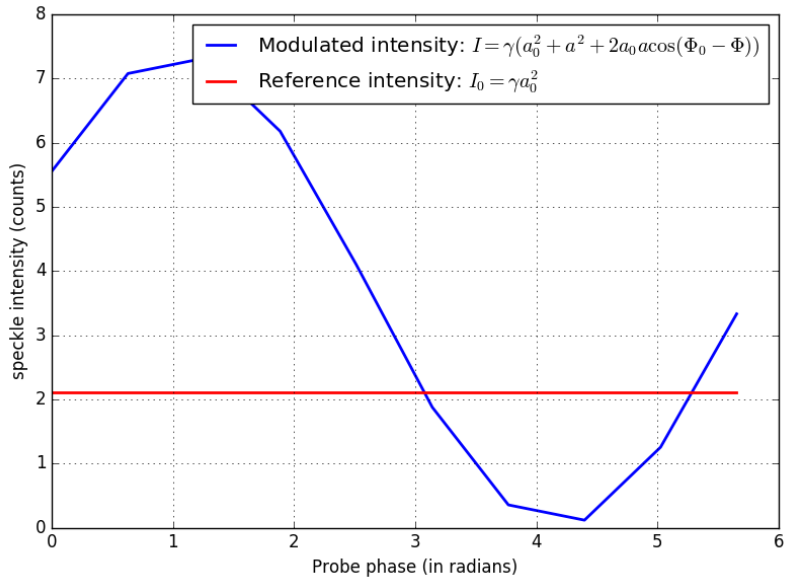


FIGURE 40: Exemple de courbe d'intensité enregistrée lors de la modulation d'un speckle.

Un exemple idéal de modulation est représenté dans la figure 40. Le niveau de référence, représenté en rouge, correspond à l'intensité enregistrée sur le speckle en l'absence de modulation. A partir de la relation calibrée 13, l'amplitude du speckle a_0 est estimée pour produire une sonde d'amplitude a, qui est normalement proche de la valeur réelle a_0. On scanne ensuite la phase de la modulation entre 0 et 2π ce qui conduit à l'obtention de la courbe représentée en bleu. Avec le nombre important (ici $n = 10$) de modulations, il est aisé de trouver la phase du speckle initial, qui correspond à la position du maximum de la courbe ($\phi_0 = \pi/3$ dans l'exemple représenté). Avec quatre modulations, correspondant aux phases 0, $\pi/2$, π et $3\pi/2$, une solution analytique existe pour déterminer l'amplitude complexe du speckle, avec la méthode dite A-B-C-D. Une telle approche suppose cependant que les quatre mesures de l'intensité ne soient pas affectées par des perturbations résiduelles de la phase atmosphérique.

UNE APPROCHE PLUS GÉNÉRALE ET PLUS ROBUSTE EST POSSIBLE. Elle consiste à ajuster un modèle paramétrique de la courbe d'intensité, en minimisant un critère des moindres carrés. En pratique, en complément du vecteur I_S des intenstités enregistrées pour les N différentes valeurs de phase de la sonde, on précalcule un vecteur W contenant les puissances successives de la racine N-ième de l'unité $\omega_N = e^{i2\pi/N}$.

Une estimation de la valeur de la phase Φ_0 du speckle sondé est directement donnée par l'argument du produit scalaire de ces deux vecteurs :

$$\Phi_0 = \arg\left(I_S^\top \cdot W\right) \tag{17}$$

Le module de visibilité de la modulation décrite par l'équation 16, noté γ est relié au module de ce même produit scalaire :

$$\gamma = \frac{2a\,a_0}{a^2 + a_0^2} = \frac{2}{N}\left|\left|I_S^\top \cdot W\right|\right|. \tag{18}$$

L'amplitude a_0 du speckle est don une des deux racines de l'équation quadratique suivante :

$$a_0^2 - \frac{2a}{\gamma}a_0 + a^2 = 0, \tag{19}$$

toutes deux données par :

$$a_0 = \frac{a}{\gamma}\left(1 \pm \sqrt{1-\gamma^2}\right). \tag{20}$$

L'amplitude a de la sonde est choisie de façon à être proche de l'amplitude a_0 du speckle qu'on cherche à caractériser, ce afin de maximiser le module de visibilité γ de la courbe d'intensité et par conséquent d'optimiser le rapport signal sur bruit de l'estimation des propriétés du speckle. Cependant, sans savoir si l'amplitude de la sonde utilisée était supérieure ou inférieure à celle du speckle, il est difficile de choisir laquelle des deux solutions est la bonne,

sinon celle de plus petite amplitude (avec le signe moins), de façon à éviter une divergence de la boucle de contrôle, quitte à réduire son efficacité globale. Une autre option est de délibérément augmenter l'amplitude de la sonde, par exemple de 5 %. On perd un peu de visibilité mais on sait systématiquement qu'il faut choisir la solution avec le signe moins.

Speckle nulling

Utilisant l'algorithme qui vient d'être présenté, il est possible de créer une boucle de contrôle de front d'onde qui modifie la fonction de transfert instrumentale, en particulier pour créer une région de plus haut contraste dans le voisinage de l'objet pointé. La géométrie de la région de contrôle est principalement contrainte par les propriétés du miroir déformable. Le nombre total d'actionneurs présents dans un diamètre de la pupille impose une fréquence spatiale maximale au delà de laquelle il ne contrôle plus rien : les 50 actionneurs présents dans la pupille de SCExAO permettent de contrôler une région qui ne peut pas être plus étendue qu'un demi-disque centré sur la source pointée, et de rayon $r_{max} = 25\lambda/D$. L'utilisateur peut cependant choisir de réduire le rayon de la région contrôlée si il veut concentrer ses efforts sur un domaine spécifique de séparation angulaire.

LA BOUCLE DE CONTRÔLE PEUT ÊTRE MULTIPLEXÉE. Un algorithme itératif adressant un à la fois les nombreux speckles du *dark hole* serait relativement inefficace : la technique de sondage des speckles peut heureusement être multiplexée et adresser simultanément plusieurs dizaines de speckles. En pratique, au début de chaque itération, la boucle de contrôle en place sur SCExAO identifie et marque la position et l'intensité des n_s ($\leq$ 50) structures les plus brillantes présentes dans la région de contrôle. Autour de chaque speckle identifié, on crée une zone d'exclusion, de rayon $2\lambda/D$ afin d'éviter le risque de compétition entre fréquences spatiales trop proches les unes des autres et on continue la recherche jusqu'à avoir trouvé les n_s positions à sonder.

Si les rayons internes r_{min} et externes r_{max} de la région de contrôle sont exprimés en unités de λ/D, il est possible d'estimer assez facilement le nombre maximal de fréquences spatiales sondables simultanément. Avec une zone d'exclusion de rayon $2\lambda/D$, ce nombre total de speckles correspond à un quart de l'aire de la différence entre les demi-disques de rayon r_{max} et r_{min}, soit

$$n_{max} = 1/8 * \pi * (r_{max}^2 - r_{min}^2). \quad (21)$$

Pour une région de contrôle comprise entre 5 et 25 λ/D, ce nombre maximal théorique de fréquences spatiales simultanément contrôlables est de 235.

A ces positions sont associées des fréquences spatiales sur le miroir déformables : l'algorithme décrit précédemment peut alors être

utilisé en modulant simultanément les n_s fréquences spatiales sélectionnées. L'exemple d'une telle modulation simultanée par le miroir déformable pour $n_s = 50$ speckles est représenté dans la figure 41. Si une structure présente dans l'image n'est pas un speckle résiduel de l'étoile coronographiée mais correspond par exemple à un compagnon planétaire, la lumière ajoutée par la modulation en provenance de l'étoile ne va pas interférer avec cette structure (grâce à la propriété fondamentale de l'incohérence spatiale des sources) : l'intensité mesurée pour cette position dans l'image ne fluctue pas. Dans l'état actuel des choses, si cette structure domine par sa brillance le champ de speckle, l'algorithme va continuer à s'acharner dessus. On perd en efficacité mais la télémétrie enregistrée pendant ces tentatives ratées de modulation sont autant de mesures de la nature incohérente de l'objet mis en évidence, qui seront utiles au moment du traitement *a posteriori*.

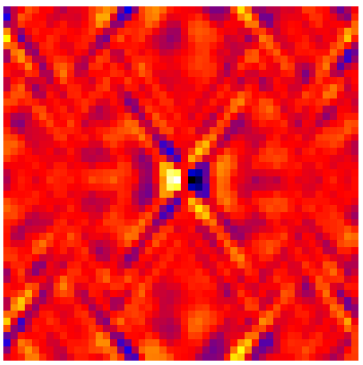


FIGURE 41: Exemple de modulation simultanée (simulée) de 50 fréquences spatiales par le miroir déformable de SCExAO, correspondant à une des premières itérations de speckle nulling, sur la région de contrôle située à droite du champ, en l'absence de coronographe.

Applications

Les articles joints à cette section présentent quelques cas d'utilisation de *speckle nulling* d'abord utilisé en laboratoire[52] pour démontrer expérimentalement les propriétés attendues d'un coronographe PIAA utilisant un inverse. Il est difficile d'imaginer qu'une telle démonstration ait été nécessaire : il faut néanmoins rappeler qu'à l'époque, certains doutaient de la pertinence du PIAA inverse, allant même jusqu'à imaginer qu'il ne ferait qu'empirer les propriétés de non-invariance de la PSF produite par le PIAA en amont. Notez au passage que le schéma optique de SCExAO était alors différent et utilisait un miroir déformable présentant un nombre plus petit d'actionneurs (32x32), placé dans un plan intermédiaire entre l'image et la pupille. Malgré cette configuration clairement sous-optimale, SCExAO a été capable de produire un *dark-hole* de contraste brut modéré (de l'ordre de $10^{-3.5}$ mais à de très faibles séparations angulaires, avec un *inner working angle* (IWA) de 2 λ/D.

52. F. Martinache, O. Guyon, C. Clergeon, and C. Blain. Speckle Control with a Remapped-Pupil PIAA Coronagraph. **PASP**, 124:1288–1294, Dec. 2012

Dans cette même configuration, l'algorithme de *speckle nulling* a ensuite été mise à l'épreuve sur le ciel[53], avant même de disposer de la correction adaptative XAO. La performance médiocre, s'explique en grande partie par le bruit de lecture important (>300e^-) du détecteur utilisé à l'époque (Xenics Xeva XS 1.7-320) pour faire ce travail. L'article montre malgré tout, que cette technique itérative de contrôle de front d'onde est robuste aux résidus turbulents et converge (lentement) vers une solution qui optimise le contraste brut aux plus faibles séparations angulaires et réduit l'amplification du bruit de speckle par les erreurs quasi-statiques, ce qui optimise le contraste des images après traitement *a posteriori*.

53. F. Martinache, O. Guyon, N. Jovanovic, C. Clergeon, G. Singh, T. Kudo, T. Currie, C. Thalmann, M. McElwain, and M. Tamura. On-Sky Speckle Nulling Demonstration at Small Angular Separation with SCExAO. **PASP**, 126:565–572, June 2014

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **124**:1288–1294, 2012 December
© 2012. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

Speckle Control with a Remapped-Pupil PIAA Coronagraph

FRANTZ MARTINACHE, OLIVIER GUYON, CHRISTOPHE CLERGEON, AND CELIA BLAIN

National Astronomical, Subaru Telescope, Observatory of Japan, Hilo, HI 96720; frantz@naoj.org

Received 2012 June 13; accepted 2012 October 23; published 2012 November 27

ABSTRACT. Phase-induced amplitude apodization (PIAA) is a well-demonstrated high-contrast technique that uses an intermediate remapping of the pupil for high-contrast coronagraphy (apodization), before restoring it to recover classical imaging capabilities. This paper presents the first demonstration of complete speckle control loop with one such PIAA coronagraph. We show the presence of a complete set of remapping optics (the so-called PIAA and matching inverse PIAA) is transparent to the wavefront control algorithm. Simple focal-plane-based wavefront control algorithms can thus be employed, without the need to model remapping effects. Using the Subaru Coronagraphic Extreme AO (SCExAO) instrument built for the Subaru Telescope, we show, using a calibration source, that a complete PIAA coronagraph is compatible with a simple implementation of a speckle nulling technique, and demonstrate the benefit of the PIAA for high-contrast imaging at small angular separation.

Online material: color figures

1. INTRODUCTION

Contrast limits for the direct imaging of extrasolar planets from ground-based adaptive optics (AO) observations are currently set by the presence of static and slow-varying aberrations in the optical path that leads to the science instrument (Marois et al. 2003). These aberrations, due to the noncommon path error between the wavefront sensor and the science camera, are responsible for the presence of long-lasting speckles in the image. Because extrasolar planets are unresolved sources, it is difficult to discriminate them among these speckles. One family of techniques, called differential imaging, is aimed at calibrating out some of these static aberrations, by using either sky rotation (angular differential imaging, or ADI), polarization (PDI), or wavelength dependence of the speckles (spectral differential imaging or SDI). Of these, ADI (Marois et al. 2006) seems very well adapted to the problem of the detection of extrasolar planets and has been successful, most notably producing the image of the planetary system around HR 8799 (Marois et al. 2008). ADI uses the rotation of the sky that naturally happens while tracking with an altitude-azimuthal telescope around transit. The position of static and slowly varying speckles, tied to the diffraction by the pupil, remains stable over long timescales, while the image of planetary companions rotates around the one of the host star.

The rotation of the field only leads to sufficient linear displacement for angular separations of the order of 1″. And in practice, below 0.5″, the performance of ADI quickly degrades below the threshold where planets can be detected.

One way to complement ADI toward small angular separation, is to use a deformable mirror (DM) to modulate speckles and introduce the diversity that will distinguish them from genuine structures like planets and lumps in disks. This type of technique is regularly used for high-contrast experiments (Guyon et al. 2010b) and appears as the technique of choice for a spaceborne mission dedicated to the direct imaging of high-contrast planets. This is also the approach we propose to use for the detection of extrasolar planets at small angular separation during ground-based adaptive optics (AO) observations. With this in mind, we are integrating and testing the Subaru Coronagraphic Extreme AO (SCExAO) project, whose optics have been described by Lozi et al. (2009) and Martinache et al. (2011).

In comparison with other extreme-AO projects (Macintosh et al. 2008; Beuzit et al. 2010), SCExAO implements an aggressive PIAA coronagraph, using a remapping of the pupil (Guyon et al. 2005; Martinache et al. 2006) optimized for high-contrast detection at small angular separations (down to 1 λ/D).

Laboratory high-contrast experiments relying on PIAA have already demonstrated high-contrast imaging capability in the 2–4 λ/D angular separation range, and achieved raw contrast of $\sim 10^{-7}$ and beyond (Guyon et al. 2010b; Kern et al. 2011; Belikov et al. 2011) that are several orders of magnitude beyond what a ground-based instrument is expected to produce (Guyon 2005). The SCExAO project attempts to apply some of the technology, tools, and techniques originally developed for spaceborne coronagraphy to ground-based AO. While space coronagraphy targets the ambitious goal of 10^{-10} raw contrast, required for the detection of an Earth-like planet around a Sun-like star, ground-based coronagraphy is done in a much less favorable and much more unstable environment. Raw contrasts achieved so far by coronagraphy downstream AO are of the order of 10^{-2} at a few λ/D, and in practice the direct maging

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

of planetary candidates relies heavily on postprocessing techniques like ADI to reach the published 10^{-5} to 10^{-6} contrast detection limits.

From a reasonably good starting point, high-contrast laboratory experiments can produce high-contrast images in a fairly small number of iterations, using an electric field conjugation (EFC) framework that relies on an acute knowledge of the system's complex amplitude response matrix (Give'On 2006; Bordé & Traub 2006). Ultimately, these techniques seem implementable at the telescope once an extreme AO system produces a continuous stream of stable high-Strehl images. For now, they remain limited to the pampered environment of the laboratory. At this stage of its development (referred to as SCExAO Phase 1), SCExAO does not include a fully functioning high-order wavefront sensor. The 32x32 DM it implements can nevertheless be used to actively generate speckle diversity and supplement ADI at small angular separations. SCExAO Phase 1 relies on iterative speckle nulling, to produce a region of higher contrast, often referred to as a "dark hole" (Malbet et al. 1995) in the field.

2. SPECKLE CONTROL WITHIN A PIAA CORONAGRAPH

The Phase-induced amplitude apodization (PIAA) coronagraph (Guyon 2003) is a high-contrast imaging device that enables the detection of faint sources down to an angular separation of $\sim 1\ \lambda/D$. A PIAA coronagraph is made of a set of aspheric optics designed to apodize the pupil by geometrically redistributing the light while preserving the overall collimation of the beam for an on-axis source. Because it alters the pupil, the point spread function (PSF) of such a system is, however, no longer translation-invariant. On SCExAO, the impact of the PIAA on the pupil is quite dramatic, as it goes as far as filling the void left by the 30% central obscuration of the Subaru Telescope pupil to produce an apodized pupil better suited for high-contrast imaging. The impact for off-axis sources is therefore also quite spectacular, as it produces strongly elongated pineapple-shaped PSFs beyond a few λ/D (Lozi et al. 2009).

While successful wavefront control experiments using PIAA have already been reported (Guyon et al. 2010b; Kern et al. 2011; Belikov et al. 2011), these experiments have all used a DM located downstream of the PIAA. The outer working angle (OWA) of this type of control is imposed by the total number of available actuators across one pupil diameter. By placing the DM downstream of the remapping optics, the OWA is reduced by a factor ~ 3, due to the plate scale change induced by the remapping (Guyon et al. 2010b). The resulting $\sim 5\lambda/D$ OWA (since there are 32 actuators across the pupil) would be a very serious limitation for any direct imaging instrument. Instead, the SCExAO project implements a complete PIAA coronagraph, with a DM located before any of the remapping optics (see Fig. 1). After the focal plane mask, a copy of the PIAA optics plugged backwards (referred to as the inverse PIAA) is introduced to undo the remapping of the pupil after the focal plane mask. This setup allows the recovery of the wide field-of-view imaging capability of the instrument (Guyon & Roddier 2002; Guyon 2003; Vanderbei & Traub 2005). Indeed, despite the drastic transformation of the wavefront operated by the PIAA, the remapping of the pupil remains a geometric operation. The inverse PIAA simply guarantees that what is known since the end of the 19th century as "Abbe's sine condition" is met after the remapping by the PIAA.

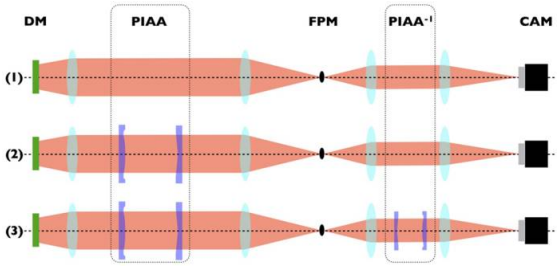


FIG. 1.—Schematic representation of three of the possible optical configurations of the SCExAO bench used for the results presented in this article. From *left* to *right*, the main components are highlighted as follows: the deformable mirror (labeled DM) located at an intermediate plane between pupil and image, the PIAA located in a collimated beam, the focal plane mask (labeled FPM) occulting the bright source observed with the system, the inverse PIAA (labeled PIAA^{-1}), also located in a collimated beam, in a plane conjugated with the PIAA so as to remap the pupil back to what it is when it enters the system, and finally, the detector (labeled CAM). From top to bottom, the three configurations are: (1) no remapping optics, equivalent to a Lyot coronagraph without a Lyot stop, (2) with PIAA only, and (3) with PIAA and inverse PIAA (the full coronagraph). See the online edition of the *PASP* for a color version of this figure.

The SCExAO instrument is a flexible platform, installed after Subaru Telescope's facility AO system, and designed to be used with the coronagraphic imager HiCIAO (Hodapp et al. 2008), and increase the size of the parameter space currently explored by the SEEDS survey (Tamura 2009; Tamura & SEEDS Team 2010), toward small angular separations. The infrared arm of SCExAO alone (Martinache et al. 2011) supports multiple optical configurations (see Fig. 1), from straightfoward imager (with or without a focal plane mask), to a complete PIAA coronagraph.

Lozi et al. (2009) have already demonstrated that the inverse PIAA, once aligned and conjugated to the PIAA, indeed restores wide field-of-view imaging capability of the system, at least up to 20 λ/D, which is beyond the outer working angle defined by the number of DM actuators across the instrument pupil. With this in mind, it seems reasonable to assume that the remapping optics should therefore be completely transparent for the wavefront control. The results presented in this article demonstrate that indeed, with a complete PIAA coronagraph (including the inverse PIAA), a simple wavefront control loop (speckle nulling) converges while being oblivious to the two remappings of the pupil.

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

2.1. Optical Configurations of SCExAO

To illustrate the impact of the remapping optics on the wavefront control, we use three distinct optical configurations of the SCExAO infrared arm, shown in Figure 1. All results reported in this paper were obtained with the SCExAO instrument fed by an external Subaru Telescope simulator (single-mode fibered laser source, $\lambda = 1.55$ μm + static turbulence phase screen) in a unstabilized environment. The reader will observe that in its current implementation, the 32x32 MEMS DM of SCExAO is not located in a pupil plane. This design was chosen for the simplicity of the optical layout, as well as to minimize the total number of reflections in the system. The design, however, also has some drawbacks. First, it requires a somewhat conservative beam size projected on the DM, to account for beam walk effects that would otherwise significantly vignette the field. The DM also needs to be horizontally tilted (~24°), so the density of actuators is not the same in horizontal and vertical directions. In practice, instead of illuminating all 32 available actuators across one pupil diameter, the beam spreads over a 27.2×24.8 actuator ellipse. SCExAO can afford to have its DM away from the pupil, since it is used downstream from an AO system that not only stabilizes the pupil but also considerably reduces the total amount of aberrations the system needs to correct for.

While the DM is always part of the optical train, both the PIAA and the inverse PIAA are mounted on motorized stages that can swing in and out of the beam with excellent repeatability, thus allowing for quick alternation between the three configurations: no remapping, PIAA only, and PIAA + inverse PIAA (see Fig. 1). In a speckle control loop, the DM is used to introduce spatial frequencies that destructively interfere with speckles induced by wavefront aberrations from the input beam. The DM can also be used to create speckles and, in turn, provide a very instructive demonstration of how (see Fig. 2) aberrations propagate through the SCExAO coronagraph.

Panel (*a*) of Figure 2 shows the 32x32 voltmap sent to the DM for this experiment. In addition to the voltmap that has been determined to produce the best output wavefront in the absence of input aberrations (what will now be referred to as the DM flat-map), two sine waves of amplitude 0.4 radians and maximal spatial frequency (in the Nyquist sense) are added, along the horizontal and vertical directions. These additional sine functions on the DM create four bright speckles, clearly visible in the focal plane visible on Panel (*b*) of Figure 2, in the absence of remapping optics. The oversized focal plane mask effectively hides the brightest area, but in the absence of Lyot stop, the diffraction by the spider arms bearing the secondary mirror of the telescope are left. Because the speckles result from the highest spatial frequency that can be introduced by the DM, their locations in the image also mark the edges of the region that can be controlled by the DM. In the absence of amplitude defects, appropriate actuation of the DM could clear the speckles over the entire 27.2×24.8 λ/D box centered on the on-axis target. In practice, because the beam includes amplitude as well as phase defects, the DM can only operate over a half of this entire region. For the results reported in this work, we arbitrarily chose to work on the left-hand side of the field. Speckles located beyond 13.6 λ/D along the horizontal axis and 12.4 λ/D along the vertical axis simply will not be affected by the the DM. Note that with a $0.04''$ diffraction limit in the H band, 12.4 λ/D on the Subaru Telescope closely matches the range of angular separation where ADI becomes usable ($12.4 \times 0.04 \simeq 0.5$ arcsecond).

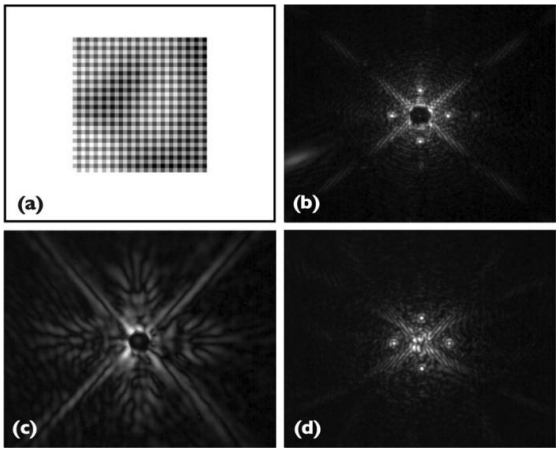


FIG. 2.—Demonstration of the effect of remapping optics on the image recorded by SCExAO's science camera. Panel (*a*) shows the voltmap (ranging from 85 to 125 V) applied to the MEMS DM to produce the images that follow. Panels (*b*), (*c*), and (*d*) use a nonlinear (square root) intensity scaling to highlight the faint diffraction structures surrounding the bright central area. Panel (*b*) shows the science camera image when the system is in configuration (1) (see Fig. 1). Panel (*c*) shows the dramatic effect of the PIAA on the focal plane image. Panel (*d*) shows the image obtained with the full coronagraph. Because the PIAA + inverse PIAA successfully restores the image of the off-axis speckles, it is expected that any type of focal-plane based wavefront control can ignore the presence of these remapping optics.

2.2. Impact of the Remapping Optics

The introduction of the PIAA dramatically impacts the structure of these image (see panel (*c*) of Fig. 2). While the diffraction spikes created by the spider remain, most of the diffraction rings due to the sharp edge of the telescope pupil disappear, and are more effectively hidden by the focal plane mask whose size now better matches the on-axis PSF. Off-axis, however, the bright diffraction-limited images of introduced speckles are turned into complex pineapple-shaped aberrated structures. The benefit of the PIAA remapping will be better demonstrated when comparing the performance of the speckle nulling loop with and without using the remapping optics in § 3.2.

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

The complete PIAA coronagraph (configuration 3) includes, after the focal plane mask, a copy of the PIAA optics mounted backwards to restore the geometry of the pupil back to what it was before entering the coronagraph. From geometric optics principles only, one expects this restoration to be perfect, but diffraction and the presence of the occulting mask in the focal plane will impact the output wavefront. Panel (*d*) of Figure 2 shows the resulting image: the bright off-axis speckles have been fully reconstructed by the inverse PIAA, and have recovered their Airy-disk shape. The dark disk left by the focal plane mask that was visible in the other two images is no longer obvious, as the disk is remapped by the inverse PIAA.

The PIAA + inverse PIAA combination successfully restores the image of the off-axis speckles back into translation invariant copies of the central PSF. For speckles that would be sufficiently far from the edge of the focal plane mask, it is expected that any type of focal-plane-based wavefront control can ignore the presence of these remapping optics. It is, however, not obvious for the speckles located nearby the original footprint of the focal plane mask that they should behave as simply. Section 3 will show, however, that speckle control can be achieved within the entire control region of the DM.

3. SPECKLE NULLING RESULTS

3.1. The PIAA Coronagraph is Transparent to the Speckle Control

To test the impact of the remapping (PIAA + inverse PIAA) optics for a focal-plane-based wavefront control loop, we use a simple iterative speckle nulling algorithm inspired from the original "dark hole" coronagraphy work by Malbet et al. (1995). Our implementation however targets specific speckles (up to a dozen speckles are simultaneously probed). This approach seems fairly inefficient in comparison with more sophisticated electric field conjugation (EFC) approaches that can achieve the same result in just a few iterations (Give'On 2006). Speckle nulling was nevertheless chosen because of its robustness and ease of implementation.

Indeed, complete EFC-based techniques rely on the knowledge of a response matrix that relates the complex amplitude of the on-axis source in the focal plane to the input wavefront. With an extreme AO loop effectively stabilizing the wavefront entering the coronagraph into something predictable, this matrix is likely to be quite stable and therefore reliable. However in its first phase of deployment at the telescope, SCExAO does not include the fast wavefront sensor. The simple speckle nulling approach should, despite the presence of a dynamic atmospheric component, calibrate out the wavefront features responsible for the presence of static speckles in the images (Guyon et al. 2010b).

Iterative speckle nulling works as follows: in a given image, up to n speckles are identified and their positions marked, relative to the central source. With a conventional imaging system, each speckle position corresponds to a two-component (x, y) spatial frequency on the DM, while its brightness indicates the amplitude a_0 of this spatial frequency. The only real unknown is the phase φ of each speckle, that can take any value between 0 and 2 π. In the following four acquisitions, to each speckle is added a speckle probe of same amplitude a_0, but each time with a different phase: 0, $\pi/2$, π and 3 $\pi/2$. The intensity of the four speckles resulting from the interference of the original speckle and the probes is used to determine its true phase ϕ_0. For the next image, also used as the input for the next iteration, a spatial frequency of opposite phase $\phi_0 + \pi$ and of amplitude $g \times a_0$, where g is the loop gain ($0.0 < g < 1.0$) is added so as to suppress the speckle.

The result of a series of 50 such speckle nulling iterations is presented in Figure 3. The starting point, shown in the left panel, shows the structure of the speckles due to amplitude and phase defects that filtered through the coronagraph. One will observe that all ring-like structures due to the sharp inner and outer edge of the pupil have been erased from the image, showing that the PIAA—focal plane mask—inverse PIAA combination achieves its purposes. Although faint (the average contrast over the entire control region is 2×10^{-3}), the most striking features can almost entirely be attributed to the spider arms bearing the secondary. After about 50 speckle nulling iterations (see right panel of Fig. 3), the average contrast inside the control region is brought down to 4×10^{-4}. The algorithm manages to suppress the diffraction features due to the spider arms, along which the gain in raw contrast is over 10^2. Overall, the speckle nulling loop improves the contrast by a factor of 10 over the entire control region.

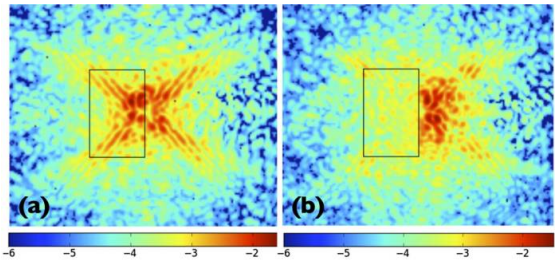


FIG. 3.—Example of high contrast result achieavable with the SCExAO coronagraph using a simple speckle nulling control loop. Panel (*a*) shows the starting point of the loop, with the deformable mirror in its nominal flat-map configuration. Note that in addition to some low-spatial frequency aberrations (created by a static turbulence plate), most of the speckles present at the starting point are located along the diffraction spikes created by the spider arms of the telescope pupil. Panel (*b*) shows the result of about 50 speckle nulling iterations, working on up to 10 speckles at a time, effectively clearing a box-shaped region of speckles, from 0 to 14 λ/D in the horizontal direction and within ± 14 λ/D in the vertical direction. See the online edition of the *PASP* for a color version of this figure.

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

3.2. The Coronagraph Relieves the Wavefront Control

Figure 4 shows the DM voltmaps resulting from the speckle nulling algorithm for the Lyot coronagraph (*left panel*) and the PIAA coronagraph (*right panel*) configurations of SCExAO. Despite being ignorant of the pupil geometry, the speckle nulling produces a DM voltage-induced phase pattern that resembles the original telescope simulator pupil.

In addition to the spider vanes, in the Lyot configuration the algorithm attempts to cancel the diffraction rings induced by the sharp inner and outer edge of the pupil. The corresponding voltmap exhibits some sharp voltage changes from one actuator to its direct neighbor near the projection of these edges on the DM. Despite these sharp changes, this configuration fails to seriously cancel the innermost rings that would require more range than the DM can actually provide. The PIAA coronagraph voltmap, on the other hand, exhibits smoother features near the edge of the pupil, and only the effort to interfere with the speckle due to the spider vanes are very obvious. The coronagraph succeeds in relieving the wavefront control device by effectively suppressing a major fraction of the diffraction in the focal plane.

Note that the regions of the DM that obviously fall outside of the pupil do not need to be actuated at all. A more sophisticated DM control software should regularize the DM shape so as to preserve stroke on the DM.

3.3. Inner Working Angle

To further characterize the performance of the SCExAO coronagraph, and estimate its inner working angle, we look at the evolution of the throughput as a function of the position of the source relative to the focal plane across the field. While applying the DM voltage map resulting from a sequence of speckle nulling iterations, the source is translated along the horizontal axis. Figure 5 shows six snapshots of the PSF as it moves across the "dark hole" created by the speckle nulling loop for the first

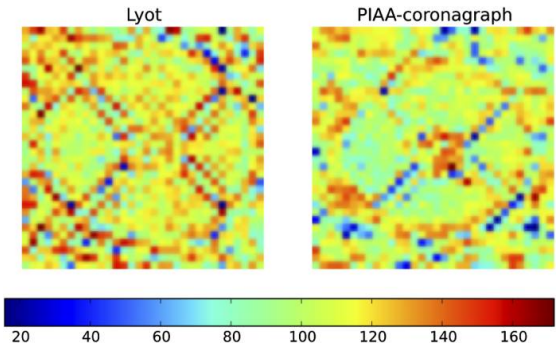


FIG. 4.—Speckle nulling voltmaps for configurations (1) Lyot coronagraph and (3) PIAA coronagraph of SCExAO (see Fig. 1). See the online edition of the *PASP* for a color version of this figure.

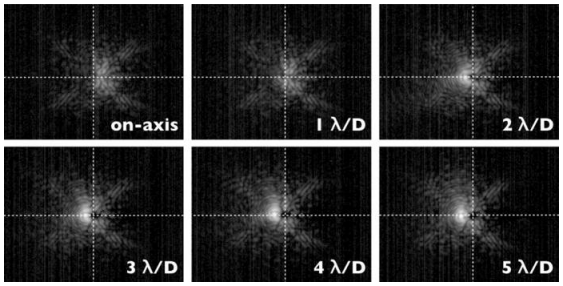


FIG. 5.—Evolution of the PSF as a function of off-axis position near the focal plane mask from on-axis to 5 λ/D off-axis. At 2 λ/D, the core of the PSF has almost entirely cleared the focal plane mask, and the diffraction rings associated with it become visible again. Beyond 3 λ/D, the PSF morphology is essentially translation-invariant. Images use a common logarithmic intensity scale to reveal both bright and faint features of the PSFs.

5 λ/D away from on-axis. Within 1 λ/D and a little beyond, ring-like structures remain erased by the coronagraph. At 2 λ/D, the core of the PSF clears the coronagraph and diffraction rings reappear in the image. Beyond 2 λ/D, the structure of the PSF is essentially translation invariant, and only the shadow of the focal plane mask betrays the presence of the coronagraph.

To complete this qualitative analysis of post coronagraph images, Figure 6 shows a curve of the system throughput, integrated over the left-hand side of the field, as the source is moved across the entire 14 λ/D control region and slightly beyond. The 100% throughput reference is determined on-axis by

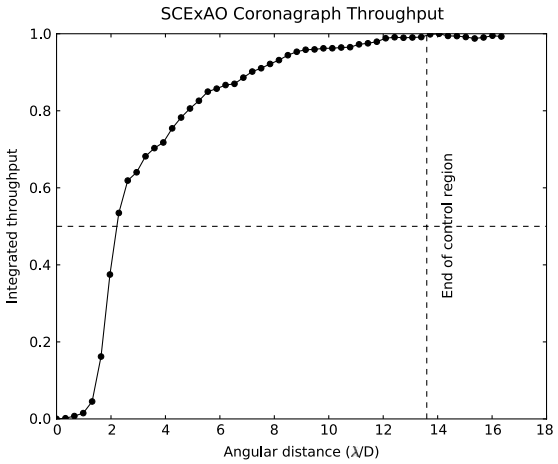


FIG. 6.—Throughput of the SCExAO coronagraph as a function of angular separation. Using an iterative speckle nulling algorithm, the voltage map that cancels speckles within the control region of the DM has been identified. This curve shows how the throughput of the coronagraph evolves as the source is progressively translated off-axis, across the entire 14 λ/D-wide control region.

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

removing the focal plane mask. Matching our qualitative description of snapshots in Figure 5, we confirm that within 1 λ/D, the throughput only varies at the percent level. The most rapid variation of throughput happens around 2 λ/D, where the PSF core clears the focal plane mask as shown in Figure 5: the throughput quickly rises from a few percent to 60%.

One convenient definition of the inner working angle (IWA) of a coronagraph is the angular separation at which its transmission reaches 50% (Guyon et al. 2006). According to the measurements summarized in Figure 6, SCExAO's current IWA is 2.2 λ/D. The focal plane mask currently used is 40% oversized, compared to the 1.6 λ/D IWA the PIAA optics have been manufactured for. It will therefore be replaced before SCExAO's next observing run.

4. CONCLUSION

We have, for the first time, demonstrated focal plane wavefront control capability in a complete PIAA—focal plane mask—inverse PIAA system. Using a simple iterative speckle nulling algorithm, we successfully produced a high-contrast region in the field (the so-called dark hole), and therefore confirmed that the complete set of remapping optics (PIAA and inverse PIAA) are transparent to the focal-plane-based wavefront control. We have also confirmed that the outer working angle allowable by the number of actuators of the deformable mirror is preserved during the double remapping. This implies that PIAA-type coronagraphs can be very well be used on ground- and space-based telescopes with no loss of outer working angle and do not require complex wavefront control algorithms. High-performance, high-efficiency coronagraphy therefore does not translate into increased system complexity.

Note that while we tested PIAA with a large mask imposing a 2.2 λ/D inner working angle, the conclusions of this article will hold for more IWA-optimized PIAA coronagraphs, including the PIAACMC (Guyon et al. 2010a) that exhibit IWA $<1\lambda/D$. Aggressive designs combining focal-plane-based wavefront control with what this article demonstrates make coronagraphy on board small telescopes (~2 m) in space a relevant option for the direct detection of extrasolar planets. They also offer the potential for imaging reflected-light planets at very small separation with the forthcoming generation of extremely large telescopes.

The results presented in this article do not address broadband operation of the coronagraph and control system, which is necessary for efficient on-sky operation. At the moderate raw contrast at which the system is designed to operate (~10^{-3} to 10^{-4} at a few λ/D), the PIAA lenses and relay optics are indeed designed to operate in a 20% wide band (H band, centered around 1.65 μm) in the absence of aberrations. Other issues that are not limited to SCExAO but are common to all high-contrast coronagraphs operating in broad band must nevertheless be carefully considered for broadband operation of the system:

1. Wavefront sensing in the focal plane must be robust against speckle elongation in broadband light, an issue that becomes important when attempting to measure the complex amplitude of speckles beyond ~5λ/D from the optical axis in a 20%-wide band. We note that in the near future, SCExAO will feed an integral field spectrograph, which is not limited by this effect and can therefore be used for broadband light focal plane wavefront sensing.
2. The focal plane mask size, if fixed, should ideally scale with wavelength. Our current 2.2 λ/D IWA design is compatible with the 20% band at the required contrast level.
3. Atmosphere-induced wavefront errors are chromatic, principally due to the differential atmospheric refraction and the chromatic diffractive propagation between turbulence layers. While this effect is of concern when using a visible wavefront sensor for correction in the near-IR, it does not limit the instrument performance when both sensing and correction are performed, like in speckle nulling, within the same near-IR band, where variations of the refractive index of air remain small.
4. Wavefront chromaticity due to optical defects is expected at some level in any system using refractive optics. These are believed to be sufficiently small to not affect the instrument performance with our current design, since the relative variation in refractive index for most optical glasses is ~0.3% across the 20% wide H band. Defects on the surface of refractive elements are expected produce a chromatic error which is 0.3% as large as the defect itself. Even if such defects limit the contrast to 10^{-2} without wavefront control (a conservative assumption for our system), the chromatic residual after wavefront correction will correspond to a $10^{-2} \times 0.03^2 \sim 10^{-7}$ contrast limit, which is well below our raw contrast goal.

The coronagraph currently in place in SCExAO does not take into account the spider vanes in the pupil, while these are responsible for most of the light observed in the final focal plane of the instrument. Consequently, the wavefront control algorithm must use pupil phase introduced by the deformable mirror to remove, over half of the focal plane, light diffracted by the spider vanes, as shown in Figure 4. This approach is simply not optimal in a broad spectral band, and the SCExAO coronagraph will be updated to a PIAACMC type coronagraph to solve this issue.

Finally, while the results presented in this article demonstrate the raw contrast required for ground-based system, it is unclear to which extent the approach adopted in this work is suitable for the ~10^{-9} contrast level that future space-based missions hope to reach in order to image and study Earth-like planets. While remapping propagation effects specific to the PIAA coronagraph can be included in the wavefront control algorithm (Krist et al. 2011), it is unknown to what extent such effects need to be accounted for. The approach presented in this article should be tested, both in simulations and in laboratory experiments, at higher contrasts than required for a ground-based system. This work will be conducted over the next 2 years on existing PIAA coronagraph testbeds.

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

REFERENCES

Belikov, R., Pluzhnik, E., Witteborn, F. C., et al. 2011, Proc. SPIE, 8151, 815102-815102-7

Beuzit, J.-L., Feldt, M., Mouillet, D., et al. 2010, in In the Spirit of Lyot 2010: Direct Detection of Exoplanets and Circumstellar Disks, ed. A. Boccaletti (Paris: University of Paris Diderot)

Bordé, P. J., & Traub, W. A. 2006, ApJ, 638, 488

Give'on, A. 2006, Ph.D. thesis, Princeton University

Guyon, O. 2003, A&A, 404, 379

———. 2005, ApJ, 629, 592

Guyon, O., Martinache, F., Belikov, R., & Soummer, R. 2010a, ApJS, 190, 220

Guyon, O., Pluzhnik, E. A., Galicher, R., et al. 2005, ApJ, 622, 744

Guyon, O., Pluzhnik, E., Martinache, F., et al. 2010b, PASP, 122, 71

Guyon, O., Pluzhnik, E. A., Kuchner, M. J., Collins, B., & Ridgway, S. T. 2006, ApJS, 167, 81

Guyon, O., & Roddier, F. 2002, A&A, 391, 379

Hodapp, K. W., Suzuki, R., Tamura, M., et al. 2008, Proc. SPIE, 7014, 701419-701419-12

Kern, B., Guyon, O., Give'On, A., Kuhnert, A., & Niessner, A. 2011, Proc. SPIE, 8151, 815104-815104-9

Krist, J. E., Belikov, R., Pueyo, L., et al. 2011, Proc. SPIE, 8151, 81510E

Lozi, J., Martinache, F., & Guyon, O. 2009, PASP, 121, 1232

Macintosh, B. A., Graham, J. R., Palmer, D. W., et al. 2008, Proc. SPIE, 7015, 701518-701518-13

Malbet, F., Yu, J. W., & Shao, M. 1995, PASP, 107, 386

Marois, C., Doyon, R., Nadeau, D., Racine, R., & Walker, G. A. H. 2003, in EAS Pub. Ser. 8, Astronomy with High Contrast Imaging, ed. C. Aime, & R. Soummer (Cambridge: Cambridge University Press), 233–243

Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348

Martinache, F., Guyon, O., Garrel, V., et al. 2011, Proc. SPIE, 8151, 81510Q-81510Q-10

Martinache, F., Guyon, O., Pluzhnik, E. A., Galicher, R., & Ridgway, S. T. 2006, ApJ, 639, 1129

Tamura, M. 2009, in AIP Conf. Proc. 1158, Exoplanets and Disks: Their Formation and Diversity (New York: AIP), 11

Tamura, M., & SEEDS Team 2010, in In the Spirit of Lyot 2010: Direct Detection of Exoplanets and Circumstellar Disks, ed. A. Boccaletti (Paris: University of Paris Diderot)

Vanderbei, R. J., & Traub, W. A. 2005, ApJ, 626, 1079

This content downloaded on Fri, 1 Feb 2013 22:30:00 PM
All use subject to JSTOR Terms and Conditions

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **126**:565–572, 2014 June
© 2014. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

On-Sky Speckle Nulling Demonstration at Small Angular Separation with SCExAO

FRANTZ MARTINACHE,[1,2] OLIVIER GUYON,[1,3] NEMANJA JOVANOVIC,[1] CHRISTOPHE CLERGEON,[1] GARIMA SINGH,[1] TOMOYUKI KUDO,[1] THAYNE CURRIE,[4] CHRISTIAN THALMANN,[5,6] MICHAEL MCELWAIN,[7] AND MOTOHIDE TAMURA[8]

Received 2013 September 16; accepted 2014 May 04; published 2014 June 11

ABSTRACT. This paper presents the first on-sky demonstration of speckle nulling, which was achieved at the Subaru Telescope in the context of the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) Project. Despite the absence of a high-order high-bandwidth closed-loop AO system, observations conducted with SCExAO show that even in poor-to-moderate observing conditions, speckle nulling can be used to suppress static and slow speckles even in the presence of a brighter dynamic speckle halo, suggesting that more advanced high-contrast imaging algorithms developed in the laboratory can be applied to ground-based systems.

Online material: color figures

1. INTRODUCTION

The detection of high-contrast features such as disks and companions in diffraction-limited images relies for the most part on the characterization and calibration of diffraction effects, whether static (induced by the geometry of the pupil and the aberrations in the optics), quasi-static (telescope pointing and slowly variable instrument alignment), or dynamic (atmosphere-induced). For ground-based observations, the best performance comes from a combination of adaptive optics (AO), which address the dominant atmospheric term, turning the seeing-limited image into a diffraction-limited one, and post-processing techniques based on differential imaging. These include spectral, polarimetric or angular differential imaging (ADI), the first two relying upon some properties of the target, such as the presence of specific spectral features or a polarimetric signature.

ADI (Marois et al. 2006; Lafrenière et al. 2007) provides the means to calibrate the static and quasi-static features of the PSF by observing a target with an alt-azimuthal telescope, for which the sky rotates relative to the telescope pupil. If enough field rotation occurs over a time that is less than the characteristic quasi-static aberration time-scale, the diversity between the orientation of the sky and the quasi-static point spread function (PSF) allows the calibration of the diffraction pattern, leading to greatly enhanced detection limits.

The calibration requires that the angular rotation induces sufficient local linear displacement of the genuine features, such as faint companions, relative to the quasi-static PSF (Marois et al. 2006) in order to avoid self-subtraction. While highly efficient at angular separations greater than 1″, it is difficult to benefit from this technique at angular separations smaller than 0.5″. For addressing such small angular separations, two approaches are possible: interferometric calibration and additional active wavefront control.

Interferometric calibration of diffraction-limited images is a nascent field that inherits from the ideas and techniques of sparse aperture masking (SAM), a.k.a., nonredundant masking (NRM) interferometry with AO (Tuthill et al. 2006). The fundamental idea behind these techniques is that although the content of images may be corrupted by residual aberrations, it is possible to construct a subset of information that is independent from these aberrations. One possible such subset of information is the kernel-phase (Martinache 2010) which, just like sparse aperture masking using closure phase, goes as far as enabling the direct detection of high contrast features beyond the diffraction limit, a regime referred to as super-resolution (Pope et al. 2013). The fundamental advantage of this approach is that it requires no additional differential "trick" like ADI, although it is for now restricted to well-corrected AO images. Recent improvements of the technique, such as the statistically independent kernel-phase and better calibration procedures

[1] National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720, frantz@naoj.org.
[2] Laboratoire Lagrange, UMR7293, Université de Nice Sophia-Antipolis, CNRS, Observatoire de la Côte d'Azur.
[3] Stewart Observatory, University of Arizona, Tucson, AZ 85721.
[4] Department of Astronomy and Astrophysics, University of Toronto, Canada.
[5] Astronomical Institute "Anton Pannekoek", University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands.
[6] Department for Astronomy, ETH Zurich, CH-8093 Zurich, Switzerland.
[7] Goddard Space Flight Center, Greenbelt, MD.
[8] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan.

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

described by Ireland (2013), show promise of performance that compares to what is currently achieved by ADI at small angular separations.

The alternative approach to the passive calibration by interferometry is to employ additional wavefront control to modulate the diffraction at small angular separation and create the diversity that enables the separation of genuine structure from diffraction features in the image. This concept, which can be referred to as coherence differential imaging (CDI), is very relevant to the upcoming generation of so called extreme AO (XAO) projects: Palm-3000/P1640 at Palomar (Oppenheimer et al. 2012), GPI at Gemini (Macintosh et al. 2008), SPHERE at VLT (Sauvage et al. 2010), and SCExAO at the Subaru Telescope (Martinache et al. 2011).

Among these projects, the Subaru Coronagraphic Extreme AO (SCExAO) project is unique with its incremental deployment schedule that enabled the on-sky testing of fundamental features of the instrument (Jovanovic et al., in preparation) prior to completion of its closed-loop XAO system (Clergeon et al. 2013). In this paper, we present the first convincing demonstration of on-sky closed-loop diffraction modulation, after an upstream AO correction is performed, providing a Strehl on the order of 20%. While the demonstrated gain is moderate and there is room for improvement, the paper nevertheless demonstrates that it is possible to control the features of a coronagraphic PSF near the inner working angle of the coronagraph using the speckle nulling approach introduced in previous work.

2. SCEXAO BEFORE XAO

What is usually referred to as an XAO system is a unit capable of producing a very stable and high-quality PSF, characterized by a Strehl ratio on the order of 90%. This very high level of Strehl, corresponding to residual wavefront errors on the order of $\lambda/20$ (where λ is the wavelength of observation), is indeed required for the high-contrast technique of coronagraphy, which aims to block the light of an on-axis bright source and reveal its circumstellar environment, in particular disk and/or faint companions.

Particularly in monochromatic light, it is possible to design coronagraphs that provide near-perfect extinction, and an inner working angle (IWA) that approaches the limit of diffraction (λ/D), where D represents the diameter of the telescope aperture. Examples of such coronagraphs include the vortex (Mawet et al. 2010) and the PIAA (Guyon 2003), both implemented on SCExAO. In practice, however, their performance is entirely dictated by the residual wavefront errors. Under poor Strehl conditions, the coronagraph does little more than allow the exposure time to be increased before saturation of the detector.

SCExAO does not currently support a high-order closed-loop AO system, expected to be first commissioned in 2014 fall. In this first phase of the project, SCExAO therefore relies entirely on the upstream Subaru Telescope facility AO system called AO188 (Minowa et al. 2010). AO188 and the coronagraphic imager HiCIAO (Hodapp et al. 2008) form the workhorse of the Strategic Exploration of Exoplanets and Disks with Subaru (SEEDS) observing campaign (Tamura 2009). SCExAO was conceived as a replacement upgrade for the current fore-optics used by HiCIAO, that should include a small IWA PIAA-based coronagraph as well as additional wavefront manipulation capability.

Therefore, while SCExAO does not currently qualify as an XAO system, it nevertheless already implements a 1-k actuator deformable mirror (DM). We have recently shown, using a calibration source in a stable laboratory environment, that the DM and the PIAA coronagraph used together can produce a high-contrast region in the field-of-view, with a 2.2 λ/D IWA (Martinache et al. 2012), with a simple iterative speckle-nulling loop. Using the same approach on-sky, relying on AO188 to produce a 20% Strehl ratio PSF, we demonstrate that it is possible to use speckle nulling to drive down the brightness of the diffraction features at small angular separation. We also show that the DM provides a flexible calibration tool that notably enables astrometric calibration as well as direct measurements of the contrast in images.

3. ON-SKY CORONAGRAPHIC CALIBRATION

Figure 1 provides a schematic overview of the optical layout of the near-IR arm of SCExAO, as it was used on the observing night of 2012 November 12. The upstream AO188 system feeds the instrument with a partially corrected $f/14$ convergent beam, intercepted by a first tip-tilt controlled mirror located in an image plane and mounted on a focusing stage, and further steered by a second tip-tilt controlled mirror. This combination provides full control of the origin and incidence of the beam entering the coronagraph. The other important feature is the focal plane mask, reflecting the light of any on-axis source onto a camera used to sense the pointing, a subsystem called the coronagraphic low-order wavefront sensor (CLOWFS), as described by Guyon et al. (2009) and Vogt et al. (2011).

In addition, the system also comprises the PIAA remapping optics (Guyon et al. 2005) located before the focal plane mask and their inverse counterpart located after the focal plane mask. Recent work (Martinache et al. 2012) showed that when used together, the presence of these two sets of remapping optics can be ignored by the focal plane based wavefront control algorithm. Reflecting upon this conclusion, this discussion will ignore the remapping optics, which for the results presented in this paper, are always in the beam.

One of the two tip-tilt mirrors that steers the beam entering the coronagraph is the 1024-actuator DM itself which, while not exactly located in a pupil plane, can nevertheless be used to introduce/compensate for diffraction features in the coronagraphic image. The square grid geometry of the actuators underneath the DM membrane makes it particularly well-suited to generate

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

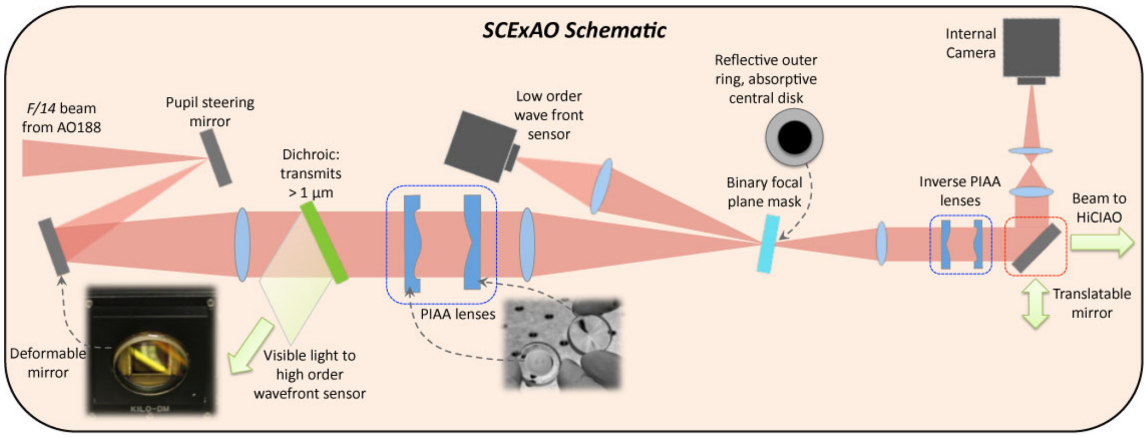


FIG. 1.—Schematic representation of the optics inside the SCExAO IR coronagraph used to produce the results presented in this paper. See the electronic edition of the *PASP* for a color version of this figure.

pairs of speckles that interfere with already present diffraction features in the coronagraphic image. The total number of actuators across the telescope pupil footprint imposes the size of the region of the coronagraphic image that can be probed by the DM, which will be referred to as the control region. As explained by Martinache et al. (2012), the control region with this current implementation fits within a $27.2 \times 24.8\ \lambda/D$ rectangular box. It is possible to generate a pair of speckles centered around the optical axis within this region by applying a sinusoidal displacement map on the DM.

Each such sinusoidal displacement map on the DM is characterized by four numbers: two cartesian coordinates of spatial frequency (k_x, k_y), one amplitude α (in radians) limited by the stroke of the DM, and a phase φ that can take any value between 0 and 2π. When applying a sinusoidal displacement map, the DM acts like a diffraction grating that creates off-axis copies of the on-axis bright source. Figure 2 illustrates the relation between DM modulation and speckles in the image. The larger the spatial frequency parameters, the faster the sinusoidal modulation of the DM surface and the further the corresponding pair of speckles lies away from the center of the control region. The larger the amplitude of the modulation, the brighter the resulting speckles. The phase does not change the intensity of the speckles unless they interfere with other coherent features in the image. Sinusoidal modulations of distinct properties can be added to the DM as long as the total requested range of displacement does not go beyond the DM stroke, creating in turn multiple pairs of speckles in the image.

The ability to produce symmetric pairs of speckles of controllable positions and contrast is a very convenient feature that facilitates otherwise difficult calibration procedures when observing with a coronagraph. Indeed, unlike classical imaging techniques and interferometric methods for which absolute control of the pointing is not a strict requirement (since the central star remains visible), the image of the star needs to be precisely driven and stabilized on the axis of the coronagraph for an efficient suppression.

Sivaramakrishnan & Oppenheimer (2006) astutely proposed to address this issue by inserting a reticulate grid of wires in the pupil plane that produces a predefined periodic diffraction pattern in the image. The grid geometry of deformable mirrors such as the 1k-actuator MEMS DM used in SCExAO can advantageously be used as an adaptive version of this reticulate grid. Just as described by Sivaramakrishnan & Oppenheimer (2006), the added satellite speckles provide a very reliable way to position the star behind the focal plane mask during target acqusition. Once in the ideal position, the low-order wavefront sensor (Guyon et al. 2009) keeps the image in a fixed position relative to the mask. During the post-processing stage, in case noncoherent structures such as companions are identified in the image, the satellite speckles offer a reliable astrometric reference.

In addition to offering astrometric and pointing aids (Hinkley et al. 2013), the generated speckles can be used for contrast calibration purposes. This can efficiently be achieved by an offline predetermination of the satellite speckle contrast relative to the central PSF as a function of spatial frequency and amplitude. Figure 3 shows one example of such on-sky contrast calibrated images, recorded by the coronagraphic imager HiCIAO fed light by SCExAO. The top panel of Figure 3 shows a nonsaturated low-exposure time image with the SCExAO focal plane mask out of the way. A sinusoidal modulation of the SCExAO DM of known characteristics creates a pair of speckles of controlled contrast. The bottom panel shows an image acquired with the same exposure time, this time with the SCExAO focal plane mask placed on-axis. Except at small angular separation, one can observe very little impact of the focal plane mask on the halo PSF features.

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

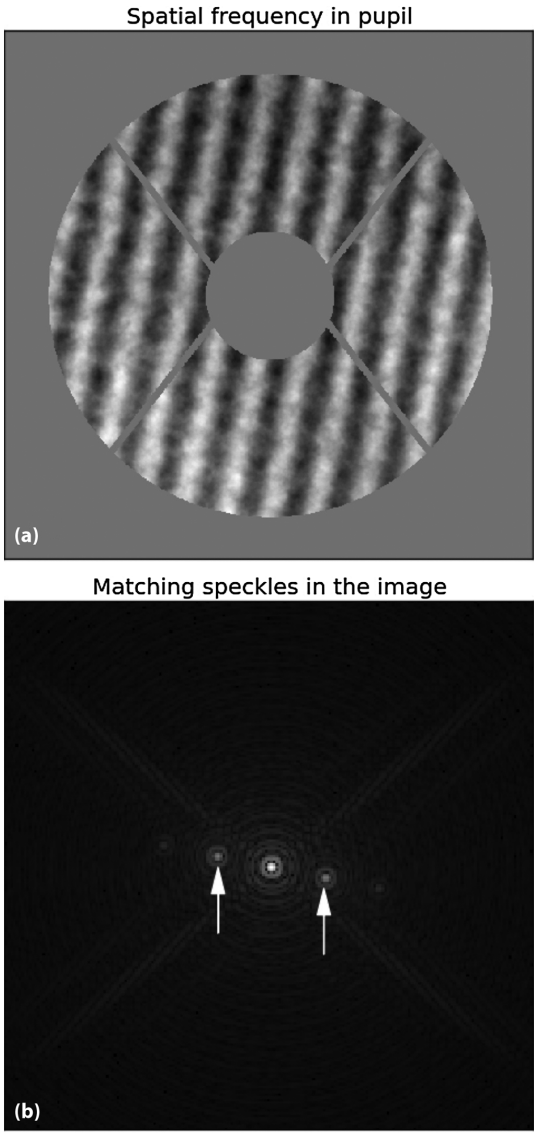


FIG. 2.—Illustration of the relation between sinusoidal modulation of the DM shape in the pupil and the resulting diffraction pattern in the image. In panel (*a*), one can count approximately 10 cycles of the sinusoid across the entire pupil. In the image (*b*), this modulation of the DM creates a pair of speckles (pointed at by white arrows) located 10 λ/D away from the center of the field. These additional speckles interfere with the already present diffraction features such as rings, spikes, and phase defects. However, they do not interfere with incoherent structures like planets and disks.

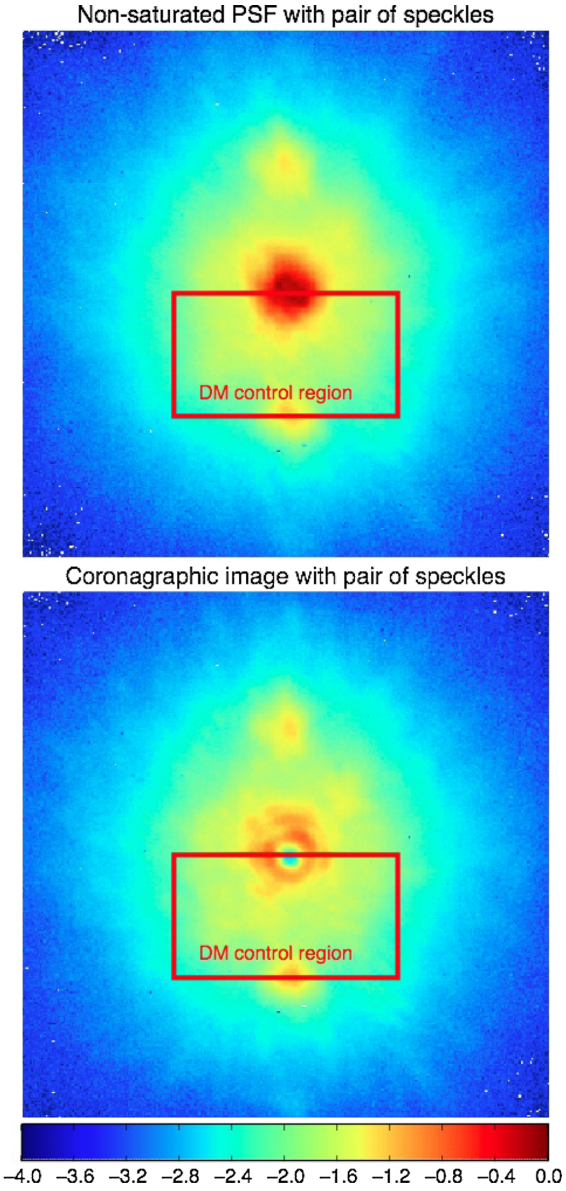


FIG. 3.—Comparison of a noncoronagraphic (*top panel*) and a coronagraphic (*bottom panel*) image acquired by HiCIAO behind SCExAO, with an added pair of speckles at max angular separation (log contrast scale). The red box highlights the area controlled by the DM during the experiments reported in this paper. Added speckles like these enable the calibration of the image in contrast, even if the central star is blocked by the coronagraph. See the electronic edition of the *PASP* for a color version of this figure.

4. ON-SKY SPECKLE NULLING

Martinache et al. (2012) showed, using SCExAO under stable laboratory conditions, that it is possible to actuate the DM before the PIAA coronagraph to create diffraction features that destructively interfere with speckles in the field, creating

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

a uniform high contrast region, often referred to as a "dark hole" in the image. Here, we apply the same iterative speckle nulling algorithm during actual on-sky observations. The major difference with the laboratory scenario is that the static (or quasi-static) features of the PSF probed by the speckle nulling algorithm are buried underneath a very strong, fast-varying dynamic component, due to the incomplete AO correction in the absence of a properly closed XAO loop.

The reported performance of ADI calibration suggests that high-contrast detection limits are set by long-lived aberrations, with a characteristic time-scale ∼1 hr or less. In this experiment, the science detector (HiCIAO) and the speckle-nulling camera cannot simultaneously observe: a retractable mirror sends the light either way. To be useful, the speckle nulling loop must therefore converge within a time-scale on the order of 15–20 minutes, so that sufficient time is left for the acquisition of frames by the (slow) science detector, before the quasi-static aberrations start changing again.

4.1. Technical Constraints

This on-sky speckle nulling experiment was performed using an uncooled InAsGa type detector (Xeva XS 1.7-320, sold by Xenics), and an H-broadband filter ($1.6\ \mu$m). Such cameras exhibit frame rates up to ∼100 Hz that make them appealing for NIR wavefront control applications; however, they also have fairly high readout noise ($>300\ e^-$) as well as high dark current. In the high-gain setting, the dark current averages 10000 counts in 20 ms exposure (almost the entire usable dynamic range). Coronagraphic observations with these cameras are therefore limited to very bright stars only.

The images showed here were obtained while observing the bright K-type star Pollux ($V = 1.15$, $R = 0.6$, $H = -0.845$). For this object, the best compromise was to use a 10-ms exposure time in low-gain mode, with a median dark current level ∼4200. While the speckle nulling loop is running, the camera continuously acquires 10-ms exposures at the frame rate of 30 Hz.

In a given image, up to n speckles are identified within the control region and their positions marked, relative to the central source, hidden by the coronagraph, but nevertheless located by the satellite speckles as described in section 3. The first step is to identify the two-component spatial frequency (k_x, k_y) of the DM that corresponds to each of the speckles identified inside the control region (cf. Fig. 2). The amplitude α_0 of this spatial frequency is estimated from the speckle brightness, proportional to the square of the amplitude α_0. The only remaining unknown is the phase φ of the speckle, which can take any value between 0 and 2π.

To determine this unknown, one modulates the DM with a sinusoidal function so as to generate a speckle probe of constant amplitude α_0, but whose phase is varied from 0 to 2π radians from image-to-image. By tracking the evolution of the resulting speckle brightness as a function of the probe phase, one is able to determine the true phase φ_0 of the original speckle. Once the speckle phase is determined, a correction of opposite phase $\varphi_0 + \pi$ and of amplitude $g \times a_0$ (with $g \sim 0.1$ the loop gain), is permanently applied before the loop is allowed to hunt for other speckles. A minimum of four probes with phase 0, $\pi/2$, π, and $3\pi/2$ is required to unambiguously estimate a long-lived quasi-static speckle phase. Unfortunately, the 10-ms exposure time imposed by the characteristics of the detector happens to be of the order of the time scale of the AO188 wavefront correction. To make the speckle nulling loop robust against fast fluctuations associated to the dynamic speckle component, the algorithm uses 30 probes of phase uniformly varying between 0 and 2π.

After the speckle nulling loop has converged, a mirror located in a collimated beam after the coronagraph (cf. Fig. 1) is translated out of the optical path so as to direct the light toward the HiCIAO imager.

Given the chosen number of probes, frame rate, and processing time, we get one iteration completed in 4 s. We increase the efficiency of the loop by simultaneously probing as many speckles as possible. However, the image-to-image variance due to the dynamic aberrations and the poor sensitivity of the camera set a limit to the total number of speckles that can be probed. In practice, the algorithm tested on-sky was able to simultaneously probe an average of five of the brightest speckles in the field.

4.2. Speckle Nulling Data

Twelve minutes of the speckle nulling loop running were sufficient to allow us to visually observe a darkening of one side of the field of view. To test the efficiency of the wavefront modulation resulting from the speckle nulling loop, several series of images were acquired with HiCIAO, alternating between a "flat" and an "optimized" SCExAO DM shape. The flat DM shape resulted from an optimization routine maximizing the Strehl in the noncoronagraphic image using Zernike modes: astigmatism, focus, coma, spherical, trefoil, and quadrifoil on an internal calibration source. The optimized shape was the result of speckle nulling corrections applied in addition to the flat. The complete acquisition sequence consisted of four such cycles (to make the analysis robust against slow trend effects), resulting in a total of twenty 5-second exposure frames per DM shape. These HiCIAO images were dark-subtracted and flat-fielded, using standard data reduction procedures established for the reduction of the SEEDS data, assembled into two data cubes.

The images gathered in Figure 4 summarize the statistical properties of these two datacubes, looking at the median (*left column*) and the standard deviation (*right column*) for the flat (*top row*) as well as the optimized (*bottom row*) DM shape, using common colorbars. For each figure, the extent of the control region is highlighted by a red box, which, for this experiment, is located in the lower part of the HiCIAO image.

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

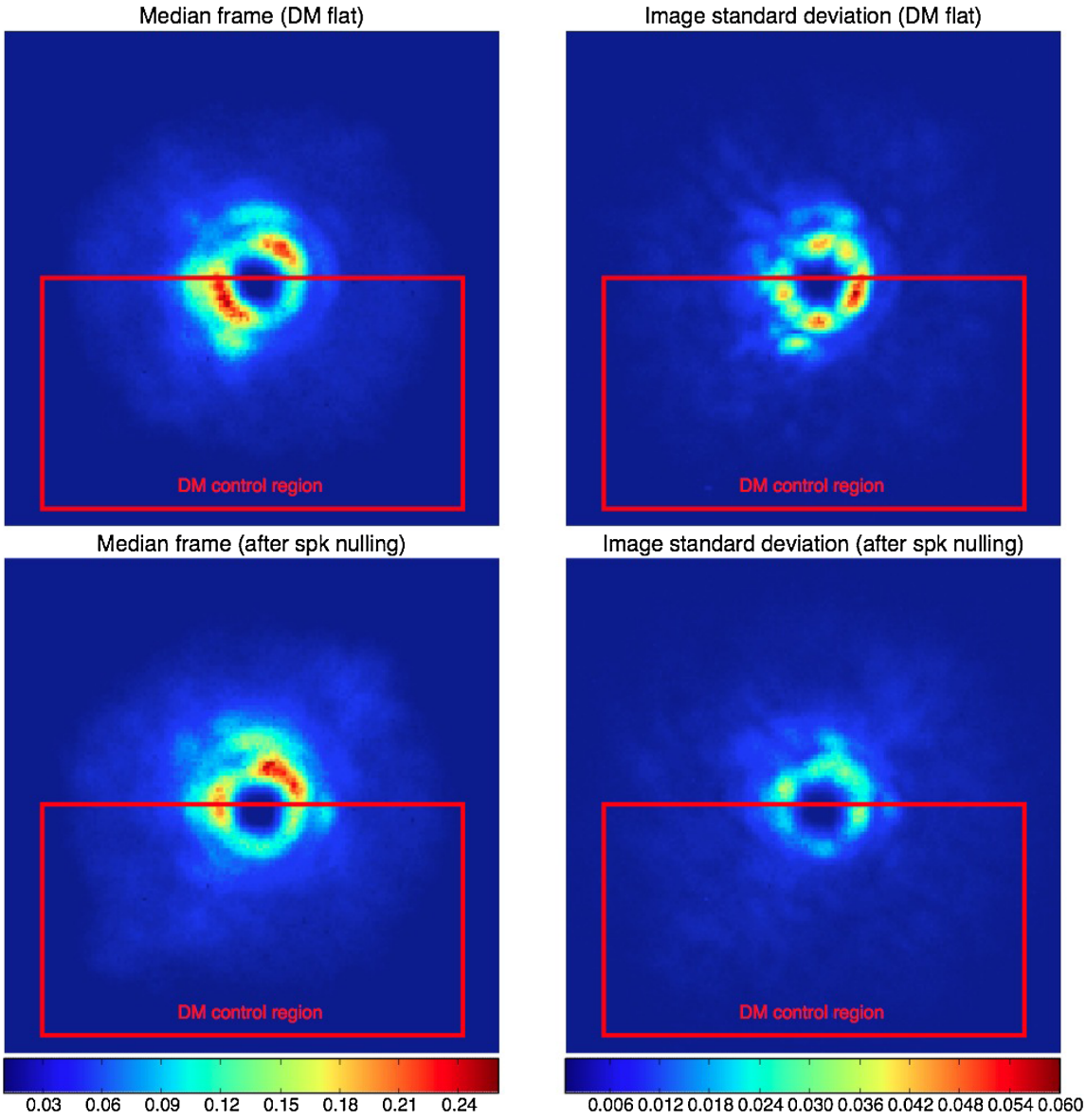


FIG. 4.—Statistical properties (median and standard deviation) of coronagraphic images acquired by HiCIAO for two SCExAO DM shapes. The top row is for a flat DM shape that maximizes the Strehl ratio. The bottom row is for a DM shape resulting from a speckle nulling loop. The colorbar scale used for the images on the left (median) is calibrated in contrast. See the electronic edition of the *PASP* for a color version of this figure.

The combination of exposure time (5 s) and neutral density (ND0.1) for HiCIAO was chosen to acquire nonsaturated images in the H-broadband filter. The low and fast varying seeing experienced during the acquisition and the lack of a yet to come high-order close-loop adaptive correction results in exposures that exhibit no obvious individual speckles in the field, a major difference from what is already achieved by XAO equiped instruments (see, e.g., Fig. 1 of Oppenheimer et al. [2013]). The first bright ring, located at a radius of 2.3 λ/D, is the dominating feature in the images. Note that this radius is very close to the quoted IWA of SCExAO (2.2 λ/D), defined as the angular separation for which the coronagraph throughput reaches 50%

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

(Martinache et al. 2012), and imposed by the focal plane mask. The standard deviation of the DM flat datacube shows that most of the speckle variance is localised onto the diffraction rings.

The speckle nulling algorithm unsurprisingly focused its efforts onto this area of the image. With the optimized DM shape (bottom images in Fig. 4), the median flux for the part of the diffraction ring overlapping with the control region is reduced. The improvement in regards to the median is moderate as the brightest parts of the ring are attenuated by a factor of 2.

4.3. Benefit for High Contrast Imaging

The detection of faint structures in a speckle-noise dominated problem is usually achieved by some form of PSF subtraction. The PSF can be obtained from more or less sophisticated methods: from the direct observation of a reference star to a synthetic PSF obtained after ADI. PSF subtraction is obviously appropriate to calibrate static features of the PSF. It is also good for fast-varying speckles that average down quickly to a smooth halo.

However, PSF subtraction fails to calibrate speckles characterized by a timescale of a few minutes which manifest by residual variance in PSF subtracted images: the speckle noise. The top row of Figure 4 illustrates this phenomenon and shows that, over the course of the twelve minutes covered by the experiment, on the most prominent image features the fluctuations are on the order of 25–50% of the local median value, and that the overall structure of the standard deviation image follows that of the median; static aberrations do appear to mix with and amplify the local speckle noise, resulting in speckles pinned on the static diffraction pattern (Aime & Soummer 2004).

In actively suppressing the static features of the coronagraphic PSF (cf. bottom row of Fig. 4), speckle nulling achieves two goals: it removes slow and static speckles (measurable by comparing the flat and optimized median frames) and reduces the speckle noise (measurable by comparing the standard deviation frames). The PSF mean level over the region corrected by speckle nulling is not only lower by a factor of 2, it is also more stable, with a standard deviation reduced by a factor of 3.

5. CONCLUSION

This paper presents the first demonstration of speckle nulling performed on-sky with partial AO correction. In the context of the SCExAO project, using a PIAA coronagraph and a deformable mirror to modulate the focal plane, the results presented here show that it is indeed possible to complement ADI and its variants at small angular separations, even in the absence of a high strehl AO system, and therefore in the presence of a brighter dynamic speckle halo.

Speckle nulling offers improvement at two levels: (1) it removes slow speckles, resulting in images with enhanced contrast before processing, and (2) it reduces the speckle noise amplification by the static diffraction, resulting in enhanced post-processing detection limits. Moreover, it enables this in a regime of angular separation that cannot be addressed by ADI, which remains the reference technique (Brandt et al. 2013) for the calibration of ground based AO PSF. In spite of unfavorable observing conditions, the control loop managed to remain stable and converged to improve the statistical properties of the coronagraphic image.

While consistent, the reported improvement remains modest, and this can in part be attributed to the choices of integration time, constrained by the noise properties of the detector used during the speckle nulling loop. Improving the sensitivity and the dynamical range of the image by using a cooled detector will increase the flexibility of the technique and allow a more efficient use of the observing time. The real performance improvement, however, will come from the use of an actual close-loop XAO system that will improve the raw Strehl of individual images.

Advanced focal plane based wavefront techniques such as electric field conjugation (Give'On 2006; Bordé & Traub 2006) have been used extensively on laboratory high contrast test-beds over the world, including with a PIAA coronagraph like the one used on SCExAO (Guyon et al. 2010) providing access to very high contrast (10^{-7} and beyond) detection limits in a stable environment. The fact that, even in difficult observing conditions, a speckle nulling driven wavefront control algorithm originally envisioned for stable test-benches remains operational is very encouraging. This indeed suggests that, with only a few adaptations, all of the high contrast laboratory work that has so far been geared toward space borne applications is relevant to ground based observations, even with partial AO correction.

REFERENCES

Aime, C., & Soummer, R. 2004, ApJ, 612, L 85
Bordé, P. J., & Traub, W. A. 2006, ApJ, 638, 488
Brandt, T. D., McElwain, M. W., Turner, E. L., Abe, L., Brandner, W., Carson, J., Egner, S., Feldt, M., et al. 2013, ApJ, 764, 183
Clergeon, C., Guyon, O., Martinache, F., Veran, J., Correia, C., Garrel, V., & Jovanovic, N. 2013, American Astronomical Society Meeting Abstracts, 221
Give'On, A. 2006, PhD thesis, Princeton University
Guyon, O. 2003, A&A, 404, 379
Guyon, O., Matsuo, T., & Angel, R. 2009, ApJ, 693, 75
Guyon, O., Pluzhnik, E. A., Galicher, R., Martinache, F., Ridgway, S. T., & Woodruff, R. A. 2005, ApJ, 622, 744
Guyon, O., Pluzhnik, E., Martinache, F., Totems, J., Tanaka, S., Matsuo, T., Blain, C., & Belikov, R. 2010, PASP, 122, 71
Hinkley, S., Pueyo, L., Faherty, J. K., Oppenheimer, B. R., Mamajek, E. E., Kraus, A. L., Rice, E. L., Ireland, M. J., et al. 2013, ApJ, 779, 153

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

Hodapp, K. W., Suzuki, R., Tamura, M., Abe, L., Suto, H., Kandori, R., Morino, J., Nishimura, T., et al. 2008, Proc. SPIE, 7014
Ireland, M. J. 2013, MNRAS, 433, 1718
Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770
Macintosh, B. A., Graham, J. R., Palmer, D. W., Doyon, R., Dunn, J., Gavel, D. T., Larkin, J., Oppenheimer, B., et al. 2008, Proc. SPIE, 7015
Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556
Martinache, F. 2010, ApJ, 724, 464
Martinache, F., Guyon, O., Clergeon, C., & Blain, C. 2012, PASP, 124, 1288
Martinache, F., Guyon, O., Garrel, V., Clergeon, C., Groff, T., Stewart, P., Russell, R., & Blain, C. 2011, Proc. SPIE, 8151
Mawet, D., Serabyn, E., Liewer, K., Burruss, R., Hickey, J., & Shemo, D. 2010, ApJ, 709, 53
Minowa, Y., Hayano, Y., Oya, S., Watanabe, M., Hattori, M., Guyon, O., Egner, S., Saito, Y., et al. 2010, Proc. SPIE, 7736
Oppenheimer, B. R., Baranec, C., Beichman, C., Brenner, D., Burruss, R., Cady, E., Crepp, J. R., Dekany, R., et al. 2013, ApJ, 768, 24
Oppenheimer, B. R., Beichman, C., Brenner, D., Burruss, R., Cady, E., Crepp, J., Hillenbrand, L., Hinkley, S., et al. 2012, Proc. SPIE, 8447
Pope, B., Martinache, F., & Tuthill, P. 2013, ApJ, 767, 110
Sauvage, J.-F., Fusco, T., Petit, C., Meimon, S., Fedrigo, E., Suarez Valles, M., Kasper, M., Hubin, N., et al. 2010, Proc. SPIE, 7736
Sivaramakrishnan, A., & Oppenheimer, B. R. 2006, ApJ, 647, 620
Tamura, M. 2009, in AIP Conf. Proc. 1158, Exoplanets and Disks: Their Formation and Diversity, ed. T. Usuda, M. Tamura, & M. Ishii (New York: AIP), 11–16
Tuthill, P., Lloyd, J., Ireland, M., Martinache, F., Monnier, J., Woodruff, H., ten Brummelaar, T., Turner, N., et al. 2006, Proc. SPIE, 6272
Vogt, F. P. A., Martinache, F., Guyon, O., Yoshikawa, T., Yokochi, K., Garrel, V., & Matsuo, T. 2011, PASP, 123, 1434

This content downloaded from 192.54.176.192 on Wed, 06 Jan 2016 10:58:07 UTC
All use subject to JSTOR Terms and Conditions

Speckle nulling sans coronographe

L'article présentant une vue d'ensemble des fonctionalités de l'instrument SCExAO introduit plus haut dans ce mémoire inclut un autre exemple d'utilisation de cette boucle de contrôle, avec un autre détecteur InGaAs plus récent (Axiom Optics OWL SW1.7HS) proposant un bruit de lecture qui bien que toujours élevé (114 e^-) est un peu plus favorable. Le système bénéficie depuis d'un miroir déformable avec plus d'actionneurs ($\sim$50 dans un diamètre de pupille), positionné dans un plan conjugué avec la pupille du télescope. Cette fois ci sans coronographe, les figures 42 (source interne de calibration) et 43 (étoile brillante) permettent d'apprécier le potentiel de la méthode si elle était mise en oeuvre avec un détecteur combinant vitesse et sensibilité.

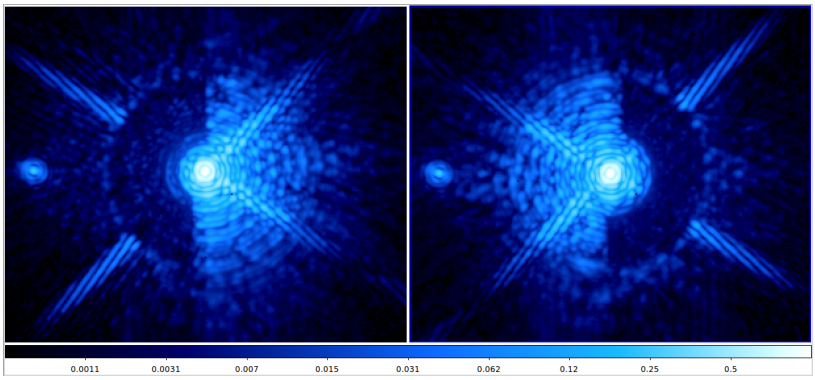


FIGURE 42: Exemples par SCExAO (bande H) de deux résultats de l'application la boucle de contrôle de speckle nulling sur la source interne de calibration (laser super-continuum). La région de contrôle est consituée d'un demi-disque, excluant la région centrale du champ. L'utilisateur peut choisir de sonder n'importe lequel des quatre quadrants (Nord - Sud - Est - Ouest) du champ.

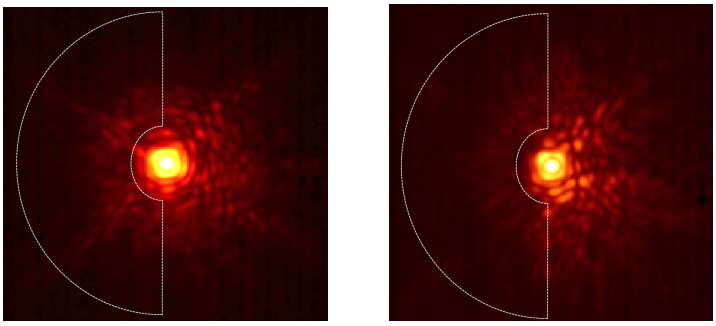


FIGURE 43: Exemple par SCExAO (bande H) de résultat de l'application de la boucle de speckle nulling sur l'étoile RX Boo (observations du 2 juin 2014). Chaque image est un composite de 5000 acquisitions pour un temps de pose de 50 μs. A gauche, RX Boo avant mise en oeuvre du speckle nulling. A droite, RX Boo, après speckle nulling. La région de contrôle est mise en évidence par la ligne pointillée.

Superflat

Dans l'exemple présenté en figure 42, l'algorithme de *speckle nulling* sert essentiellement à corriger les speckles induits par la géométrie de la pupille : aigrettes et anneaux de diffraction respectivement induites par les araignées et l'ouverture circulaire obstruée du télescope. Cependant, les speckles induits par des erreurs résiduelles de phase présentes dans la pupille sont également corrigés. L'algorithme de *speckle nulling* ne dispose d'aucune information lui permettant de discriminer les speckles selon leur origine. On peut malgré

tout au moins écrire la correction globale C^R_{DH} appliquée sur le miroir déformable pour créer un *dark-hole* à droite [54] comme la somme de deux types de correction de phase C^R_P et d'amplitude C^R_A :

$$C^R_{DH} = C^R_A + C^R_P \tag{22}$$

Même si la boucle ne peut pas distinguer ces deux contributions, les speckles induits par la géométrie de la pupille (qu'on désignera sous le nom de "speckles d'amplitude") présentent cependant une caractéristique qui les différencient des speckles résultants d'erreurs de phase (qu'on appellera "speckles de phase") : le schéma global de répartition des speckles d'amplitude est pair. On peut le justifier grâce aux propriétés de parité de la transformée de Fourier de la fonction réelle décrivant l'ouverture. En l'absence d'aberrations, à chaque speckle dans la région de contrôle correspond par conséquent un autre speckle de même amplitude et phase. Cette propriété est manifeste dans l'allure des images présentées dans la figure 42 : l'algorithme de speckle nulling, en atténuant les speckles d'un côté du champ, les amplifie de l'autre côté.

54. C^R_{DH} (right) permet de distinguer le dark-hole du côté droit de celui du côté gauche noté C^L_{DH} (left)

Pour les speckles de phase, l'application du speckle nulling est une compensation directe de la source de l'aberration : l'énergie initialement contenue dans le speckle est uniformément redistribuée dans toute la PSF dont l'intensité augmente légèrement.

Comme montré plus haut, il est possible de créer un *dark-hole* des deux côtés du champ. Les propriétés de symétrie font que la correction appliquée aux speckles d'amplitude pour le *dark-hole* de droite est l'opposé de la correction à appliquer pour les speckles d'amplitude pour le *dark-hole* de gauche : $C^R_{DH} = -C^L_{DH}$.

En présence d'erreurs résiduelles de phase, chacune des corrections appliquées pour créer ces *dark-holes* contient une estimation de la correction de phase, déterminée après analyse du quadrant correspondant de l'image. Les deux corrections peuvent donc être combinées : la moyenne de ces deux corrections contient une estimation de la phase résiduelle instrumentale :

$$C_P = 1/2 * (C^R_{DH} + C^L_{DH}) \tag{23}$$

Cette technique est utilisée sur SCExAO pour produire une compensation de toutes les erreurs cumulées de phase instrumentale jusqu'au détecteur scientifique : c'est une calibration que j'appelle le **superflat**. Il vient en complément de la compensation des bas-ordres apportée par d'autres méthodes : ZAP, qui sera introduit dans la dernière partie de ce mémoire, et le LOWFS qui a été présenté plus haut. Le **superflat** est une calibration très utile, un produit secondaire de l'algorithme de *speckle nulling* lorsqu'il est mis en oeuvre en l'absence de coronographe, et qui permet de mesurer (et donc de compenser) efficacement l'erreur de chemin optique non-commune (NCPA pour non-common path aberration) pour des fréquences spatiales comprises entre $\sim$4 et 25 λ/D. Un exemple d'image produit par SCExAO

après application du **superflat** est présenté dans la figure 44. Cette image a servi de base à la création du logo de SCExAO (figure 45).

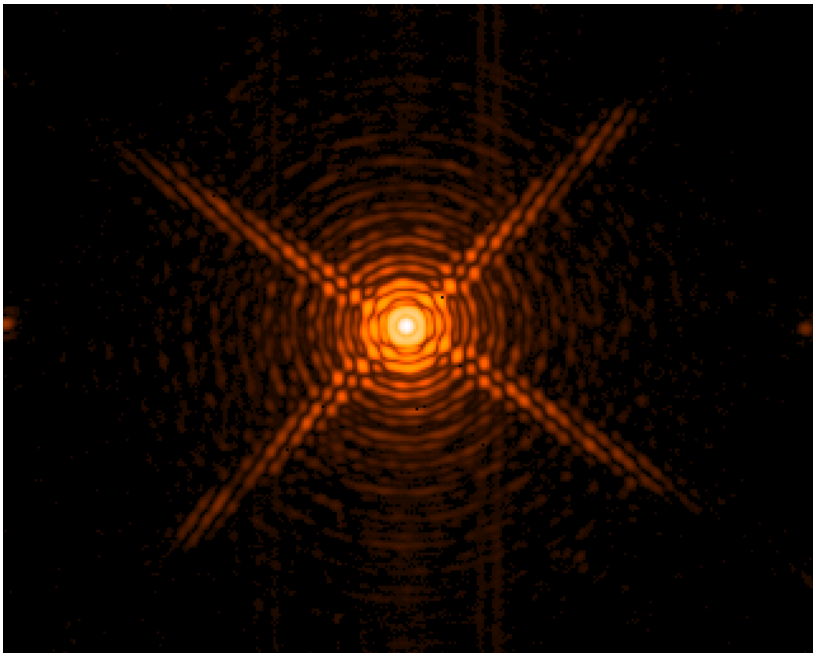


FIGURE 44: Résultat de l'application du superflat. Une nouvelle figure sera faite en deux parties : l'image de départ et l'image après correction, et contiendra une estimation de la correction (en nm) apportée par la technique.



FIGURE 45: Logo de l'instrument SCExAO, accompagné d'un de ses mottos : "What could go wrong ?" résumant l'esprit de créativité décomplexée qui caractérise ce projet.

Constraste et qualité de surface d'onde

La formule 12 reliant, dans le cas des faibles aberrations, le contraste d'un speckle à l'amplitude de l'aberration à son origine utilisée plus tôt dans une boucle de contrôle de front d'onde, peut être inversée pour donner une idée de la qualité de surface d'onde nécessaire à garantir pour arriver à un niveau de contraste donné, dans l'hypothèse d'un coronographe idéal. Cette inversion :[55]

$$a = \frac{\sqrt{c}}{2\pi} \times \lambda,$$

accompagnée d'une petite application numérique devrait achever de convaincre le lecteur de la difficulté de l'objectif de la détection directe. Ainsi, pour atteindre dans l'image coronographique brute, un contraste brut local de 10^{-6}, requiert un contrôle de front d'onde de la fréquence spatiale correspondante avec une stabilité meilleure que $10^{-3}/2\pi \times \lambda$, ce qui en bande H ($\lambda = 1.6\,\mu$m) correspond a mieux que 0.25 nm !

55. Contrairement à l'utilisation précédente, comme on parle du front d'onde et pas de la surface d'un miroir, le facteur 4π est remplacé ici par 2π

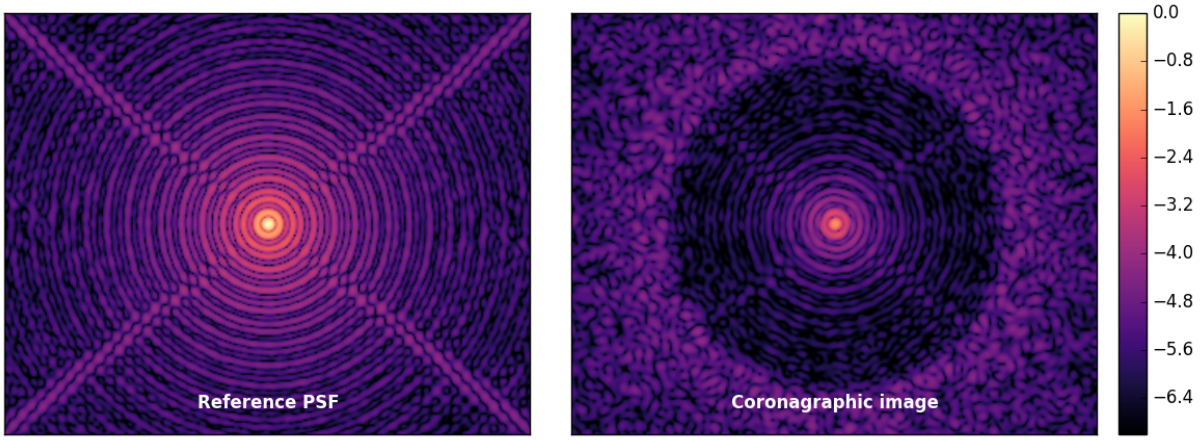


FIGURE 46: Exemple d'image simulée (monochromatique) utilisant le modèle théorique du coronographe parfait. Panneau de gauche : PSF non-coronographique, en apparence idéale, et pourtant affectée par un résidu de surface d'onde de 50 nm RMS. Panneau de droite : image coronagraphique correspondante affectée par la même erreur de surface d'onde. Les deux panneaux partagent la même échelle de couleur.

Dans l'hypothèse d'un coronographe idéal (voir la figure 46) supprimant parfaitement la contribution statique de la diffraction sur l'axe, la fuite coronographique induite par une perturbation de la surface d'onde, caractérisée par une valeur RMS globale σ, se manifeste par la présence d'une réplique de la PSF originale, atténuée par un facteur :

$$c_b = \frac{1}{\sqrt{2}} \left(\frac{2\pi\sigma}{\lambda} \right)^2.$$

Dans le cas des meilleures performances rapportées des systèmes XAO en opération à l'heure d'aujourd'hui ($\sigma \sim$50 nm RMS)[56], ce facteur d'atténuation de la PSF est $c_b \approx 0.027$.

CE CHIFFRE MÉRITE À ÊTRE COMPARÉ aux incertitudes des mesures de clôtures de phase qui ont été présentées dans le chapitre précédent : bénéficiant seulement d'une correction adaptative classique, il

56. J.-F. Sauvage, T. Fusco, C. Petit, A. Costille, D. Mouillet, J.-L. Beuzit, K. Dohlen, M. Kasper, M. Suarez, C. Soenke, A. Baruffolo, B. Salasnich, S. Rochat, E. Fedrigo, P. Baudoz, E. Hugot, A. Sevin, D. Perret, F. Wildi, M. Downing, P. Feautrier, P. Puget, A. Vigan, J. O'Neal, J. Girard, D. Mawet, H. M. Schmid, and R. Roelfsema. SAXO: the extreme adaptive optics system of SPHERE (I) system overview and global laboratory performance. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(2):025003, Apr. 2016

n'est pas rare de mesurer des incertitudes inférieures au demi degré, ce qui correspond à une incertitude équivalente de quelques nanomètres sur le front d'onde. Même si l'interférométrie à masque non-redondant ne bénéficie pas du gain apporté par la réduction du bruit de photon de la source coronographiée, la propriété auto-étalonnante de la clôture de phase reste très avantageuse. La dernière partie de ce mémoire va par conséquent revenir sur cette notion et expliquer comment cette idée peut être généralisée à l'analyse d'images corrigées par de l'optique adaptative sans utiliser de masque non-redondant.

“SO, IN THE FACE OF OVERWHELMING ODDS, I’M LEFT WITH ONLY ONE OPTION : I’M GOING TO HAVE TO SCIENCE THE SHIT OUT OF THIS.”

ANDY WEIR, *THE MARTIAN*

Les noyaux de phase



FIGURE 47: Appliquer les méthodes de l'interférométrie aux données acquises par des télescopes équipés d'optique adaptative pour atteindre le plein potentiel de résolution angulaire des télescopes : voilà l'ambition principale du projet KERNEL.

Ce mémoire a rappelé à plusieurs reprises que le pouvoir de résolution d'un télescope est fondamentalement contraint par le phénomène de diffraction, dont l'ordre de grandeur caractéristique est donné par le rapport λ/D (voir équation 2). Les livres d'introduction à l'astronomie, présentent la contrainte de résolution angulaire de façon très simpliste, basée sur l'utilisation d'un critère binaire, considérant que tout ce qui est de taille inférieure à λ/D est non résolu. Les résultats observationnels présentés dans la première partie de ce mémoire ont, je l'espère, réussi à montrer que ce critère binaire ne réflète pas le réel potentiel des techniques modernes d'interprétation des données à haute résolution angulaire comme dans le cas de l'interférométrie à masque non-redondant.

DANS UNE IMAGE DOMINÉE PAR LES STRUCTURES DE DIFFRACTION induite par la forme de la pupille du télescope lui même, la capacité à résoudre ou ne pas résoudre une structure dans le voisinage direct d'un objet est plutôt une question de rapport signal sur bruit, sur des mesures de visibilité complexe auxquelles la géométrie d'une pupille donne accès. A signal sur bruit infini ou si la réponse impulsionnelle (notée PSF, pour point spread function) d'un instrument est parfaitement connue, il n'y a pas de limite effective à la résolution qu'on peut atteindre en déconvoluant une image [57]. C'est cette argumentation qui avait conduit, dans la première partie du manuscrit de cette HDR, à masquer la majorité de la pupille d'un instrument. On a pu voir que l'insertion dans la pupille d'un masque non-redondant simplifie considérablement l'interprétation des données même lorsque la PSF n'est pas parfaitement caractérisée et permet l'extraction de quantités observables auto-étalonnées appelées clôtures de phase qui permettent de s'affranchir des erreurs résiduelles de la correction adaptative.

57. à condition de suffisamment échantillonner cette image

CETTE SIMPLIFICATION A UN COÛT : la condition stricte de non-redondance de la pupille impose que typiquement 90 % de la surface collectrice totale soit masquée. À cette perte de transmission s'ajoute un effet plus marqué de diffraction par des trous, qui s'accompagne par une perte significative de sensibilité des observations. Tant que la perte de signal s'accompagne d'un gain global de rapport signal sur bruit, l'impact global reste positif. Mais à mesure que les ambitions scientifiques des programmes d'observation par interféromé-

trie augmente, le bruit de photon des structures à détecter dans le voisinage d'une source centrale non-résolue devient un terme dominant de l'équation de rapport signal sur bruit. Cette perspective est suffisamment sérieuse pour que les promoteurs principaux de l'utilisation de masques non-redondants se soient lancés dans l'étude et l'expérimentation de remapping de la pupille comme FIRST [58] à base de fibres optiques ou DragonFly [59] à base de circuit en optique intégrée, qui peuvent en théorie permettre d'utiliser l'intégralité de la lumière collectée en réorganisant la pupille redondante d'un instrument classique en une ou plusieurs pupilles non-redondantes.

58. G. Perrin, S. Lacour, J. Woillez, and É. Thiébaut. High dynamic range imaging by pupil single-mode filtering and remapping. **MNRAS**, 373:747–751, Dec. 2006

59. N. Jovanovic, P. G. Tuthill, B. Norris, S. Gross, P. Stewart, N. Charles, S. Lacour, M. Ams, J. S. Lawrence, A. Lehmann, C. Niel, J. G. Robertson, G. D. Marshall, M. Ireland, A. Fuerbach, and M. J. Withford. Starlight demonstration of the Dragonfly instrument: an integrated photonic pupil-remapping interferometer for high-contrast imaging. **MNRAS**, 427:806–815, Nov. 2012

La puissance pratique de la clôture de phase m'a, au fur et à mesure de son utilisation, fortement impressionné. L'exploitation de cet observable m'a permis lors de la première partie de mon parcours post-thèse, de faire beaucoup d'observations intéressantes et inédites dans le domaine de la détection de compagnons, couvrant une région de l'espace des paramètres qui n'est toujours pas abordée par la coronographie depuis le sol. J'ai été tenté d'en étendre le domaine d'application, et ai commencé à réfléchir à une extension de cette notion, qui soit utilisable pour des pupilles ne respectant pas la condition de non-redondance. Un des projets de ma thèse concernait la méthode de cophasage dite des speckles dispersés [60], que j'avais partiellement réussi à appliquer aux pupilles redondantes, mais avec une méthode peu généralisable. Fort de cette expérience de l'utilisation pratique des clôtures de phase, j'ai révisé la "recette" de ma thèse en lui donnant une fondation saine et solide, reposant sur l'algèbre linéaire. Grâce à cette approche, j'ai pû démontrer l'existence de nouvelles quantités observables, baptisées noyaux de phase (ou kernel-phase en anglais), car elles forment une base pour le noyau de l'application linéaire qui permet de modéliser le système optique. Quelques années plus tard, j'ai complété cette étude initiale par une description d'une application littéralement orthogonale à celle des noyaux, qui exploite l'information contenue dans le plan de Fourier à des fins de métrologie.

60. F. Martinache. Global wavefront sensing for interferometers and mosaic telescopes: the dispersed speckles principle. *Journal of Optics A: Pure and Applied Optics*, 6:216–220, Feb. 2004

Ces développements ont été rendus possibles grâce à un changement majeur : la disponibilité de l'optique adaptative (OA). En effet, contrairement à ce qui se faisait encore vers la fin des années 1980 [61], le masque non-redondant n'est maintenant que très rarement utilisé en l'absence d'OA. Combiner l'interférométrie NRM à l'OA était une étape facilement justifiable : l'analyse de Fourier des interférogrammes NRM requiert des temps d'intégration plus courts que le temps de cohérence associé à la turbulence atmosphérique. Cette contrainte se traduit rapidement par une faible sensibilité, limitant la technique à l'observation des sources les plus brillantes. L'OA, même si elle reste imparfaite en l'absence de feedback direct en provenance du plan focal (voir chapitre précédent), stabilise le front d'onde et permet d'intégrer plus longtemps, rendant possible l'observation interférométrique de sources considérablement plus faibles et permet-

61. A. C. S. Readhead, T. S. Nakajima, T. J. Pearson, G. Neugebauer, J. B. Oke, and W. L. W. Sargent. Diffraction-limited imaging with ground-based optical telescopes. **AJ**, 95:1278–1296, Apr. 1988; and T. Nakajima, S. R. Kulkarni, P. W. Gorham, A. M. Ghez, G. Neugebauer, J. B. Oke, T. A. Prince, and A. C. S. Readhead. Diffraction-limited imaging. II - Optical aperture-synthesis imaging of two binary stars. **AJ**, 97:1510–1521, May 1989

tant de mesurer des clôtures de plus en plus petite amplitude [62].

Nous avons vu dans la première partie du mémoire comment l'information de cohérence potentiellement collectable par un télescope est détruite (voir la figure 15) lors d'un process impliquant la turbulence et la redondance. L'OA change cependant considérablement la donne : la figure 48 illustre son impact sur la mesure de la visibilité complexe résultant de la somme de cinq phaseurs (ie. correspondant à une base interférométrique redondante d'ordre cinq. En l'absence d'OA, les phases individuelles des phaseurs sont effectivement aléatoires mais dès que l'OA offre une correction même partielle de la surface d'onde, les phaseurs tendent à s'aligner et la phase résultante peut être approximée comme la moyenne de la phase des phaseurs individuels. Ce changement va permettre de reformuler la relation de convolution objet-image (equation 1) qui est fondamentalement dégénérée, en une équation linéaire, en s'intéressant à la phase dans l'espace de Fourier.

62. P. Tuthill, J. Lloyd, M. Ireland, F. Martinache, J. Monnier, H. Woodruff, T. ten Brummelaar, N. Turner, and C. Townes. Sparse-aperture adaptive optics. In *Advances in Adaptive Optics II. Edited by Ellerbroek, Brent L.; Bonaccini Calia, Domenico. Proceedings of the SPIE, Volume 6272, pp. (2006).*, July 2006

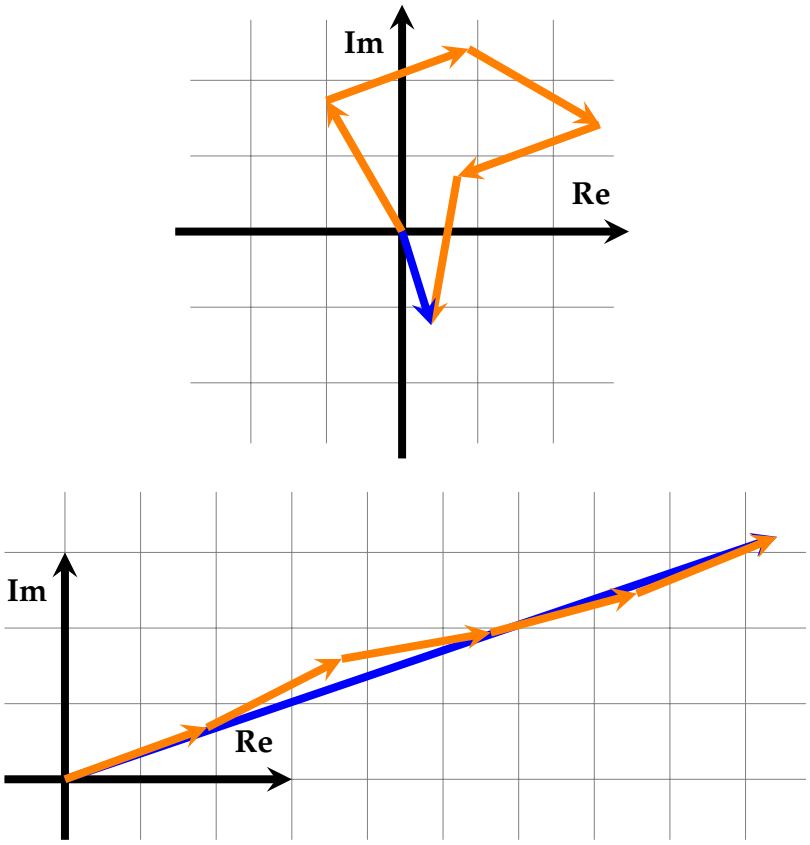


FIGURE 48: Somme de cinq phaseurs dans le plan complexe : le premier panneau reproduit une figure présentée au premier chapitre. La phase aléatoire de chaque phaseur conduit à une visibilité complexe globale dont la phase et la visibilité peuvent prendre des valeurs aléatoires. Le second panneau présente une situation beaucoup plus avantageuse où une optique adaptative stabilise partiellement la phase individuelle des phaseurs. La phase résultante peut être approximée comme étant la moyenne des phases individuelles des phaseurs contribuant à cette base.

Généralisation de la notion de clôture

Pour comprendre comment construire cette généralisation de la clôture de phase, on peut passer en revue une petite série de modèles élémentaires de complexité croissante. Le premier système est la configuration interférométrique permettant l'acquisition d'une unique quantité de clôture : trois ouvertures A,B et C, formant un triangle tel

que les lignes de base (A-B), (B-C) et (C-A) soient non redondantes, comme représenté dans la figure 49.

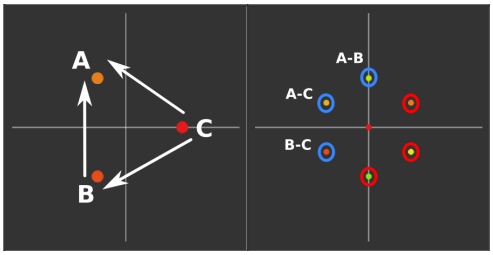


FIGURE 49: A gauche : représentation schématique d'un interféromètre à trois sous-ouvertures A,B et C formant un triangle non-redondant. A droite : représentation du plan de Fourier couvert par ce système : trois mesures de phase pour les lignes de base (A-B), (A-C) et (B-C) sont possibles.

Les équations décrivant la phase qui peut être mesurée dans l'espace de Fourier pour ces trois bases interférométriques s'écrivent de la façon suivante :

$$\begin{aligned}\Phi(A-B) &= \Phi_O(A-B) + (\varphi_A - \varphi_B)\\ \Phi(A-C) &= \Phi_O(A-C) + (\varphi_A - \varphi_C)\\ \Phi(B-C) &= \Phi_O(B-C) + (\varphi_B - \varphi_C).\end{aligned}$$

où $\Phi(X-Y)$ représente la phase mesurée pour la base (X-Y), $\Phi_O(X-Y)$ la phase attendue de la source pointée et la différence $(\varphi_X - \varphi_Y)$, le piston instrumental affectant cette même base. Ce système d'équations peut être réécrit de façon plus compacte sous forme matricielle, si l'on introduit les vecteurs à trois composantes Φ, Φ_O et φ, ainsi qu'une matrice $\mathbf{A}$ tels que :

$$\Phi = \Phi_O + \mathbf{A}\cdot\varphi \tag{24}$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \tag{25}$$

LA MATRICE $\mathbf{A}$ EST APPELÉE MATRICE DE TRANSFERT DE PHASE. Elle décrit comment la phase instrumentale φ se propage dans le plan de Fourier pour polluer la mesure des phases intrinsèques à l'objet que l'on souhaite caractériser. La relation de clôture classique, qui combine les lignes du système d'équations de façon à créer une relation ne contenant aucun terme dépendant de φ, peut être repensée comme l'application d'un opérateur linéaire $\mathbf{K}$, s'appliquant par la gauche, de telle sorte que $\mathbf{K}\cdot\mathbf{A} = 0$. La solution évidente ici : $\mathbf{K} = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$, appliquée par la gauche à l'équation 24, vérifie :

$$\mathbf{K}\cdot\Phi = \mathbf{K}\cdot\Phi_O, \tag{26}$$

ce qui est une reformulation algébrique de la propriété désirée de la clôture de phase : le produit $\mathbf{K}\cdot\Phi$ ne dépend plus que des propriétés de l'objet (Φ_O). Une fois ce modèle et cette réécriture compris, on peut passer au scénario suivant, utilisant lui aussi trois ouvertures A, B et C mais cette fois ci placées en ligne, de telle sorte que les bases (A-B) et (B-C) soient de la même longueur, comme représenté

dans la figure 50. Un tel système est redondant : les bases (A-B) et (B-C) donnent accès à la même information $\Phi_O(B-C)$ sur la source. Chaque mesure est par contre polluée par son terme de phase instrumentale respectif ($\Delta\varphi_{A-B}$ et $\Delta\varphi_{B-C}$). Dans un mode de recombinaison interférométrique de type Fizeau (qui est le seul considéré ici), ces informations vont se trouver mélangées au même point de l'espace de Fourier, ce qui se traduit par un système d'équations différent du cas précédent[63] :

63. On introduit la notation $\Delta\varphi_{XY} = \varphi_X - \varphi_Y$ qui permet de réduire la longueur des équations.

$$\Phi(A-C) = \Phi_O(A-C) + \Delta\varphi_{AC} \quad (27)$$

$$\Phi(B-C) = Arg\left(e^{i(\Phi_0(B-C)+\Delta\varphi_{AB})} + e^{i(\Phi_0(B-C)+\Delta\varphi_{BC})}\right). \quad (28)$$

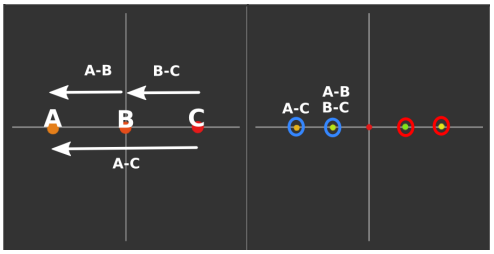


FIGURE 50: A gauche : représentation schématique d'un interfèromètre à trois sous-ouvertures A, B et C formant un système redondant en ligne. A droite : représentation schématique du plan de Fourier couvert par ce système. Cette configuration redondante ne donne accès qu'à deux éléments d'information de phase sur l'objet. L'équation de la phase pour la base (B-C) doit être modifiée pour prendre en compte cette redondance.

Dans sa forme de base, l'équation de phase 28 est problématique. Les deux termes exponentiels complexes, peuvent, tout à fait, selon les valeurs des pistons instrumentaux s'ajouter de façon constructive ou desctructive, ce qui se traduit par une mesure de la phase pouvant prendre des valeurs complètemet aléatoires (voir la figure 48). C'est précisément ce qui a justifié l'utilisation d'un masque imposant la condition de stricte non-redondance. Cependant, en présence d'optique adaptative, cette équation peut être linéarisée. Les termes instrumentaux supposés être suffisamment petits ($\Delta\varphi << 1$) pour permettre d'écrire le nouveau système[64] :

64. Avec la phase instrumentale linéarisée, des simplifications s'opèrent, telle qu'ici où les deux termes $\Delta\varphi_{AB}$ et $\Delta\varphi_{BC}$ s'additionnent pour ne laisser que $\Delta\varphi_{AC}$ (théorème de Chasles).

$$\Phi(A-C) = \Phi_O(A-C) + \Delta\varphi_{AC} \quad (29)$$

$$\Phi(B-C) \approx \Phi_O(B-C) + \frac{1}{2}\Delta\varphi_{AC}. \quad (30)$$

Comme dans le scénario précédent, ce système d'équation peut être réécrit sous forme matricielle, requérant l'introduction d'une nouvelle matrice (diagonale) $\mathbf{R}$ appelée matrice de redondance, telle que :

$$\Phi = \Phi_O + \mathbf{R}^{-1}\cdot\mathbf{A}\cdot\varphi \quad (31)$$

avec cette fois :

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \quad (32)$$

On peut définir une nouvelle relation de clôture, matérialisée par un nouvel opérateur $\mathbf{K}$ s'appliquant par la gauche tel que $\mathbf{K}\cdot\mathbf{R}^{-1}\cdot$

$\mathbf{A} = 0$. La solution : $\mathbf{K} = \begin{bmatrix} 1 & -2 \end{bmatrix}$ vérifie cette équation. C'est une forme généralisée de clôture de phase (généralisée car elle fait intervenir la redondance), que j'ai choisi d'appeler noyau de phase, ou kernel, car le vecteur contenu dans cet opérateur fait partie du noyau de la matrice de transfert de phase $\mathbf{A}$.

Pour cet exemple spécifique, le passage par une écriture matricielle est une complication discutable et le véritable potentiel de ce formalisme devient plus évident lorsqu'on l'utilise pour traiter le cas d'ouvertures plus riches, comme va le montrer le dernier des scénarios élémentaires, représenté dans la figure 51, n'impliquant qu'une ouverture de plus que les cas précédents.

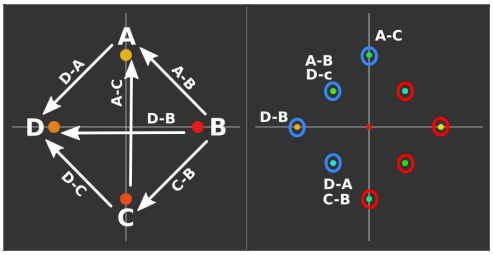


FIGURE 51: A gauche : représentation schématique d'un interfèromètre à quatre sous-ouvertures A, B, C et D ; formant un système redondant en forme de losange. A droite : le plan de Fourier couvert par ce système. Cette configuration redondante donne accès à quatre lignes de bases distinctes, dont deux redondantes d'ordre 2.

Avec les matrices de transfert de phase $\mathbf{A}$ et de redondance $\mathbf{R}$ en place, la même équation 31 s'applique. Le contenu des matrices doit être mis à jour. Il faut choisir une convention pour l'ordre dans lequel on écrit les équations correspondant aux différentes bases dans le plan de Fourier : ma convention est de parcourir le plan de Fourier de la gauche vers la droite d'abord, et de haut en bas ensuite, ce qui donne la séquence suivante : (D-B), (A-B), (D-A) et (A-C). Il faut de la même façon, choisir un ordre pour les sous-ouvertures : pour aider à la lecture, je choisis ici de prendre les ouvertures dans l'ordre alphabétique. Avec ces conventions, les matrices s'écrivent :

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{33}$$

Cette fois ci, deux noyaux ou relations de clôture généralisée sont possibles : on peut ainsi vérifier que $\mathbf{K} \cdot \mathbf{R}^{-1} \cdot \mathbf{A} = 0$ lorsque :

$$\mathbf{K} = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 1 & 0 & -2 & -1 \end{bmatrix}. \tag{34}$$

Recherche automatisée des kernels

TOUT L'ART DE L'APPROCHE PROPOSÉE REPOSE dans la bonne tenue d'un modèle qui permet de savoir quelles combinaisons d'ouvertures contribuent aux termes de pistons $\Delta\varphi_i^k$ pour la composante de Fourier d'indice k : cette connaissance est contenue dans la matrice de transfert de phase $\mathbf{A}$. Au delà des exemples élémentaires

proposés, la construction à la main de ces modèles devient vite fastidieuse : la suite des opérations est confiée à un algorithme automatique qui va construire les matrices **A** et **R** à partir d'un fichier de coordonnées proposant une représentation discrète de la pupille d'un instrument [65]. Si l'on appelle N_A le nombre de sous-ouvertures virtuelles dans la pupille et N_{UV}, le nombre de fréquences spatiales distinctes que ces sous-ouvertures permettent de construire, la matrice de transfert **A** est une matrice à N_{UV} lignes et $N_A - 1$ colonnes [66].

65. Le code XARA (eXtreme Angular Resolution) réalisant ce travail est disponible sur `http://github.com/fmartinache/xara`

66. Même si elle est bien affectée par les pistons individuels affectant les différentes bases virtuelles, une image Fizeau est insensible à un terme de piston global s'appliquant à l'intégralité de la pupille : on ne mesure pas le front d'onde absolu, mais relativement à un point de référence choisi arbitrairement dans la pupille. Cela revient à éliminer une des colonnes de la matrice de tranfert de phase, d'où le $N_A - 1$.

Un exemple concret de modèle utilisé pour interpréter des données en provenance du télescope de Hale au Mont Palomar est présenté dans la figure 52. Ce modèle, construit par le code de calcul des kernels XARA, montre la représentation discrète de la pupille et la couverture du plan de Fourier qui lui est associée. Pour cet exemple spécifique, la pupille continue du télescope est représentée par un vecteur constitué de 332 points d'échantillonnage, choisis selon une maille régulière de façon à réfléter la nature redondante de cette pupille et incluant la connaissance de la présence d'un masque placé par l'observateur en plan pupille ("Medium-Cross"). Cet échantillonnage donne accès à un total de 1124 fréquences spatiales distinctes dans le plan de Fourier. Cela veut dire que la matrice de transfert de phase **A** correspondant à ce modèle de l'instrument est faite de 1124 lignes et 331 colonnes. La matrice de redondance **R** qui y est associée, contient sur sa diagonale des termes allant de 1 (quelques bases sont en effet non-redondantes) à 280 (pour les bases les plus courtes, également les plus redondantes). Notez au passage que les deux matrices $\mathbf{R}^{-1}$ et **A** peuvent tout à fait en pratique être combinées en une seule matrice [67]. L'avantage à garder les deux termes séparés est que sous cette forme, **A** ne contient que les valeurs 0, 1 ou -1.

67. A l'avenir, je ne ferai d'ailleurs plus la distinction entre **A** et $\mathbf{R}^{-1} \cdot \mathbf{A}$

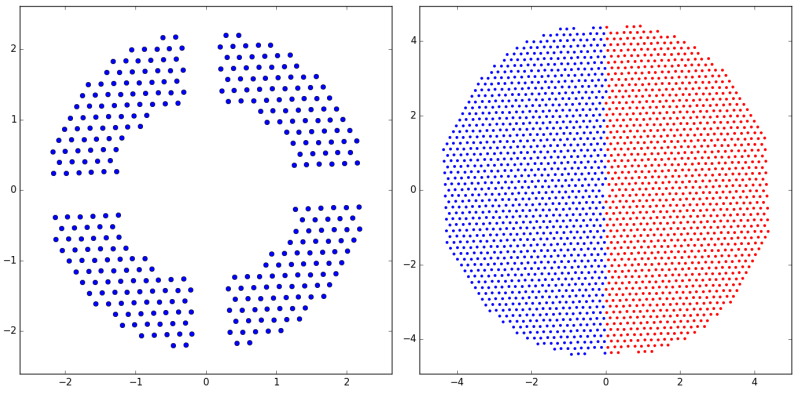


Figure 52: A gauche : modèle discret de la pupille dite "Medium-Cross" de l'instrument PHARO équipant le télescope de Hale au Mont Palomar, constitué de 332 points d'échantillonnage. A droite : couverture du plan de Fourier offerte par cette pupille : les 332 éléménts discrets de la pupille forment 1124 fréquences spatiales distinctes dans le plan de Fourier. Les unités des axes horizontaux et verticaux sont en mètres.

Une fois cette matrice construite, la recherche de ses kernels se fait également de façon automatique, en utilisant sa décomposition en valeurs singulières (SVD pour Singular Value Decomposition). Cet algorithme permet de décomposer la matrice **A** de telle

sorte que :

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T, \tag{35}$$

où $\mathbf{U}$ et $\mathbf{V}$ sont des matrices unitaires et $\mathbf{\Sigma}$ une matrice diagonale contenant les valeurs singulières de $\mathbf{A}$. Ces valeurs singulières sont réelles, posisives et ordonnées de manière décroissante jusqu'à atteindre exactement zéro à un point qui dépend des propriétés de la pupille. La figure 53 représente les valeurs singulières associées au modèle de PHARO introduit plus haut.

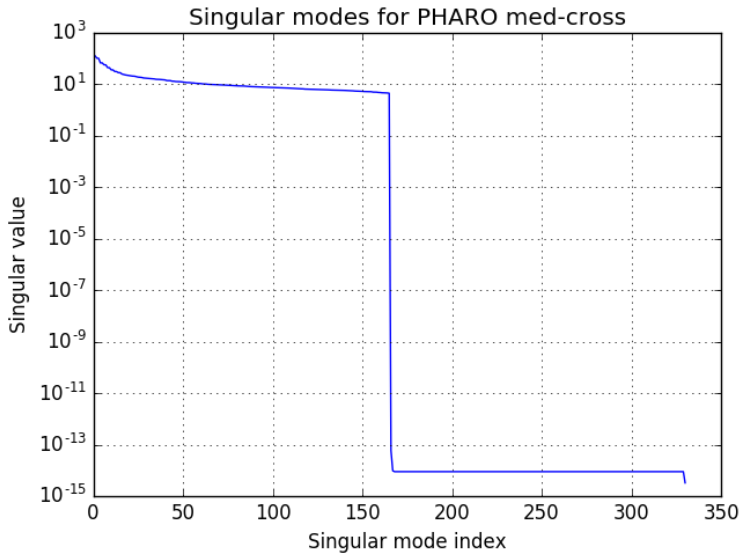


FIGURE 53: Représentation des valeurs singulières de la matrice de transfert du modèle de l'instrument PHARO introduit précédemment. On observe pour ce cas précis que les 166 premières valeurs singulières sont strictement positives et que les 166 suivantes sont au niveau du bruit numérique, ie. zéro.

Dans notre contexte, les kernels recherchés sont les combinaisons de lignes de la matrice de tranfert $\mathbf{A}$, regroupées dans un opérateur $\mathbf{K}$, qui vérifient $\mathbf{K} \cdot \mathbf{A} = \mathbf{0}$. Ces combinaisons se trouvent être contenues dans les colonnes de la matrice $\mathbf{U}$ correspondant à des valeurs singulières nulles. On peut s'en convaincre en multipliant les deux côtés de l'équation 35 par la gauche par $\mathbf{U}^T$:

$$\mathbf{U}^T \cdot \mathbf{A} = \mathbf{\Sigma} \cdot \mathbf{V}^T, \tag{36}$$

et bien constater que les dernières colonnes de $\mathbf{U}$ produisent des zéros. La phase instrumentale peut donc théoriquement être entièrement éliminée et produire, à partir des N_{UV} observables originaux, pollués par les résidus de contrôle de surface d'onde, un sous-ensemble de N_K nouveaux observables qui n'en dépendent pas : les noyaux de phase.

La relation de convolution revisitée

Dans tous les cas particuliers qui viennent d'être présentés, on voit qu'on peut, en faisant l'hypothèse des petites aberrations, écrire que l'équation générale de la phase d'une composante d'indice k de Fourier, de redondance r :

$$\Phi^k = \Phi^k_O + Arg\left(\sum_{i=0}^{r} \exp\left(j\Delta\varphi^k_i\right)\right), \tag{37}$$

peut être linéarisée :

$$\Phi^k \approx \Phi^k_O + \frac{1}{r}\sum_{i=0}^{r} \Delta\varphi^k_i, \tag{38}$$

Il est possible de connaître *a priori* les différentes parties de l'ouverture qui contribuent à chaque point de mesure dans le plan de Fourier et d'encoder cette connaissance dans la matrice de transfert **A** et la matrice de redondance **R**. Un équivalent de la relation classique de convolution objet - image :

$$I = O \otimes PSF \tag{39}$$

peut être écrit pour la phase dans l'espace de Fourier, s'écrivant comme [68] :

$$\Phi = \Phi_0 + \mathbf{R}^{-1} \cdot \mathbf{A} \cdot \varphi. \tag{40}$$

La relation de convolution est dégénérée : une image donne accès à autant de quantités observables qu'elle n'offre de pixels. Déconvoluer l'image, en l'absence d'apriori fort sur les propriétés de la PSF, revient à trouver deux fois plus d'inconnues que de contraintes disponibles, ce qui n'est pas possible. Le modèle proposé permet un traitement alternatif des données : ce sont les propriétés de **A** (et donc celles du modèle discret représentant la pupille) qui vont déterminer la part d'information sur la source qui est recouvrable par les kernels [69] ou les contraintes métrologiques qu'on peut en tirer [70].

CONTRAIREMENT À LA CONVOLUTION le modèle proposé n'offre qu'une description partielle de l'information, sur la phase. L'information d'amplitude, même si elle est également accessible par un traitement similaire [71], ne sera pas traitée ici.

68. Le code de couleur utilisé permet de faire la correspondance entre les deux formes.

69. F. Martinache. Kernel Phase in Fizeau Interferometry. **ApJ**, 724:464–469, Nov. 2010

70. F. Martinache. The Asymmetric Pupil Fourier Wavefront Sensor. **PASP**, 125:422–430, Apr. 2013

71. B. J. S. Pope. Kernel phase and kernel amplitude in Fizeau imaging. **MNRAS**, 463:3573–3581, Dec. 2016

Pupille et nombre de kernels

La construction du modèle discret d'un instrument est l'étape clé du process, qui conditionne toute la suite des opérations. Il n'aura peut être pas échappé au lecteur, dans l'exemple illustré par la figure 53, que le modèle utilisé pour décrire la pupille de l'instrument PHARO produit un nombre de valeurs singulières nulles qui est exactement la moitié du nombre de points originaux. Cette propriété remarquable est due aux propriétés de symétrie de la géométrie de la pupille. En pratique, une pupille symétrique discrétisée en un réseau de N_A sous-ouvertures produira systématiquement $N_A/2$ valeurs singulières non-nulles. Si on connaît le nombre N_{UV} de points dans le plan de Fourier, on peut donc connaître, pour une pupille symétrique, le nombre N_K de kernels :

$$N_K = N_{UV} - N_A/2. \tag{41}$$

Une justification de cette propriété a été donnée dans la section annexe d'un article qui sera introduit un peu plus loin[72]. En plus des propriétés de symétrie de la pupille (sur lesquelles nous reviendrons plus tard), les détails de la géométrie vont avoir un impact sur le nombre total de points distincts accessibles dans le plan de Fourier. Il est impossible de prédire tous les cas de figure possibles, néanmoins quelques cas particuliers peuvent guider notre compréhension de la dimension du problème en nous donnant les bons ordres de grandeur.

72. F. Martinache. The Asymmetric Pupil Fourier Wavefront Sensor. **PASP**, 125:422–430, Apr. 2013

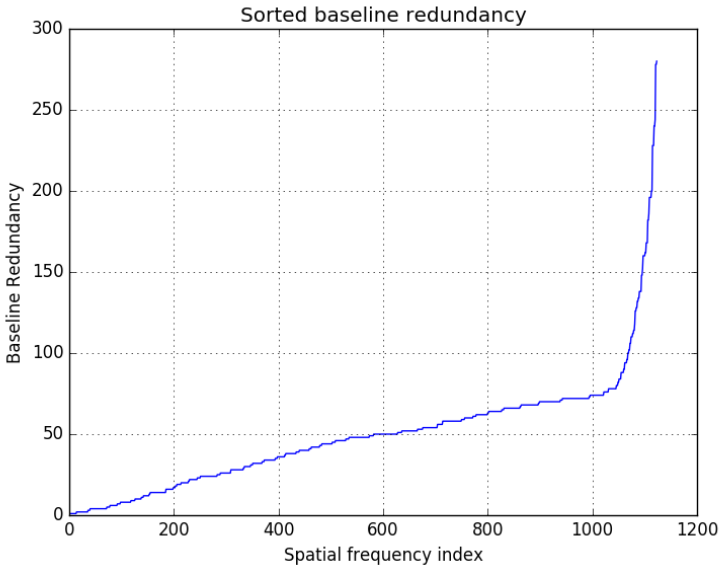


FIGURE 54: Répartition ordonnée de la redondance pour le modèle de pupille de PHARO à 332 sous ouvertures virtuelles. La redondance moyenne pour ce modèle est de 48. L'intégrale de cette courbe est de 54946.

Le premier ordre de grandeur à avoir en tête est celui du nombre de fréquences spatiales maximal accessible à une pupille strictement non-redondante de N_A sous-pupilles, qui est égal au nombre de paires parmi N_A, soit $\binom{N_A}{2} = N_A \times (N_A - 1)/2$. Les 332 sous ouvertures virtuelles décrivant la pupille de l'instrument PHARO présenté

plus tôt, si on pouvait les réorganiser en un réseau non-redondant, donneraient accès à 54946 mesures distinctes ! Si ce nombre est considérablement plus grand que les 1124 points indépendants de l'espace de Fourier reportés, c'est parce que la pupille est fortement redondante, comme le rappelle la figure 54 affichant la répartition de cette redondance. L'intégrale de la courbe représentée, correspondant au total des 1124 points effectifs, pondérés par leur redondance respective, est égale à 54946, ce qui est exactement le nombre maximum reporté plus tôt.

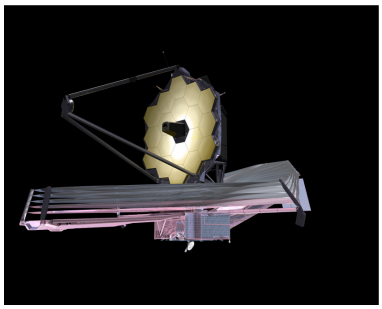


FIGURE 55: Exemple emblématique de télescope segmenté : le JWST, qui devrait être placé au point L2 du système Terre-Soleil en 2018.

UN CAS DE FIGURE RELATIVEMENT GÉNÉRIQUE sur lequel nous pouvons faire des observations intéressantes est celui d'un réseau de sous pupilles organisé en une série d'anneaux concentriques, suivant une maille hexagonale : c'est après tout une bonne approximation pour représenter une ouverture circulaire et au vu de la banalisation des ouvertures segmentées à venir (voir la Figure 55), ce modèle devient de plus en plus pertinent. Si on paramétrise cette géométrie par le nombre d'anneaux n_r qu'elle contient, le nombre total de sous-ouvertures contenu (excluant la sous-ouverture centrale) est de :

$$N_A = 3n_r \times (n_r + 1), \tag{42}$$

et le nombre de points indépendants formés dans l'espace de Fourier correspondant est de :

$$N_{UV} = 3n_r \times (2n_r + 1). \tag{43}$$

Comme de telles pupilles sont parfaitement symétriques, la propriété sur le nombre de kernels constructibles (équation 41) s'applique. On peut ainsi s'amuser à estimer la fraction totale de l'information de phase qui est récupérable par les kernels pour une pupille pleine, en calculant le rapport :

$$f(n_r) = \frac{N_K}{N_{UV}} = 1 - \frac{N_A}{2N_{UV}} = 1 - \frac{n_r + 1}{2 \times (2n_r + 1)}. \tag{44}$$

Créer un modèle discret au pas de plus en plus fin, conduit à un nombre de couronnes tendant vers l'infini : la fraction totale de l'information de phase sauvée par les kernels tend asymptotiquement vers :

$$\lim_{n_r \to \infty} f(n_r) = 0.75. \tag{45}$$

Ce qui veut dire que pour une pupille pleine, on peut compter récupérer environ 75 % de l'information de phase collectée par la pupille (mais polluée par les résidus de correction adaptative) grâce aux kernels. Ceci est une estimation pessimiste de la fraction d'information récupérable : une application numérique au cas spécifique de PHARO qui a accompagné ce chapitre révèle une meilleure performance, avec une fraction $f = (1124 - 332/2)/1124 \approx 0.85$. Ceci est dû aux propriétés spécifiques de la pupille, et en particulier de sa grande obstruction centrale.

Pupilles optimisées pour les kernels ?

L'observation précédente conduit naturellement à la question suivante : y a-t-il une géométrie de pupille qui optimise la fraction d'information de phase théoriquement récupérable par les kernels ? La capacité à comprendre la nature d'une source complexe avec un réseau interférométrique dépend de la richesse de la couverture uv et du nombre d'observables de qualité qui peuvent en être extraits. Si on exclut la possibilité de faire de la synthèse d'ouverture en utilisant la rotation du ciel pour augmenter cette couverture, la capacité à collecter plus d'observables pour un instrument dont la taille est contrainte est important.

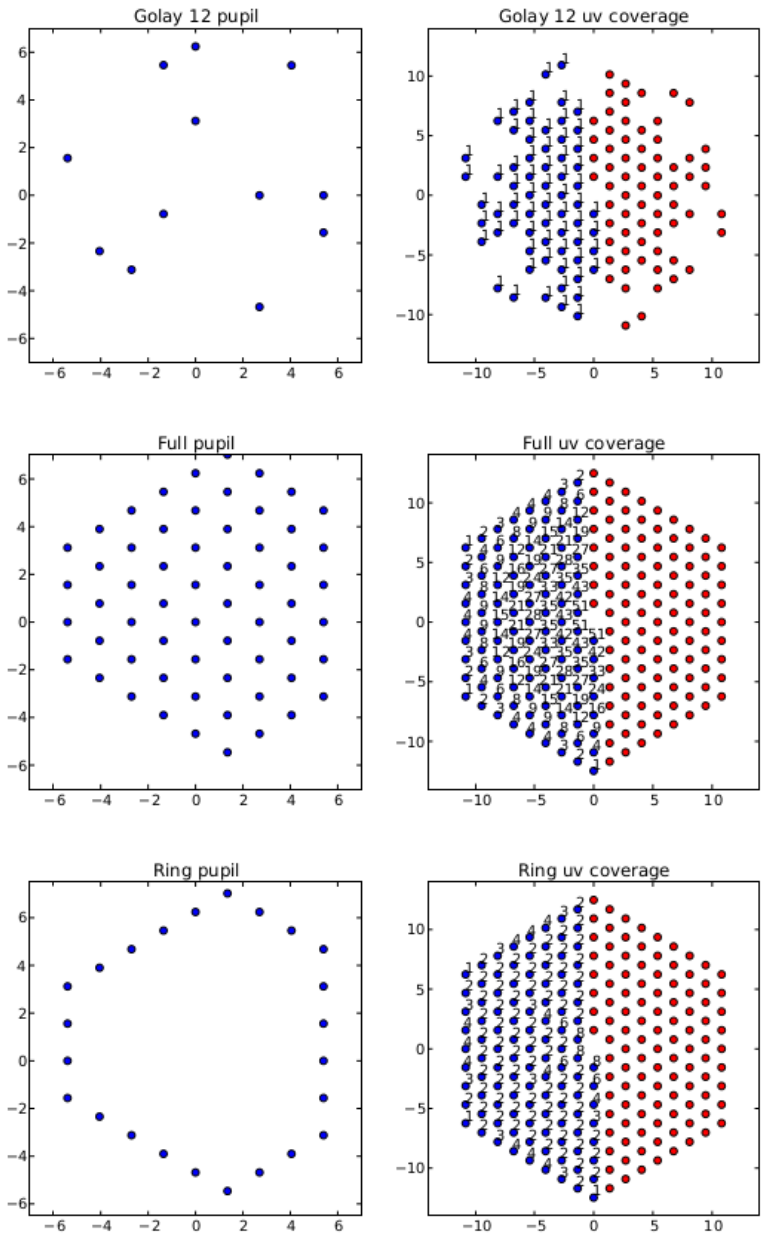


FIGURE 56: Comparaison entre la configuration compacte G12 et deux alternatives redondantes ayant la même empreinte globale.

On peut repartir d'une configuration non-redondante compacte, telle que la configuration G12 de Golay[73], et chercher la plus petite pupille discrète pleine dans laquelle la configuration G12 s'insère : une proposition, offrant la même baseline minimale et maximale est visible dans la rangée du milieu de la figure 56. Avec cette configuration dense, on augmente la couverture uv, qui passe de $N_{UV} = 66$ pour G12 à $N_{UV} = 108$ pour la configuration pleine. Cependant, lorsqu'on compare le nombre de kernels ou clôtures qu'on peut construire, on voit qu'avec $N_K = 55$ clôtures de phase, la configuration G12 reste avantageuse par rapport à l'ouverture pleine pour laquelle on peut extraire $N_K = 49$ kernels.

73. M. Golay. Point arrays having compact, nonredundant autocorrelations. *JOSA*, 61:272–273, 1971

	Golay 12	**Full**	**Ring**
n_A	12	27	15
f_A	0.61 (N_{UV}=66)	1.0 (N_{UV}=108)	1.0 (N_{UV}=108)
f_R	0.83 (N_K=55)	0.45 (N_K=49)	0.79 (N_K=85)
$f_A \times f_R$	0.51	0.45	0.79

TABLE 1: Résumé des propriétés des configurations G12, pleine et annulaire. n_A, n_{UV} et n_K représentent respectivement le nombre total de sous ouvertures dans la configuration, le nombre de points distincts dans le plan uv qu'ils couvrent et le nombre de kernels qui peuvent être déduits de ces mesures brutes. La configuration annulaire est la plus avantageuse.

Une alternative beaucoup plus intéressante à la configuration pleine est de ne conserver que les ouvertures situées sur la couronne externe du réseau dense : cette nouvelle géométrie annulaire et la pleine offrent la même couverture du plan uv, avec des propriétés de redondance très différentes. Ces différences se reflètent dans le nombre plus important $N_K = 85$ de kernels qu'on peut extraire de la configuration annulaire. Les propriétés comparées des trois réseaux sont résumées dans la Table 1.

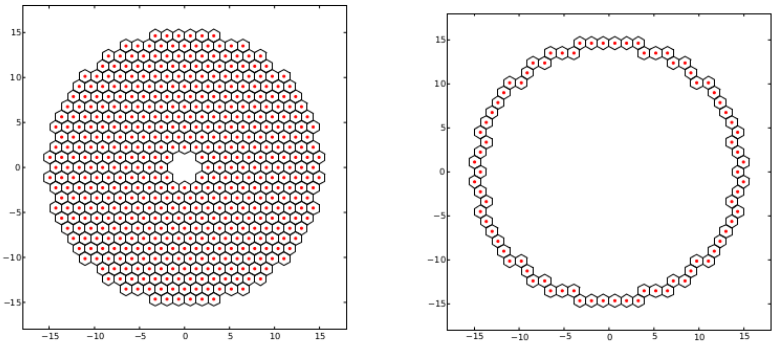


FIGURE 57: Le TMT (Thirty Meter Telescope) est un projet de télescope segmenté de 30 mètres de diamètre dont la pupille est faite de 492 segments. Si l'on considère que la base interférométrique la plus courte accessible correspond à la distance entre deux segments adjacents, cette pupille ouvre l'accès à 972 bases interférométriques distinctes. Les 78 segments de la couronne externe, permettent à eux seuls, de donner accès à toutes ces mêmes fréquences spatiales.

La configuration pleine représentée dans la figure 56 est construite sur une maille hexagonale mais n'est pas constituée d'une série d'anneaux concentriques. Si riche de ces nouvelles informations, on compare maintenant une telle configuration (voir la figure 57 qui présente le cas du TMT) à la configuration annulaire qui lui correspond, on confirme cette observation : la structure annulaire offre la même couverture du plan de Fourier en augmentant le nombre de kernels constructibles (rappel : $N_K = N_{UV} - N_A/2$) et donc la fraction totale

de l'information de phase préservable.

	Full pupil	**Ring only**
n_A	492	78
n_{UV}	972	972
n_K	726 (75 %)	933 (96 %)
Max. redundancy	462	26
Mean redundancy	124	3

TABLE 2: Propriétés comparées de la pupille du TMT et de la configuration n'utilisant que les ouvertures de la couronne externe. Avec six fois moins de segments, cette configuration annulaire offre la même couverture uv et permet l'extraction d'un nombre plus important de kernels de phase.

Si l'on reprend le calcul de la fraction total de l'information de phase récupérable par les kernels pour l'ouverture annulaire, réduite à un nombre réduit à $N_A = 6n_r$ ouvertures, cette fraction (voir l'équation 44) devient :

$$f(n_r) = 1 - \frac{1}{(2n_r + 1)}. \tag{46}$$

Cette fraction tend rapidement vers 100 % lorsque n_r augmente, suggérant fortement que les pupilles annulaires sont particulièrement adaptées à l'analyse proposée.

Cet avantage ne concerne que le nombre total d'observables sans aucune considération pour le rapport signal sur bruit qui leur est associé. Victime de l'économie, il a été évoqué à plusieurs reprises que le projet européen d'ELT pourrait au moins dans un premier temps, démarrer ses observations avec une pupille moins riche qu'initialement prévue. Un travail systématique des mérites relatifs de différentes configurations, incluant les considérations de signal sur bruit et de sensibilité aux aberrations résiduelles serait un guide utile pour décider de l'optimisation de la pupille de l'ELT, si une telle stratégie de mise en oeuvre du télescope se confirmait.

Extraction de la phase de Fourier

Héritant directement des techniques mises en oeuvre dans le pipeline de réduction des données à masque non-redondant pendant mon post-doc à Cornell, l'extraction de la phase de Fourier se fait après calcul de la FFT de l'image de départ, préalablement centrée et éventuellement zéro-paddée pour offrir une résolution suffisante dans l'espace de Fourier.

Le modèle discret du plan uv résultant du modèle de la pupille de l'instrument dont on peut voir un exemple à droite de la figure 52 donne une liste de bases interférométriques exprimées en mètres. Pour être utilisées sur un jeu de données, ces longueurs en mètres doivent être mises à l'échelle, par un facteur κ qui dépend de l'échelle focale *ps* (pour *plate-scale*, généralement exprimée en millisecondes d'arc par pixel ; *isz* de la taille totale de l'image (comprenant le *zero-padding*) en pixels ; ainsi que la longueur d'onde λ, exprimée en mètres :

$$\kappa = \frac{ps * \pi * isz}{180 * 3600 * 1000 * \lambda}. \tag{47}$$

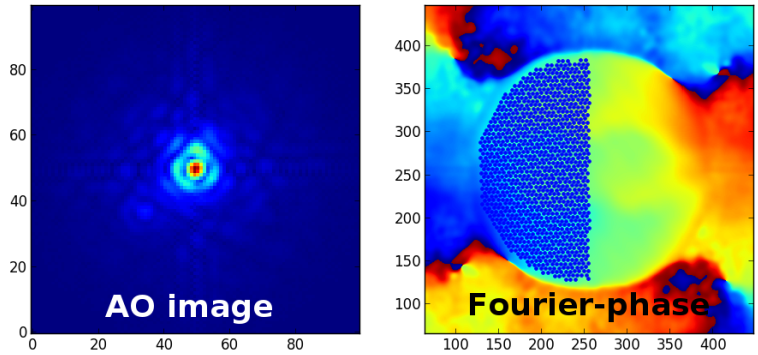


FIGURE 58: Exemple de projection du modèle de bases interférométiques sur la FFT d'une image. A gauche, image observée avec le télescope de Hale sur l'étoile α-Ophiucus. A droite, la phase de la transformée de Fourier de cette même image, calculée par FFT avec en overlay, la position des bases interférométriques virtuelles du modèle discret après mise à l'échelle.

Une fois mis à l'échelle en pixels, ce modèle des bases interférométriques est projeté sur le résultat de la FFT, et l'information est extraite aux pixels (arrondis au nombre entier le plus proche) pour peupler le vecteur Φ (voir figure 58). Bien que fonctionnelle et ayant été utilisée jusque là pour produire les résultats reportés dans ce mémoire, cette méthode a des limites : les bases projetées correspondent rarement à un nombre entier de pixels dans l'espace de Fourier et un modèle discret fin qui requiert une résolution importante dans l'espace de Fourier peut demander beaucoup de padding, pour éviter des problèmes d'aliasing.

L'approche du pixel entier le plus proche peut causer deux problèmes : dans le moins défavorable des cas, on est légèrement hors-cible et la mesure de la phase de Fourier peut être biaisée. Dans un cas moins favorable, lorsqu'on échantillonne l'information près du bord du support de Fourier, on peut accidentellement inclure des points en dehors de ce support, qui présentent des sauts de phase d'amplitude maximale $\pm\pi$. Une possibilité est, après FFT, d'interpoler les valeurs calculées pour fournir une estimation plus juste de la phase pour la base interférométrique considérée. Une telle approche est cependant lourde à mettre en oeuvre, en particulier si elle est utilisée dans un contexte de contrôle de front d'onde temps réel, qui peut avoir besoin d'être rapide.

Une approche plus satisfaisante consiste à calculer la transformée de Fourier discrète aux seuls points uv concernés par le modèle. L'approche développée est une ré-écriture de la transformée de Fourier discrète (DFT) classique $\mathbf{F}$ appliquée à une image $\mathbf{I}$ en utilisant une matrice de Vandermonde (normalisée) $\mathbf{W}$ contenant les puissances de la racine N-ième de l'unité $w = e^{-i2\pi/N}$:

$$\mathbf{W} = \left(\frac{w^{jk}}{\sqrt{N}} \right)_{j,k=0,\ldots,N-1}. \tag{48}$$

Disposant d'une telle matrice, la DFT de l'image s'écrit comme le résultat du produit matriciel suivant :

$$\mathbf{F} = \mathbf{W} \cdot \mathbf{I} \cdot \mathbf{W}^T, \tag{49}$$

où $\mathbf{W}$ et $\mathbf{W}^T$ ($= \mathbf{W}$) agissent respectivement sur les lignes et les colonnes de l'image $\mathbf{I}$[74]. Comme le modèle discret est construit à partir d'une grille périodique, le plan de Fourier est également échantillonné de façon périodique : en adaptant le contenu des matrices gauche et droite s'appliquant à $\mathbf{I}$, on peut écrire cette DFT locale (LDFT) comme le résultat d'un autre produit matriciel :

$$\mathbf{F} = \mathbf{L} \cdot \mathbf{I} \cdot \mathbf{R}, \tag{50}$$

74. En faisant l'hypothèse que l'image est carrée. Pour une image rectangulaire, les matrices gauches et droites seront différentes. C'est la disponibilité d'une version particulièrement rapide de cette DFT (la FFT) qui fait que cette technique est très employée.

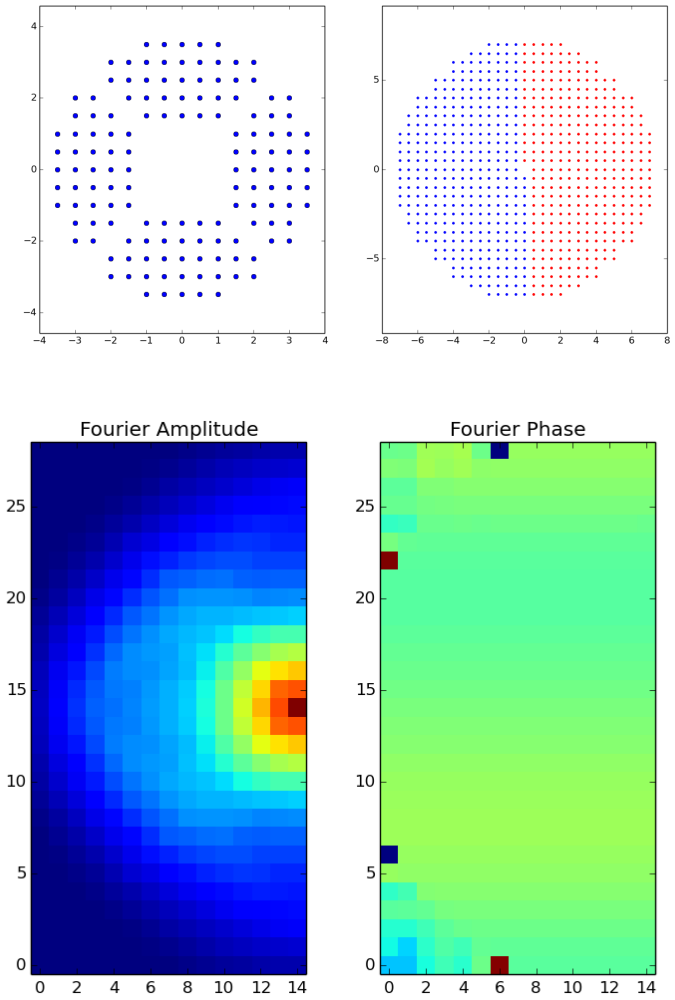


FIGURE 59: En haut : Modèle discret (pupille et plan uv) d'une ouverture circulaire avec obstruction centrale de 30 % du diamètre total et araignées se croisant perpendiculairement. En bas, calcul de la LDFT d'une image produite par le système optique, en l'absence d'aberration, séparant l'amplitude (à gauche) et la phase (à droite) calculée pour les fréquences spatiales demandées seulement. Avec une grille périodique orientée de façon à suivre la direction des pixels dans l'image, l'algorithme est particulièrement efficace car peu de points sont calculés en dehors du vrai support de la transformée de Fourier.

où les lignes de $\mathbf{L}$ (respectivement, les colonnes de $\mathbf{R}$)[75] contiennent les puissances successives (de $-isz/2$ à $+isz/2-1$) du vecteur de racines pour les coordonnées v_j (respectivement u_j) des N_b bases concernées :

$$v = \left(\frac{e^{-i2\pi\kappa u_j/isz}}{\sqrt{isz}} \right)_{j=0,N_b-1}. \tag{51}$$

Pour évaluer l'intérêt de cette implémentation de l'algorithme d'extraction, on va comparer les résidus d'extraction des deux méthodes

75. voir l'implémentation exacte sur http://github.com/fmartinache/xara/blob/master/xara/core.py#L491

FFT et LDFT, pour une image à signal sur bruit infini : une image contenant un objet binaire de propriétés connues (séparation angulaire de 30 mas et contraste de 0.25), acquise par un télescope de 8 mètres de diamètre, avec une pupille respectant le modèle présenté dans la description de la figure 59, observant à la longueur d'onde $\lambda = 1.6\,\mu$m avec une échelle focale de 10 mas par pixel. On s'attend à ce que la qualité de la transformée de Fourier discrète augmente[76] avec la taille de l'image aussi on compare deux cas : le premier où on ne conserve de l'image que 128x128 pixels et un deuxième où l'on inclut 256x256 pixels.

76. sans zero-padding

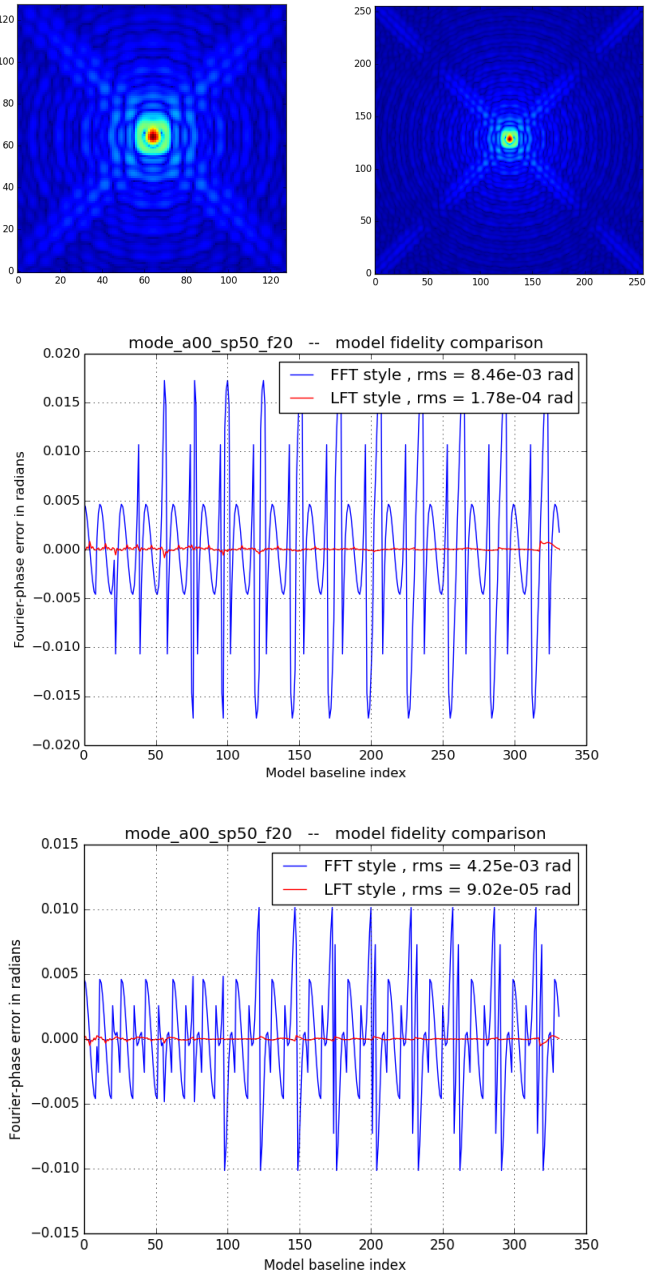


FIGURE 60: Comparaison de la fidélité de la procédure d'extraction de la phase de Fourier. La rangée du haut présente les deux images de travail, la première contenant les 128x128 pixels centraux d'une image idéale sans bruit et la seconde proposant, avec le même échantillonnage, 256x256 pixels de cette même image. La rangée centrale compare pour l'image 128x128, le résidu de phase pour la FFT (en bleu) à celui de la LFDT (en rouge). L'effet d'aliasing induit par l'arrondi au pixel entier le plus proche dans le cas de l'extraction à base de FFT est visible dans les oscillations pseudo-périodiques observées sur le résidu. La LDFT offre une meilleure performance. La rangée du bas présente les mêmes informations, pour l'image 256x256. Doubler la taille de l'image réduit le résidu des deux approches par un facteur deux. La LDFT offre dans les deux cas de figure, une meilleure fidélité par rapport au résultat attendu.

Dans les deux cas, on observe (voir figure 60) que le résidu, défini comme la différence entre la phase extraite par les deux algorithmes

FFT et LDFT et la phase théorique attendue pour l'objet simulé est systématiquement en faveur de la LDFT, par un facteur $\sim$50. Des expériences supplémentaires incluant du zéro-padding réduisent le résidu de l'algorithme à base de FFT, par un facteur $\sim$2 en doublant (avec des zéros) la taille du tableau. Regagner l'avantage de justesse offerte par l'approche LDFT demanderait beaucoup de zéro-padding : j'utiliserai par conséquent désormais ce nouvel algorithme.

Démonstration de l'intérêt des kernels

Maintenant qu'on dispose d'un algorithme propre d'extraction de la phase dans le plan de Fourier, on peut se lancer dans quelques expériences qui vont montrer l'intérêt de toute la cuisine qui est élaborée dans cette partie du manuscrit. On va pour cela, simuler l'acquisition de l'image d'un système binaire, simple à modéliser, dont on va chercher à détecter la signature dans l'espace de Fourier. L'objet est caractérisé par les paramètres suivants :

— séparation angulaire : 31.63 mas

— angle de position : 18.43 degrés

— contraste : 0.25.

L'objet est observé par un télescope de 8 mètres, identique à celui décrit dans la figure 59. Ces valeurs de paramètres ont été choisies de façon à faire tomber la position du compagnon sur un nombre entier de pixels (échelle focale de 10 mas par pixel) sans être dans la direction priviligiée des pixels. La simulation est cette fois affectée par un écran de phase instrumentale, caractérisé par un spectre de type Kolmogorov mais avec une relativement faible amplitude (140 nm RMS). Cet écran de phase et l'image qui en résulte (simulation à signal sur bruit infini) sont représentés dans la figure 61. L'image fait 256 pixels de côté : l'erreur numérique estimée sur la phase extraite dans l'espace de Fourier est de l'ordre de 10^{-4} radians.

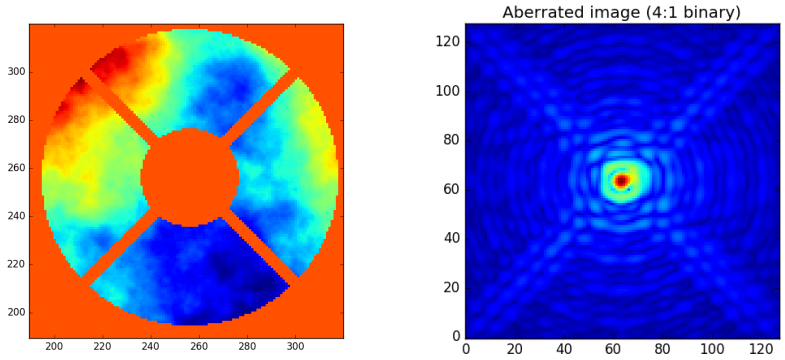


FIGURE 61: A gauche : structure de l'écran de phase projeté sur la pupille du télescope introduisant 140 nm d'aberration RMS. A droite : image simulée d'un système binaire affectée par cet écran de phase. Le compagnon dont le contraste est de 4 :1 pour une séparation angulaire plus petite que λ/D n'est pas directement distinguable à cause des aberrations.

Dans l'espace de Fourier (voir la figure 62), la signature binaire attendue : une modulation sinusoïdale de la phase le long de la direction donnée par l'angle de position, est rendue invisible par la

trace de la phase instrumentale. A l'instar de la phase de franges d'interférences affectées par du piston atmosphérique (voir chapitre 1), la phase de Fourier, sous cette forme brute est inexploitable.

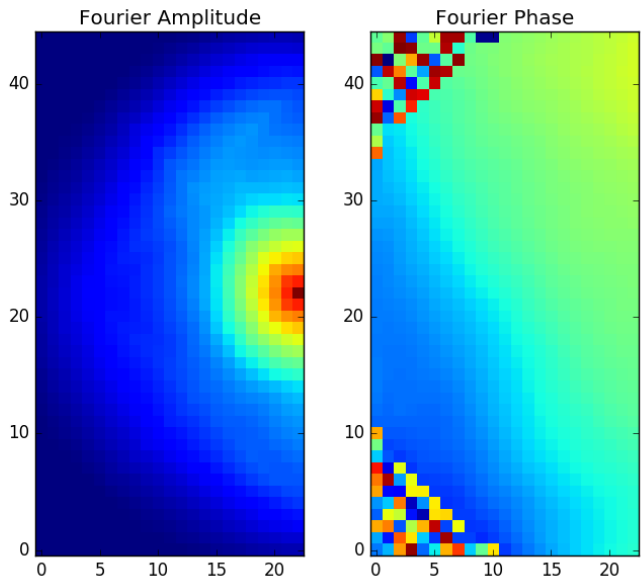


FIGURE 62: Traduction, dans l'espace de Fourier, de l'information contenue dans l'image qui vient d'être présentée.

Bien que dominante dans l'espace de Fourier, l'erreur de phase instrumentale est suffisamment faible pour justifier la linéarisation ayant conduit à l'obtention de l'équation 40. La procédure de recherche automatisée des kernels (à partir de la SVD de la matrice de transfert), conduit à l'obtention d'une matrice $\mathbf{K}$ dont les lignes contiennent les relations de clôture. Cette matrice, appliquée par la gauche aux données, conduit théoriquement à l'obtention de quantités qui ne dépendent plus des erreurs de phase instrumentale, tels que :

$$\mathbf{K} \cdot \Phi = \mathbf{K} \cdot \Phi_O. \tag{52}$$

Les figures 63 et 64 comparent chacune : d'un côté, la phase brute extraite du plan de Fourier pour les coordonnées u,v du modèle à celle théoriquement attendue pour l'objet simulé ; et de l'autre côté, les kernels construits à partir de ces mesures brutes à ceux attendus pour ce même objet simulé. L'allure générale en dent de scie de la courbe de phase dans l'espace de Fourier est due à la convention retenue pour ordonner les informations extraites : de la gauche vers la droite et du haut vers le bas dans le plan représenté dans la figure 62.

Si l'on compare la déviation standard entre les phases brutes expérimentales et attendues ($\sigma_{raw} = 0.6$ radians) à celle entre les kernels expérimentaux et attendus ($\sigma_K = 0.03$ radians), l'intérêt des kernels est évident : au prix d'une réduction de $\sim 20\%$ du nombre total de quantités observables (de 332 phases uv à 262 kernels), on récupère des informations qui sont des estimateurs justes des propriétés de la

cible observée et qui peuvent être utilisés par exemple pour ajuster les valeurs des paramètres d'un modèle.

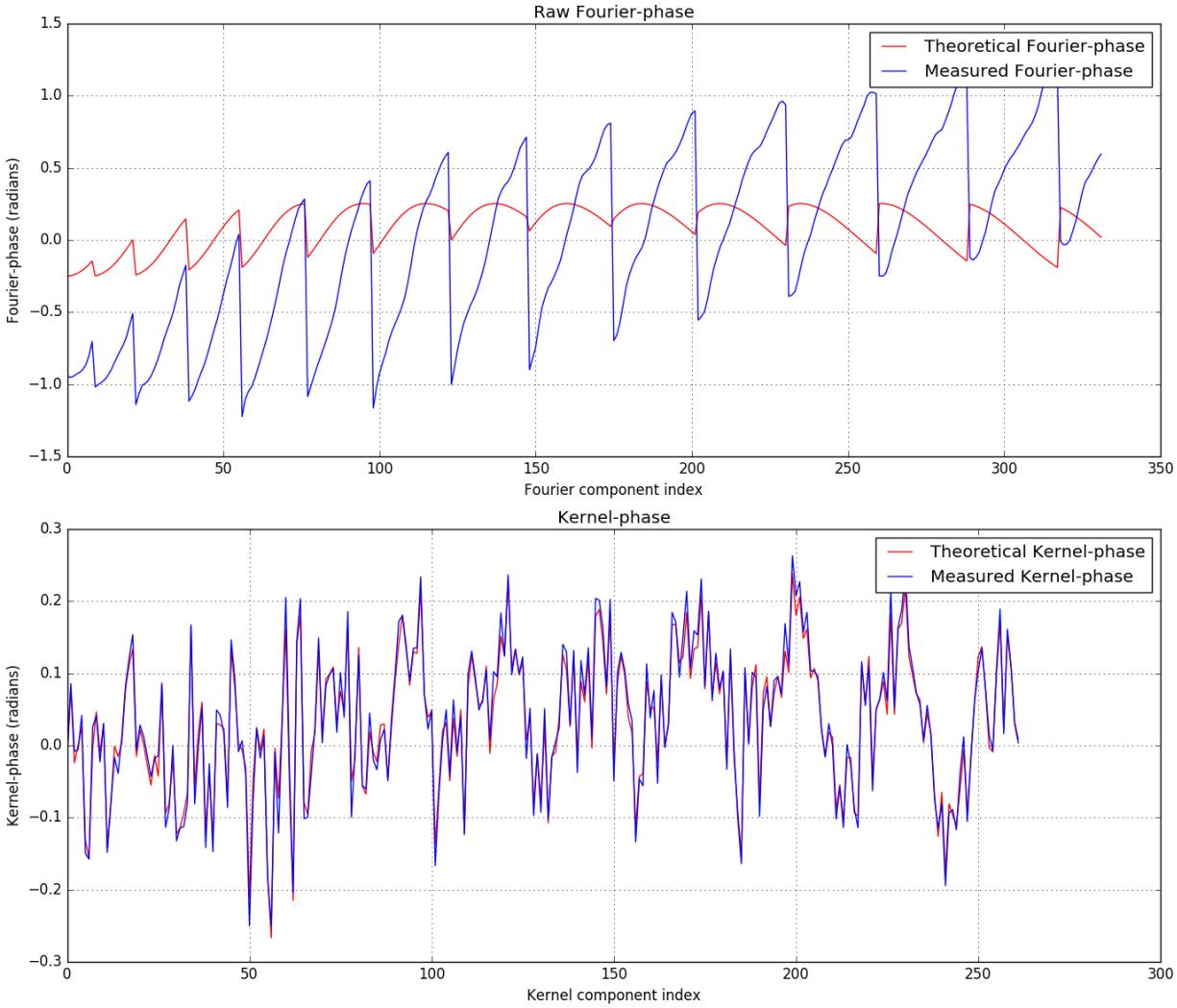


FIGURE 63: Figure du haut : comparaison entre la phase théorique de Fourier attendu le système binaire simulé (courbe rouge) et celle effectivement mesurée lorsqu'une erreur de phase instrumentale s'y retrouve couplée (courbe bleue). Figure du bas : comparaison semblable, cette fois ci entre les kernels de phase théoriquement attendus (en rouge) pour cet objet et ceux effectivement construits à partir de la phase de Fourier brute. La meilleure adéquation entre les deux courbes de kernels justifie leur utilisation.

La représentation des kernels proposée par la figure 63 n'est pas particulièrement facile à interpréter : contrairement à la clôture de phase pour laquelle les astronomes interférométristes ont réussi à développer une compréhension intuitive, les kernels, même s'ils mesurent quelque-chose de similaire, sont une construction plus abstraite. Nous verrons plus loin comment on peut utiliser les kernels pour rapidement produire des pseudo-images des sources observées. Pour le moment, on se contentera de ce genre de représentation, ou de représentation sous forme de diagramme de corrélation, tel que proposé dans la figure 64.

Deux publications sont incluses ici car elles présentent deux cas d'utilisation de cette méthode à des fins astrophysiques. La méthode est mise en oeuvre sur des données d'archive de la caméra NICMOS1 embarquée sur HST. Le premier de ces articles permet de revisiter l'étoile binaire GJ164, décrite dans la première partie de ce manuscrit.

La deuxième publication, est un travail que j'ai supervisé, utilise cette même approche sur un relevé homogène de naines brunes de type L, également observées avec HST/NICMOS1.

The Astrophysical Journal, 724:464–469, 2010 November 20 doi:10.1088/0004-637X/724/1/464
© 2010. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

KERNEL PHASE IN FIZEAU INTERFEROMETRY

Frantz Martinache
Subaru Telescope, Hilo, HI, USA
Received 2010 June 20; accepted 2010 September 16; published 2010 November 2

ABSTRACT

The detection of high contrast companions at small angular separation appears feasible in conventional direct images using the self-calibration properties of interferometric observable quantities. The friendly notion of closure phase, which is key to the recent observational successes of non-redundant aperture masking interferometry used with adaptive optics, appears to be one example of a wide family of observable quantities that are not contaminated by phase noise. In the high-Strehl regime, soon to be available thanks to the coming generation of extreme adaptive optics systems on ground-based telescopes, and already available from space, closure phase like information can be extracted from any direct image, even taken with a redundant aperture. These new phase-noise immune observable quantities, called kernel phases, are determined a priori from the knowledge of the geometry of the pupil only. Re-analysis of archive data acquired with the *Hubble Space Telescope* NICMOS instrument using this new kernel-phase algorithm demonstrates the power of the method as it clearly detects and locates with milliarcsecond precision a known companion to a star at angular separation less than the diffraction limit.

Key words: binaries: close – methods: data analysis – stars: low-mass – techniques: high angular resolution – techniques: image processing – techniques: interferometric

1. PHASE IN THE FOURIER PLANE

Only two parameters essentially determine whether a source is detectable during an observation: its brightness at the wavelength λ of interest and the angular resolution necessary to separate the source or feature from its direct environment. The angular resolution is ultimately constrained by the diffraction of the telescope, and astronomers usually follow the rule of thumb known as the Rayleigh criterion, stating that to be resolved, two sources need to be separated by $1.22\lambda/D$, where D is the diameter of the telescope used, to design their observations.

The development of optical interferometry has however made this criterion obsolete: thanks to the exquisite level of calibration it permits, interferometry indeed makes it possible to detect sources or constrain the extent of features around objects at separations significantly smaller than the diffraction limit. Even at the scale of one single telescope, the results obtained with the technique known as non-redundant masking (NRM) interferometry, first, seeing-limited (Haniff et al. 1987; Readhead et al. 1988) and more recently used with adaptive optics (AO) systems (Tuthill et al. 2000, 2006; Lloyd et al. 2006; Ireland et al. 2008; Kraus et al. 2008; Martinache et al. 2009) demonstrate the relevance of this technique for the detection of structures at small angular separation, that would not be accessible from conventional AO images (Rajagopal et al. 2004).

Even if it only uses one single telescope in order to reach this level of resolution, one however needs to accept that the familiar product called "image" may not necessarily constitute the best final data product. Instead, when interested in high-angular resolution properties of partially resolved objects, it is convenient to derive information not from the image itself, but from its Fourier-transform counterpart. This information, known as complex visibility, is extracted from the Fourier plane, calibrated and then tested against a model of the observed object.

In optical interferometry, this approach is often mandatory: the paucity of apertures ($N \sim 2$–5) and baselines make the content of a direct (Fizeau) image of limited value. Information-rich images can be reconstructed after extraction of the complex visibility function from the (u, v)-plane, but only with a large ($N > 10$) number of apertures as in radio interferometry, or after using image synthesis. The optical image reconstruction known as pupil densification that is used in hypertelescopes (Labeyrie 1996) does provide an alternative, but again, only becomes compelling if a large number of apertures are used (Labeyrie et al. 2008). But even when an image can be reconstructed from optical interferometry measurements, e.g., the images of the binary Capella by Baldwin et al. (1996), the intensity map of the surface of Altair by Monnier et al. (2007) or the spectacular images of the disk eclipsing ϵ Aurigae (Kloppenborg et al. 2010), quantitative characteristics of the sources can only be deduced from the fit of the interferometric data by parametric models. In the case of a marginally resolved binary star, precise measurements of angular separation, orientation and contrast, with confidence intervals, deduced from a model fit of complex visibilities carry much more scientific value than an image of "blurry blobs."

Visibilities in the Fourier plane are complex numbers, whose amplitude and phase are usually considered separately. This paper focuses on the treatment of the phase and ignores the amplitude. In general, the power contained at given spatial frequency is the result of the coherent sum of R random phasors, with R being a scalar coding the redundancy of the spatial frequency. In the presence of residual optical path differences (OPDs), this coherent sum of R random phasors loses the phase information and results in the formation of speckles in seeing-limited images with a visible/IR telescope. NRM interferometry solves this problem, by discarding light with a pupil mask designed so that each baseline is unique ($R = 1$), which makes the extraction of the phase possible.

The phases alone, being corrupted by residual OPDs, are of restricted interest. It is however possible to combine them to form what is known as closure phase (Jennison 1958), that is the sum of three phases measured by baselines forming a closed triangle. This remarkable interferometric quantity (cf. the introduction to closure phase by Monnier 2000) exhibits a compelling property: it rejects all residual pupil-plane phase errors. Moreover, because it is determined from the analysis of the final science detector and not on a separate arm wavefront

sensor, it is also immune to non-common path errors between the wavefront sensor and the science camera.

Once extracted and calibrated, the closure phases can then be compared to a parametric model, for instance of a binary star, to confirm or infirm the presence of a companion around a given source, while uncertainties provide contrast detection (i.e., sensitivity) limits. This approach was successfully used by Lloyd et al. (2006), Martinache et al. (2007), Ireland et al. (2008), Kraus et al. (2008), and Martinache et al. (2009), who typically report sensitivity of 5–6 mag in the near-infrared at separations ranging from 0.5 to 4 λ/D.

This paper aims at generalizing the notion of closure phase, and shows that closure phase like quantities, i.e., sharing the same property of independence to pupil-plane phase errors, can be constructed even in the case of redundant apertures.

2. KERNEL PHASE

2.1. Linear Model

Whether contiguous (i.e., single dish) or not (i.e., interferometric), the two-dimensional pupil of an imaging system can be discretized into a finite collection of N elementary sub-apertures. One of these elementary sub-apertures taken as zero-phase reference, the pupil-plane phase of a coherent point-like light source can be written as an $(N-1)$-component vector φ. Given that the image, or interferogram, of this source is sufficiently sampled, then in the Fourier plane (a.k.a. (u, v)-plane in interferometry) one will be able to sample up to M phases, where M is a function of the pupil geometry only. For a non-redundant array made of N elementary sub-apertures, the number of sampled (u, v) phases is maximum $M = \binom{N}{2}$. The same number of sub-apertures organized in a redundant array, for instance following a regular grid, produces significantly less distinct (u, v) sample points as each point receives the contribution of several pairs of sub-apertures.

In most cases, since each point receives the sum of several random phasors, both phase and amplitude are lost and cannot be simply retrieved: this results in the formation of speckles. However, if the Strehl is high enough, the complex amplitude associated with the instrumental phase in one point of the pupil, φ_k, can be approximated by $e^{i\varphi_k} \approx 1 + i\varphi_k$. Direct application of the approach is therefore for now restricted to space-borne diffraction-limited optical and mid-IR telescopes like *HST* (cf. Section 3), but should also prove relevant to the upcoming generation of extreme AO systems.

Given that the proposed approximation holds, while observing a point source, the unknown (instrumental) phase distribution in the pupil φ can be related to the phases Φ measured in the Fourier plane with a single linear operator. To build an intuitive understanding of this relation, let us consider the following scenarios.

1. If the phase is constant across the entire pupil, then none of the baselines formed by any pair of elementary sub-apertures does record a phase difference, and the phase in the Fourier plane is zero everywhere.
2. If a phase offset δ_ϕ is added to one single sub-aperture, then each baseline involving this aperture records a phase difference, which is exactly $\pm\delta_\phi$. Figure 1 represents several such scenarios.
3. If the pupil-plane phase vector φ is completely random, each of the M samples in the Fourier plane is then the average of R phase differences on the pupil, where R is the redundancy of the considered baseline.

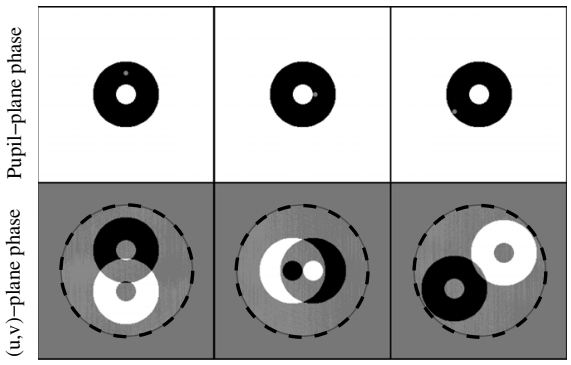


Figure 1. Iterative process for the determination of the transfer matrix **A**. The top row shows the sub-aperture of the full two-dimensional pupil (circular aperture with 30% central obscuration), where a phase offset is applied (three cases are represented). The bottom row shows the resulting distribution of phase in the Fourier plane. The overlaid dashed-line circle in the bottom row marks the cutoff spatial frequency of the transfer function.

To reproduce this behavior, the following linear model will be used:

$$\Phi = \mathbf{R}^{-1} \cdot \mathbf{A} \cdot \varphi, \tag{1}$$

where Φ represents the M-component Fourier plane phase vector, $\mathbf{R}$ an $M \times M$ diagonal matrix whose diagonal elements code the redundancy of the baselines, and $\mathbf{A}$ represents an $M \times (N-1)$ transfer matrix, whose properties form the core of the discussion of this work. To be complete, the model should also include the phase information intrinsic to the observed source, represented by the term Φ_O that simply adds on top of the instrumental phase. One can then multiply both sides of the equation by the matrix $\mathbf{R}$ so that it becomes

$$\mathbf{R} \cdot \Phi = \mathbf{A} \cdot \varphi + \mathbf{R} \cdot \Phi_O. \tag{2}$$

While $\mathbf{R}$ and $\mathbf{A}$ could have been merged into one single operator, they are intentionally kept distinct. The rationale for this choice is so that the left-hand side of Equation (2), i.e., the measurements, can be acquired by reading directly the imaginary part of the complex visibility. Given that the next (quadratic) term in the Taylor expansion of $e^{i\varphi}$ being real, this makes the approximation valid to the third order in phase. This also makes $\mathbf{A}$ of striking aspect as it is then exclusively filled with values 0, 1, or -1.

If the matrix $\mathbf{A}$ were invertible, then the analysis of one unique focal plane image of a single star (case corresponding to Equation (1)) would be sufficient to determine the instrumental phase φ as seen from the detector, and drive an AO system and/or delay lines. Except for the special case of a non-redundant aperture, the problem is however known to be degenerate, despite the larger number of measures than unknowns ($M > N-1$).[1]

As demonstrated by the successes of NRM interferometry, a complete characterization of the wavefront is not essential if one can determine observable quantities that are pupil-phase independent. The closure relations used in interferometry can be related to the operator $\mathbf{A}$: these relations are simply linear combinations (modelized by an operator $\mathbf{K}$) of rows of $\mathbf{A}$ that

[1] The use of this model for wavefront sensing purposes will be the object of another paper.

produce 0, forming something known as the left null space of A:

$$\mathbf{K} \cdot \mathbf{A} = \mathbf{0}. \quad (3)$$

For a non-redundant array, each closure relation will fill a row of **K** with mostly zeroes, except in three positions corresponding to the baselines forming a closing triangle, that will contain 1 or −1. These relations are however not the only possible ones, and less trivial combinations, involving more than three rows at a time, can be constructed. The total number of independent relations however remains unchanged and is only imposed by the geometry of the array.

Although not impossible, finding the operator **K** "by hand" (i.e., finding a basis for the left null space of **A**) for a redundant aperture is a tedious task, as the matrix **A** can get quite large. A very efficient way to do this is to calculate the singular value decomposition (SVD) of $\mathbf{A}^T$. The SVD algorithm (Press et al. 2002) allows us to decompose the now $(N-1) \times M$ matrix $\mathbf{A}^T$ as the product of an $(N-1) \times M$ column-orthogonal matrix **U**, an $M \times M$ diagonal matrix **W** with positive or zero elements (the so-called singular values), and the transpose of an $M \times M$ orthogonal matrix **V**:

$$\mathbf{A}^T = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T. \quad (4)$$

One relevant property of the SVD is that it explicitly constructs orthonormal bases for both the null space and the range of the matrix $\mathbf{A}^T$. Of particular interest here are the columns of **V** that correspond to singular values equal to zero: these vectors form an orthonormal base for the null space, also referred to as Kernel of $\mathbf{A}^T$, that is exactly what is needed to fill in the rows of **K**.

If the observed target is not perfectly symmetric, and exhibits actual phase information (i.e., $\Phi_O \neq 0$, see for instance Monnier 2000), Equation (2) is required. Multiplying it with the left side operator **K** leads to a new series of new phase-like quantities that are not contaminated by instrumental phase, generalizing the notion of closure phase (Baldwin et al. 1986) on which NRM interferometry from the ground (Tuthill et al. 2000) and from space (Sivaramakrishnan et al. 2009) entirely rely.

While not as immediately tangible as the notion of closure phase, this proposed generalization, hereafter referred to as kernel phase (or Ker-phase) since it relates to the kernel of the matrix $\mathbf{A}^T$, exhibits a unique advantage over the classical closure phase: it is not restricted to non-redundant apertures and makes it possible to extract phase-residual immune information from images acquired from telescopes of arbitrary pupil geometry.

This has some obvious advantages over the restrictive non-redundant case:

1. Throughput: non-redundant aperture masks exhibit a typical 5%–15% throughput (Martinache et al. 2007), and photon noise of the companion one tries to detect may be the dominant source of noise. Given that it benefits from the same phase-noise canceling properties as closure phase, for a given readout noise and exposure time, kernel phase on an unmasked aperture offers an immediate boost in sensitivity (or dynamic range) on faint sources.
2. Number of observable quantities: a common non-redundant aperture mask design exhibits nine sub-apertures, therefore forming $\binom{9}{2} = 36$ baselines and $\binom{8}{2} = 28$ independent closure phases (Monnier 2000). More independent kernel phases can be extracted from the Fourier transform of a full-aperture image, which will provide a better characterization of the target.

Another incidental advantage is that, being a product of the SVD, all the kernel-phase relations contained in **K** form an orthonormal basis, and therefore do not introduce correlation in the data. A consequence is that manipulating Ker-phases does not require keeping track of the covariance matrix used for closure phases in masking interferometry, which simplifies their interpretation.

2.2. Calibration

In discretizing the pupil into a finite number of sub-apertures, one important assumption is made: the phase (or more generally, the complex amplitude of the electric field) is assumed to be uniform within each sub-aperture. Yet even for a space-borne telescope, in the absence of atmosphere, this is only an idealization as small-scale structures, like polishing imperfections of the primary mirror for instance, will impact, to some extent, the value of the Ker-phases. This issue is not proper to Ker-phases and also affects closure phases. Thus, unless perfect (i.e., single mode) spatial filtering is performed within each sub-aperture of a non-redundant array, the closure phase on a point source is not exactly be zero.

This effect can somewhat be mitigated by subtracting from the Ker-phases of a science target, the Ker-phase signal measured on a point source observed in identical conditions. NRM interferometry results reported in Martinache et al. (2009), for instance, make extensive use of this kind of calibration: from the ground, this approach is very powerful as it makes it possible to calibrate other sources of systematic errors like the effect of broadband filters which smear out the Fourier plane and differential atmospheric refraction. From space, this may not be as essential depending on the science goal: if the Ker-phases obtained on a binary system are non-calibrated, then they will simply contain a systematic error term that will limit the achievable contrast.

3. KERNEL-PHASE ANALYSIS OF *HST*/NICMOS DATA

While the kernel-phase approach may prove difficult to apply to ground-based observations until extreme AO become available, it can readily be applied to diffraction-limited observations made from space. It is tested here on a series of non-coronagraphic narrowband images acquired with the Near-Infrared Camera and Multiobject Spectrometer (NICMOS) onboard the *Hubble Space Telescope*. Two data sets acquired with the NICMOS1 in the *F190N* filter on two distinct objects are used: the first target is a calibration star, SAO 179809, which was observed in 1998; the second is the high-proper motion star GJ 164, around which a companion was astrometrically discovered and whose existence was confirmed after point-spread function (PSF) modelization and subtraction of these NICMOS1 images by Pravdo et al. (2004). This latter target is an ideal benchmark: given its expected <10:1 luminosity contrast, one should expect a strong, unambiguous Ker-phase signal.

Moreover, ground-based infrared aperture masking interferometry measurements reported by Martinache et al. (2009) combined with the astrometry have led to strong constraints on the orbit of the companion around the primary. The location of GJ164 B measured from the Ker-phase analysis of the data can be compared to the orbit prediction.

Figure 2 shows the model of the pupil used for this exercise. The *HST* pupil exhibits a 30% central obscuration as well as 90° spider arms (actual dimensions were taken from the NICMOS1 configuration file in the TinyTim PSF simulation package for

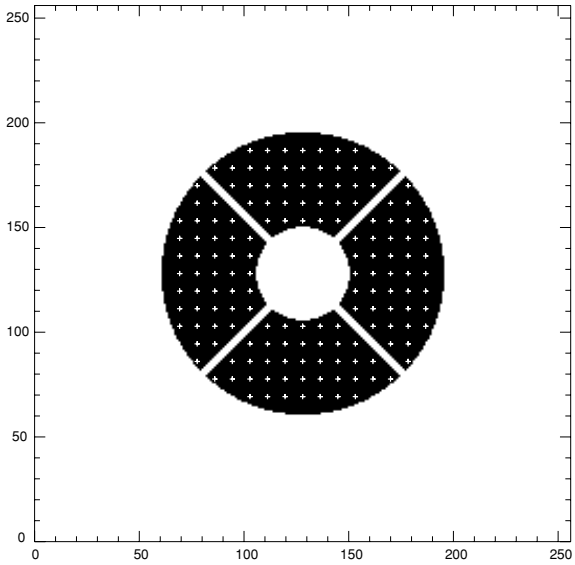


Figure 2. Model for the geometry of the *HST*/NICMOS pupil and location of the sample points for the determination of the Ker-phase relations. The 156 sample points of the pupil fall on a regular square grid with a step of 1/16 of a pupil diameter that do not intersect with the central obstruction or the spider arms.

HST). The phase across the pupil is discretized into a 156 elementary sub-aperture array, whose locations fall on a regular square grid of step of 1/16 of the outer pupil diameter. The phase sample is assembled into a 155-component $(N-1)$ vector φ.

These 156 pupil phase samples map in the (u, v)-plane onto a square grid of 366 distinct elements.[2] The resulting (u, v)-sampling is illustrated in Figure 3. For this analysis, **A** (cf. Equation (1)) is therefore a 155×366 rectangular matrix, whose SVD reveals that 78 singular values are non-zero, leaving $366 - 78 = 288$ Ker-phase relations.

The GJ164 data consist of a total of 80 frames, acquired at average Julian Date 2453049.3 (2004 February 14 UT). Each image is a non-saturated 32 s exposure, and the target was acquired in a total of 10 distinct dither positions. Note that this data set does not include images on a point source and therefore, the Ker-phases calculated from this data set are non-calibrated. Images corresponding to one dither position were simply co-added forming a final total of ten 250 s exposure images, and assembled into a data cube. The SAO 179809 data consist of four distinct 20 s exposure frames assembled into a separate data cube.

For both data cubes, the images were then centered with sub-pixel accuracy and windowed by a super-Gaussian function as described by Kraus et al. (2008) to limit sensitivity to readout noise. The window size is about $25\lambda/D$ in diameter, which is significantly larger than the field of view in which this technique is relevant.

After this preparatory stage, the images are simply Fourier-transformed (cf. second panel of Figure 3), and the signal $\mathbf{R} \cdot \Phi$ is directly measured for each of the 366 (u, v) points by sampling the imaginary part of the local complex visibility. The uncertainty associated with the measurement of each phase is estimated from the dispersion of the signal in the direct neighborhood of the (u, v) point.

The (u, v) signals are then assembled into Ker-phases using the relations gathered in the rows of **K** and uncertainties are propagated. The procedure is repeated for each of the frames within each data cube. The final retained series of 288 Ker-phases is the weighted average for all frames.

Because the Ker-phase relations are designed to produce quantities independent from pupil phase errors, a point source is expected to exhibit zero signal within uncertainty. Despite the small number of statistics (four frames acquired on SAO 179809), the Ker-phase of the calibrator do average to zero (with a 19°.7 standard deviation), while the binary exhibits a large signal amplitude (>100°) in comparison with the uncertainty of individual Ker-phases (∼2°). The third panel of Figure 3 compares the Ker-phase histograms of both data sets.

To further investigate the GJ 164 data, a parametric model of the (u, v)-plane phase Φ_O for a binary star is needed. The parameters are the angular separation, the position angle of the secondary relative to the primary, and the luminosity contrast ratio. The model phase Φ_O is then multiplied by the diagonal matrix **R**, and finally, transformed into model Ker-phases using the relations established during the SVD.

The agreement between the data and the model is very good (cf. panel 4 of Figure 3), considering the large number of measurements (288) adjusted by only three parameters. The uncertainties on the Ker-phases, determined from the scatter of

[2] Note for reference that a non-redundant array of 156 sub-apertures would produce exactly 12,090 distinct (u, v) points.

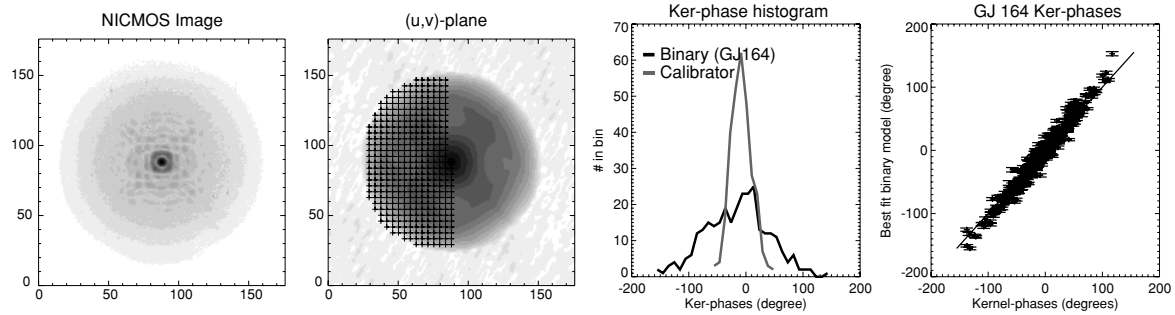


Figure 3. From left to right: example of a narrowband ($F190N$) NICMOS1 image used for this work, visualized with a nonlinear brightness scale; the Fourier transform of this image. The 366 sample points for the phase in the Fourier domain are overlaid; a comparison of histograms of the 288 Ker-phases calculated using the relations identified in Section 2.1. By design, the Ker-phases calculated from images of a single star are expected to be zero within uncertainty: the corresponding histogram (gray curve) confirms this expectation. In comparison, the Ker-phase histogram of the binary (dark curve) appears significantly larger. The same GJ164 Ker-phases plotted against the model of a binary star that best fits the data convince of the presence of a companion in the data.

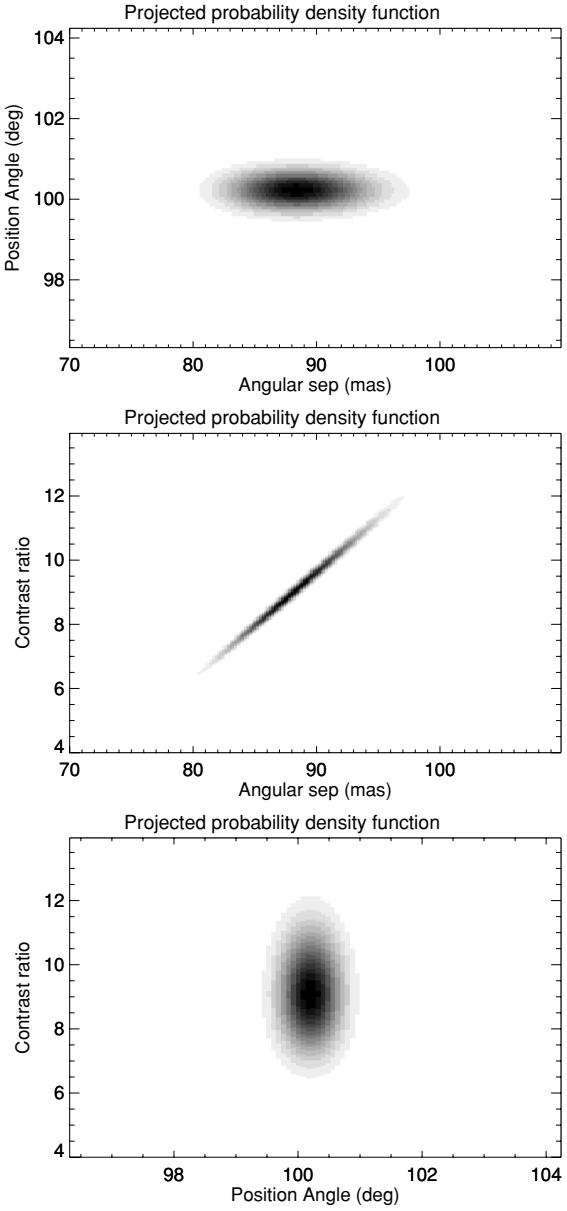


Figure 4. To determine confidence intervals for the parameters of the binary, a likelihood analysis comparable to the one presented by for closure phase was performed. These panels show the three projections of the likelihood function in the region of best fit. Except for the expected correlation between angular separation and contrast ratio for a detection within 1 λ/D, the solution is unambiguous and well constrained, demonstrating the elegance of the Ker-phase approach.

the data however, lead to a best-fit reduced χ^2 larger than 1. A global error term (10°) is then added in quadrature to the uncertainty to account for a systematic error in the non-calibrated Ker-phase and produce a final reduced $\chi^2 = 1$.

One can then proceed with determining the uncertainty on the parameters of the model fit, by close examination of the likelihood function, very much like what is described in Martinache et al. (2009). The three panels of Figure 4 show the

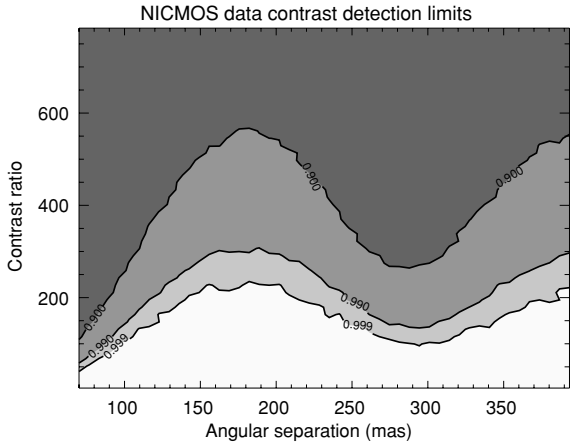


Figure 5. Level of confidence in the detection of a companion from the analysis of the *HST*/NICMOS data with the Ker-phase algorithm. A darker color indicates a region of lower confidence level. Three levels are highlighted: the 90%, 99%, and 99.9% confidence levels. At angular separation 0.5 λ/D (i.e., 80 mas at $\lambda = 1.9\,\mu$m), a contrast limit better than 50:1 is possible at the 99% confidence level.

Table 1
Ker-Phase Detection of GJ164B in Nicmos Data Compared to Prediction from Orbital Parameters

Parameter	Ker-phase fit	Prediction
Separation (mas)	88.5 ± 3.6	88.2
P.A. (deg)	100.6 ± 0.3	100.4
Contrast	9.1 ± 1.2	

evolution of this function in the parameter space region near the best solution. Just like with closure phase data, at angular separations less than 1 λ/D, contrast and separation appear to be correlated.

The uncertainty on each parameter of the model fit is determined after marginalization of the likelihood function over the other two parameters. Despite the noted correlation, the constraint on the parameters appears satisfactory, and the best fit (cf. Table 1) lies well within 1σ of the position predicted from the orbital parameters determined from NRM interferometry from the ground. It also matches the location reported by Pravdo et al. (2004), after subtraction of a simulated PSF from the same data, only with a constraint on the position angle improved by a factor of 10.

From its $(H-K)$ color index, Martinache et al. (2009) were able to conclude that GJ 164 B is of spectral type later than M8.5, while the primary is well characterized as an M4.5 dwarf. One of the most prominent spectral features for M dwarfs is the broad absorption band of water at 1.8 μm, getting deeper with later types (Jones et al. 1994; Leggett et al. 2001). The ∼5:1 contrast ratio quoted in the NRM paper was determined over the full K_s filter (bandwidth 2.0–2.3 μm). A careful examination of the spectral sequence by (Jones et al. 1994) reveals that for this combination of spectral types, the luminosity of GJ 164B relative to GJ164A seen in the NICMOS F190N filter is expected to drop by 30%–40% due to the water absorption band. The 9:1 contrast determined from the Ker-phase model (cf. Table 1) in this narrow filter reflects this evolution. The analysis of this GJ164 data demonstrates the validity of the Ker-phase approach, by positively detecting a companion whose existence was known beforehand. This <10:1 contrast detection was how-

ever expected to be easy, despite the small angular separation of the detection (0.6 λ/D).

Typical NICMOS1 data sets on a given target usually consist of four frames only. The SAO 179809 data set is then a representative example and the statistics of its Ker phase ($\sigma = 19.^{\circ}7$) can be used in a Monte Carlo simulation to determine contrast detection limits.

Because the sampling of the (u, v)-plane exhibits no gap, the sensitivity does not depend on the position angle relative to the central star. One can however expect it to be a function of angular separation. A total of 10,000 simulations were performed per point in the angular separation/contrast plane to produce the sensitivity map displayed in Figure 5. The map highlights the 90%, 99%, and 99.9 % confidence level detection thresholds.

The technique looks promising: for such a data set, at 0.5 λ/D, a 50:1 contrast detection appears possible at the 99% confidence level. The sensitivity increases and peaks at 180 mas, which unsurprisingly corresponds to the location of the first zero of the diffraction for the centrally obstructed telescope (about 1.1 λ/D), and reaches ∼200:1.

4. CONCLUSION

Classical closure phase appears to be one special case of a wider family of observable quantities that are immune to phase noise and non-common path errors. In the high-Strehl regime, it was demonstrated that closure phase like quantities, called Ker-phases, can be extracted from focal plane images, and provide high-quality "interferometric grade" information on a source, even when the pupil is redundant. The Ker-phase technique was successfully applied to a series of archive NICMOS images, clearly detecting a 10:1 contrast companion at a separation of 0.5λ/D. Non-calibrated Ker-phase appears sensitive to the presence of 200:1 contrast companion at angular separation 1λ/D. Re-analysis of other comparable NICMOS data sets with this technique might very well lead to the detection of previously undetected objects in the direct neighborhood of nearby stars.

Unlike closure phases, which are extremely robust to large wavefront errors, the use of Ker-phases is however for now restricted to the high-Strehl regime, and will only become relevant to ground-based observations when extreme AO systems become available. There is nevertheless hope to be able to extend the application of Ker-phases to not-so-well corrected AO images, using additional differential techniques. One possibility, consists in using integral field spectroscopy to follow in the Fourier plane the evolution of the complex visibilities as a function of wavelength. With enough resolution and spectral coverage, this indeed allows us to identify the phasors contributing to the power contained at one spatial frequency.

The SVD of the transfer matrix used to create Ker-phase relations can also be used to produce a pseudoinverse to the matrix, and in some cases, allows us to inverse the relation linking the (u, v) phases to the pupil phases. This means that under certain conditions, a single monochromatic focal plane image can also be useful for wavefront sensing purposes. This is particularly interesting since the measurement is happening at the level of the final science detector, which therefore allows us to calibrate non-common path errors. The application of the formalism to wavefront sensing will be the object of a future publication.

The author thanks Michael J. Ireland and Olivier Guyon for the useful discussions of the ideas presented in this work. This work was supported in part by the Jet Propulsion Laboratory under contract 1379504.

REFERENCES

Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Nature, 320, 595
Baldwin, J. E., et al. 1996, A&A, 306, L13
Haniff, C. A., Mackay, C. D., Titterington, D. J., Sivia, D., & Baldwin, J. E. 1987, Nature, 328, 694
Ireland, M. J., Kraus, A., Martinache, F., Lloyd, J. P., & Tuthill, P. G. 2008, ApJ, 678, 463
Jennison, R. C. 1958, MNRAS, 118, 276
Jones, H. R. A., Longmore, A. J., Jameson, R. F., & Mountain, C. M. 1994, MNRAS, 267, 413
Kloppenborg, B., et al. 2010, Nature, 464, 870
Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762
Labeyrie, A. 1996, A&AS, 118, 517
Labeyrie, A., Le Coroller, H., & Dejonghe, J. 2008, Proc. SPIE, 7013, 107
Leggett, S. K., Allard, F., Geballe, T. R., Hauschildt, P. H., & Schweitzer, A. 2001, ApJ, 548, 908
Lloyd, J. P., Martinache, F., Ireland, M. J., Monnier, J. D., Pravdo, S. H., Shaklan, S. B., & Tuthill, P. G. 2006, ApJ, 650, L131
Martinache, F., Lloyd, J. P., Ireland, M. J., Yamada, R. S., & Tuthill, P. G. 2007, ApJ, 661, 496
Martinache, F., Rojas-Ayala, B., Ireland, M. J., Lloyd, J. P., & Tuthill, P. G. 2009, ApJ, 695, 1183
Monnier, J. D. 2000, in Principles of Long Baseline Stellar Interferometry, ed. P. R. Lawson (Pasadena, CA: NASA), 203
Monnier, J. D., et al. 2007, Science, 317, 342
Pravdo, S. H., Shaklan, S. B., Henry, T., & Benedict, G. F. 2004, ApJ, 617, 1323
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C: The Art of Scientific Computing (2nd ed.; New York: Cambridge Univ. Press)
Rajagopal, J. K., Barry, R., Lopez, B., Danchi, W. C., Monnier, J. D., Tuthill, P. G., & Townes, C. H. 2004, Proc. SPIE, 5491, 1120
Readhead, A. C. S., Nakajima, T. S., Pearson, T. J., Neugebauer, G., Oke, J. B., & Sargent, W. L. W. 1988, AJ, 95, 1278
Sivaramakrishnan, A., et al. 2009, Proc. SPIE, 7440, 30
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555
Tuthill, P. G., et al. 2006, Proc. SPIE, 6272, 103

The Astrophysical Journal, 767:110 (14pp), 2013 April 20 doi:10.1088/0004-637X/767/2/110
© 2013. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

DANCING IN THE DARK: NEW BROWN DWARF BINARIES FROM KERNEL PHASE INTERFEROMETRY*

Benjamin Pope[1], Frantz Martinache[2], and Peter Tuthill[1]
[1] Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2226, Australia; bjsp@physics.usyd.edu.au, p.tuthill@physics.usyd.edu.au
[2] National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720, USA; frantz@naoj.org
Received 2012 November 1; accepted 2013 February 25; published 2013 April 2

ABSTRACT

This paper revisits a sample of ultracool dwarfs in the solar neighborhood previously observed with the *Hubble Space Telescope*'s NICMOS NIC1 instrument. We have applied a novel high angular resolution data analysis technique based on the extraction and fitting of kernel phases to archival data. This was found to deliver a dramatic improvement over earlier analysis methods, permitting a search for companions down to projected separations of ∼1 AU on NIC1 snapshot images. We reveal five new close binary candidates and present revised astrometry on previously known binaries, all of which were recovered with the technique. The new candidate binaries have sufficiently close separation to determine dynamical masses in a short-term observing campaign. We also present four marginal detections of objects which may be very close binaries or high-contrast companions. Including only confident detections within 19 pc, we report a binary fraction of at least $\epsilon_b = 17.2^{+5.7}_{-3.7}\%$. The results reported here provide new insights into the population of nearby ultracool binaries, while also offering an incisive case study of the benefits conferred by the kernel phase approach in the recovery of companions within a few resolution elements of the point-spread function core.

Key words: brown dwarfs – stars: formation – stars: low-mass – techniques: high angular resolution – techniques: image processing – techniques: interferometric

Online-only material: color figures

1. INTRODUCTION

A detailed picture of multiplicity is an essential element to understanding low-mass stars and brown dwarfs. Binary systems present an opportunity to determine model-independent dynamical masses when both astrometry and radial velocity data are available. Systems so characterized may then become part of the foundations for the construction of an observationally constrained mass–luminosity–age sequence applicable across the entire class.

Furthermore, the statistical properties of populations of low-mass binaries have profound implications on the basic physics of star formation and solar system assembly. Multiplicity rates are a key discriminant between hypotheses about the formation and evolution of low-mass systems, as discussed in Burgasser et al. (2007). Two main mechanisms have been proposed for the formation of brown dwarfs in the field: embryo ejection and gravoturbulent collapse (Basu 2012). Specifically, the embryo ejection hypothesis predicts a low binarity incidence (∼8%; Bate 2012), which conflicts with the observed binarity rate (∼15%; Reid & Hawley 2005). Mapping the incidence of binarity, and in particular extending completeness to smaller orbital separation, is therefore of interest in establishing the primary formation mechanism of field brown dwarfs.

Snapshot imaging is a straightforward way to discover new multiple systems. Intrinsically faint and red, L dwarfs present challenging targets for ground-based observations, typically requiring laser guide star (LGS) adaptive optics (AO). Space telescopes naturally offer high Strehl ratio imaging at diffraction-limited resolution, with the major departures from ideal performance arising from field-dependent point-spread function (PSF) changes, spacecraft jitter, and slow optical drift from thermally induced breathing modes of the mechanical structures.

Imaging campaigns with the *Hubble Space Telescope* (*HST*) have demonstrated notable success in prospecting for companions to cool objects, providing high-quality diffraction-limited images of a large number of targets (Reid et al. 2006a, 2008). These campaigns have shed light on the population of cool dwarfs in the solar neighborhood.

The simplest and most widely used method for detection of companions in snapshot imaging essentially relies on direct visual examination of images. Obvious companions are quickly identified, and traditional astronomical image analysis tools, namely aperture photometry and centroiding, provide the important astrometric and photometric characteristics of the target. Faint or close-in companions are, however, easily missed in a visual search and identifying such objects requires more sophisticated computational techniques.

For example, some stellar images exhibit an elongation along one axis as noted by Reid et al. (2006a) which may be suggestive of the presence of a barely resolved companion. Subtraction of a model PSF has been exploited to infer the presence of a companion (Krist et al. 1998; Pravdo et al. 2004; Dieterich et al. 2012), although the performance of this approach is arguably poor, and furthermore it weakly constrains the relative photometry and astrometry. We propose to look at the same images from an interferometric standpoint, leveraging the exquisite level of calibration this technique offers.

For the detection and characterization of companions at small angular separations, non-redundant masking (NRM) interferometry used in conjunction with AO has demonstrated outstanding performance, e.g., in Tuthill et al. (2006), Lloyd et al. (2006), and Kraus & Ireland (2012). The key underpinning such successes has been the robust, self-calibrating nature of the observables recovered from NRM interferometry, and in particular the *closure phase*, first suggested for the radio

* Based on observations performed with the NASA/ESA Hubble Space Telescope. The Hubble observations are associated with proposal ID 10143 and 10879 and were obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

(Jennison 1958) and later exploited in the optical (Baldwin et al. 1986). Imaging systems where the phase on any given baseline in the pupil is disturbed by random phase errors from atmospheric or instrumental aberrations suffer from degraded performance. However by summing phases around closed loops of non-redundant baselines, these random phasors cancel out and the resulting closure phases are extremely robust to wavefront aberrations. NRM interferometry from the ground (Tuthill et al. 2000) relies heavily on closure phase for high-contrast detection, and there are plans to extend the technique to space platforms (Sivaramakrishnan et al. 2009). Recent observations achieved with this technique reported by Lloyd et al. (2006), Martinache et al. (2007), Ireland et al. (2008), and Kraus et al. (2008) and Martinache et al. (2009) demonstrate that the level of calibration achieved with interferometric measurements permits the detection of companions at scales at or even somewhat beyond the diffraction limit of the imaging system. Recently, NRM interferometry succeeded in providing evidence for a low-luminosity companion in the transitional disk host systems T Cha (Huélamo et al. 2011) and LkCa 15 (Kraus & Ireland 2012).

It has recently been demonstrated that if a conventional (full-aperture) PSF is of sufficient quality (wavefront residual errors typically $\lesssim\lambda/4$), an analogous set of high-quality interferometric observables can be extracted from the images (Martinache 2010). These new quantities are the *kernel phases*, and represent a generalization of the idea of closure phase to a redundant pupil configuration. The major advance offered by kernel phase interferometry is that it is not restricted to non-redundant pupils. In brief, for small wavefront errors (i.e., high Strehl ratio), the phase errors in the pupil plane can be related to those in the Fourier plane by a linear operator. The kernel or null-space of this operator therefore singles out a subspace of baseline phases which are not affected by this error, which can then play the same role as closure phases in providing a robust set of observables to constrain image structure. Kernel phases were first successfully extracted from *HST*/NICMOS data on a single target by Martinache (2010), demonstrating significant improvement over more traditional data analysis (Pravdo et al. 2004). The technique has now also been successfully applied to ground-based AO observations (Martinache 2011).

This paper revisits a sample of nearby ultracool dwarfs observed by the *HST* NICMOS NIC1 camera and first presented in Reid et al. (2006a) and Reid et al. (2008). Our analysis allows dramatic extensions to the discovery space for putative companions, and in particular explores separation ranges down to 1 AU on targets located within 20 pc. Section 2 provides an overview of the data set and introduces the methods used for our new analysis. Section 3 discusses the results of the kernel phase analysis for the entire sample and implications for the astrophysical interpretation of brown dwarf formation.

2. SAMPLE AND METHODS

2.1. Sample of Ultracool Dwarfs

This study focuses on two samples of ultracool dwarfs, observed with the *HST*/NICMOS NIC1 camera, and whose properties were reported by Reid et al. (2006a) and Reid et al. (2008). Each target was observed in two filters: $F110W$ and $F170M$, which correspond loosely to the astronomical J and H bands. These differ in that the J and H bands sample atmospheric transmission windows, which do not constrain space-based observations. We will use J and H as shorthands for $F110W$ and $F170M$, respectively, but the difference should be noted. Table 1 summarizes the observational properties of the combined sample as stated in Table 1 of Reid et al. (2006a) and Table 1 of Reid et al. (2008).

In addition to detecting several binaries by traditional data analysis methods, these authors also provide a list of 43 and 26 apparently unresolved objects in the 2006 and 2008 samples, respectively, which we revisit in this paper. All 10 of the previously resolved binaries were independently recovered with kernel phases, and for all we report significantly improved astrometric precision. In addition to confirming the technique and software on an unambiguous sample, the dramatic improvements to the binary parameters offer the chance to determine orbital elements and therefore dynamical masses. Both the detections and the remaining unresolved binaries are used in quantitative exploration of the performance limitations of kernel phase analysis in the recovery of high-contrast systems.

2.2. Kernel Phase Analysis

Kernel phase analysis follows the principles introduced in Martinache (2010). The first step is to generate a model of the pupil of the imaging system as seen from the detector. This task is straightforward based on information contained in the TinyTim v. 7.2 PSF simulation package for NIC1 (Krist et al. 2011), available at `tinytim.stsci.edu`. For kernel phase analysis, the model pupil is discretized into a square grid array of 72 sub-apertures with a unit spacing $1/12$ th that of the pupil diameter (see Figure 1). Regions of the primary blocked by spiders or the secondary mirror are not sampled, and one can also observe that the unit baseline imposed by this sampling of the pupil imposes an outer working angle of $6\,\lambda/D$.

This geometry fills the (u, v)-plane with a regular grid of 176 distinct sample points at a cadence of 12 points across the diameter. The transfer matrix that relates instrumental phase errors to spurious (u, v)-phase information is therefore a 72×176 rectangular matrix, whose singular value decomposition (SVD) reveals 36 non-zero singular values (that is exactly one-half of the entire number of sample points in the pupil), leaving $176 - 36 = 140$ independent kernel phase relations. Kernel phase analysis can therefore recover $140/176 = 80\%$ of the available phase information present in the quantized grid.

The specific discretization chosen was found to have relatively little impact on the performance of the algorithm. If instead we adopt a finer-sampled pupil model with 20 points across the diameter, we get 1516 kernel phases out of 1632 distinct baselines, leading to a 93% phase recovery. We analyzed a portion of our data set with this finer 20 point sampling and found little improvement in the quality of fit or precision in parameter estimation, though we note somewhat better agreement between H and J bands with the finer pupil model using Levenberg–Marquardt model fitting. Because the finer grid analysis was computationally expensive but yielded only a small change in fitted binary parameters, it was judged that application over a large grid to fit the available data was unwarranted. We have therefore chosen the coarser model for our fitting routines, but note that more computer time may produce some improvements with a finer pupil model. For application to wider separation binaries, however, the finer model would be strictly required: if we have p points across the pupil, Nyquist's sampling theorem imposes an outer working angle $p\lambda/2D$. If this condition is not met, the Fourier plane fringes will not be well sampled and parameter estimates will be subject to aliasing or may not be recovered at all.

Table 1
Sample of Unresolved L dwarfs from Reid et al. (2006a) and Reid et al. (2008) (After Line Break)

2 MASS Name	Sp. Type	*J*	*H*	*K*
2MASS J00361617+1821104	L3.5	12.47	11.59	11.06
2MASS J00452143+1634446	L0	13.06	12.06	11.37
2MASS J01075242+0041563	L8	15.82	14.51	13.71
2MASS J01235905−4240073	M8	13.15	12.47	12.04
2MASS J01550354+0950003	L5	14.82	13.76	13.14
2MASS J02132880+4444453	L1.5	13.51	12.77	12.24
2MASS J03140344+1603056	L0	12.53	11.82	11.24
2MASS J03552337+1133437	L6	14.05	12.53	11.53
2MASS J04390101−2353083	L6.5	14.41	13.37	12.81
2MASS J04455387−3048204	L2	13.41	12.57	11.98
2MASS J05002100+0330501	L4	13.67	12.68	12.06
2MASS J05233822−1403022	L2.5	13.12	12.22	11.63
2MASS J06244595−4521548	L5	14.48	13.34	12.60
2MASS J06523073+4710348	L4.5	13.55	12.37	11.69
2MASS J08251968+2115521	L7.5	15.12	13.79	13.05
2MASS J08354256−0819237	L5	13.15	11.95	11.16
2MASS J08472872−1532372	L2	13.52	12.63	12.05
2MASS J09083803+5032088	L7	14.56	13.47	12.92
2MASS J09111297+7401081	L0	12.92	12.20	11.75
2MASS J09211410−2104446	L2	12.78	12.15	11.69
2MASS J10452400−0149576	L1	13.13	12.37	11.81
2MASS J10484281+0111580	L1	12.92	12.14	11.62
2MASS J10511900+5613086	L2	13.24	12.42	11.90
2MASS J11040127+1959217	L4	14.46	13.48	12.98
2MASS J11083081+6830169	L0.5	13.14	12.23	11.60
2MASS J12130336−0432437	L5	14.67	13.68	13.00
2MASS J12212770+0257198	L0	13.17	12.41	11.95
2MASS J14283132+5923354	L5	14.78	13.88	13.27
2MASS J14482563+1031590	L5	14.56	13.43	12.68
2MASS J15074769−1627386	L5	12.82	11.90	11.30
2MASS J15394189−05200428	L3.5	13.92	13.06	12.58
2MASS J15525906+2948485	L1	13.48	12.61	12.03
2MASS J16580380+7027015	L1	13.31	12.54	11.92
2MASS J17054834−0516462	L0.5	13.31	12.54	12.03
2MASS J17312974+2721233	L0	12.09	11.39	10.91
2MASS J17534518−6559559	L4	14.10	13.11	12.42
2MASS J18071593+5015316	L1.5	12.96	12.15	11.61
2MASS J19360262−5502367	L4	14.49	13.63	13.05
2MASS J20575409−0252302	L1.5	13.12	12.27	11.75
2MASS J21041491−1037369	L2.5	13.84	12.96	12.36
2MASS J22244381−0158521	L4.5	14.05	12.80	12.01
2MASS J23254530+4251488	L7.0	15.51	14.46	13.81
2MASS J23515044−2537367	L0.5	12.46	11.73	11.29
2MASS J002424.6−015819	M9.5	11.86	11.12	10.58
2MASS J010921.7+294925	M9.5	12.91	12.16	11.68
2MASS J022842.4+163933	L0	13.17	12.33	11.82
2MASS J025114.8−035245	L3	13.08	12.26	11.65
2MASS J025503.5−470050	L8	13.23	12.19	11.53
2MASS J031854.0−342129	L7	15.53	14.31	13.48
2MASS J044337.6+000205	M9	12.52	11.80	11.17
2MASS J083008.3+482848	L8	15.44	14.34	13.68
2MASS J085925.4−194926	L7	15.51	14.44	13.73
2MASS J102248.2+582545	L1	13.50	12.64	12.16
2MASS J102552.3+321235	L7	15.91:	15.59:	15.07
2MASS J104307.5+222523	L8	15.95	14.75	13.99
2MASS J105847.8−154817	L3	14.18	13.24	12.51
2MASS J115539.5−372735	L2	12.81	12.04	11.46
2MASS J120358.1+001550	L4	14.01	13.06	12.48
2MASS J130042.5+191235	L1	12.71	12.07	11.61
2MASS J142131.5+182741	L0	13.23	12.43	11.94
2MASS J142528.0−365023	L3	13.75	12.58	14.49
2MASS J143928.4+192915	L1	12.76	12.04	11.55
2MASS J150654.4+132106	L3	13.41	12.41	11.75
2MASS J151500.9+484739	L6	14.06	13.07	12.57
2MASS J172103.9+334415	L3	13.58	12.92	12.47
2MASS J200250.7−052152	L6	15.32	14.23	13.36
2MASS J202820.4+005227	L3	14.30	12.38	12.79
2MASS J214816.3+400359	L6.5	14.15	12.78	11.77
2MASS J223732.5+392239	M9.5	13.35	12.68	12.15

A shorter wavelength of observation delivers an increase of angular resolution, but with the same level of optical aberration, this also precipitates a greater degree of image degradation (lower Strehl ratio). When considering residual phase noise, we therefore expect that the kernel phase signal to noise will be accordingly higher for images taken at longer wavelengths.

We are furthermore limited by the fact that we only have single snapshots of each target: without multiple frames it is difficult to calibrate systematic errors and explore statistical uncertainties on the kernel phase observables. We therefore selected a sample of stars for which we could see no PSF abnormality or obvious Fourier phase structure, and repeatedly applied a Levenberg–Marquardt fitting algorithm to the raw kernel phase data to attempt to find binary companions. Those targets for which no companion model was significantly preferred over a single source were deemed to be "unresolved." We then used this unresolved population to establish uncertainties as ensemble standard deviations for each kernel phase, which in turn enables quantification of significance in subsequent explorations entailing χ^2 fitting. The results presented here could be considerably improved with the design of an observational campaign at the outset which delivers better diversity, by exploiting multiple exposures and dedicated point source calibrators. A more comprehensive understanding of systematic errors and noise estimates for individual targets would yield more sensitive limits on detection and better errors on fitted parameters.

2.3. Bayesian Methods

A binary system at any one epoch can be characterized by its angular separation δ, position angle θ, and contrast ratio c. The likelihood of a binary model with these parameters given the set of kernel phases $\{K\phi_j\}$ is related to the χ^2 statistic by

$$L(\delta, \theta, c|\{K\phi_j\}) \propto \exp(-\chi^2/2). \quad (1)$$

When normalized, this likelihood is the joint density probability function for all three parameters. When calculating χ^2, we found it necessary to add an additional systematic error term in quadrature to bring the minimum reduced χ^2 down to 1. Confidence intervals for any individual parameter can be calculated by integrating over the two other parameters. After this *marginalization*, we estimate the parameter and its uncertainty from the mean and the standard deviation of the one-dimensional marginal distribution, respectively.

The approach closely follows established practice with closure phase in NRM interferometry for the characterization of binaries (Martinache et al. 2009). When applied to our sample *HST* imaging data set, the final results from our algorithm were (1) confirmation of binaries already identified with other methods, (2) the determination of statistically sound constraints on the binary parameters, and (3) a robust statistical estimate for the probability that signals extracted from any given system betray the presence of a companion or can be attributed to noise.

The sampling and integration of the likelihood function given in Equation (1) is in general difficult, and is typically performed by a computationally expensive grid integration or a Monte Carlo Markov Chain (MCMC) random sampling method. In this paper we apply a recently developed alternative, namely nested sampling. This method, proposed by Skilling (2004), uses an unusual change of variables to calculate the model evidence. It has recently seen a surge of interest, e.g., for cosmological model fitting (Mukherjee et al. 2006; Mukherjee & Parkinson 2008) and the analysis of simulated gravitational wave data (Aylott et al. 2009; Feroz et al. 2009).

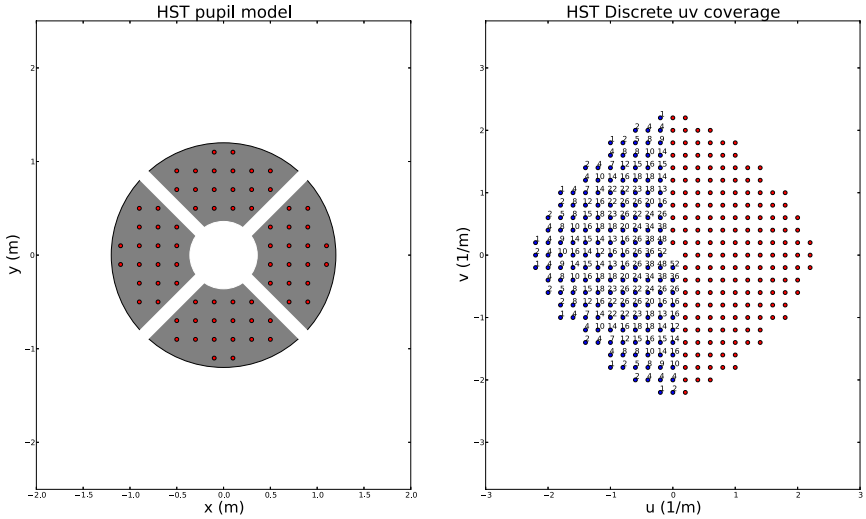


Figure 1. Diagrams showing (left) the discretized pupil model used for kernel phase analysis and (right) the resultant (u, v) sampling points.
(A color version of this figure is available in the online journal.)

The key idea of nested sampling is to populate the allowed prior space with a large number (∼100) of "active points" which are initially chosen at random and subsequently evolved toward ensemble states of successively higher likelihood using MCMC methods. Our implementation was based on Sivia & Skilling (2006) and ultimately yielded a statistical representation of the likelihood space which could be used for binary hypothesis testing and estimation of model parameter values and their uncertainties. Although a number of alternate gradient-descent and MCMC methods were benchmarked, nested sampling was found to be computationally the most efficient. A global binarity analysis of the entire sample in both the J and H bands could be accomplished quickly; however, for objects which are in the barely resolved limit, there are well-known strong parameter degeneracies—particularly between separation and brightness of a companion. This ambiguity conflates bright close companion models with somewhat more distant fainter companion models, considerably diminishing the astrophysical utility of the findings.

In addition to separately fitting image data in J and H bands, nested sampling was fast enough to enable joint four-parameter fitting of both images simultaneously $\mathcal{M}(\delta, \theta, c_H, c_J)$. The ambiguity in separation/contrast from separate fitting was found to be greatly ameliorated by the covariance of separation with contrast. Joint fitting enforces identical separation between bands, greatly restricting the size of the χ^2 valley of degeneracy with contrast ratio. These findings are promising for the coming generation of Integral Field Unit cameras working with extreme AO systems which naturally deliver spatio-spectral data cube observations.

For the joint fitting, an additional error term was added in quadrature to represent unknown noise sources. This was found iteratively such that each band separately had a minimum reduced χ^2 of 1 at the best joint fit parameters. In cases where the existing error estimates resulted in a minimum reduced $\chi^2 < 1$, no adjustment was made.

For this study, we searched a delimited parameter space for companions. We initially searched up to a contrast ratio of 200, somewhat beyond the limits established in Section 3.4.

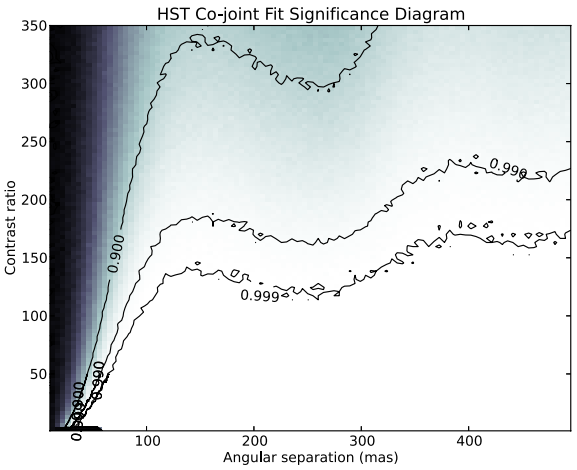


Figure 2. *HST* snapshot contrast-detection limits for co-joint fitting as a function of separation, averaged over position angle. The 90%, 99%, and 99.9% contours are overplotted with labels. Paler regions indicate higher significance.
(A color version of this figure is available in the online journal.)

Candidate binaries were then compared to Plots 2 and 3 to establish significance. The range of separation explored ran from 30 to 200 mas and all position angles were considered.

3. RESULTS AND DISCUSSION

Images for all objects in Table 1 in the two filter bands ($F110W$ and $F170M$) were recovered in digital form from the *HST* MAST Archive, where they are listed under Proposals 10143 and 10879. All data were processed using our kernel phase techniques, and in the discussion which follows, we divide our results into four subsections: binaries already reported (Section 3.1), new binaries uncovered by kernel phase (Section 3.2), marginal detections meriting further study (Section 3.3) and sample detection thresholds and the incidence of unresolved sources (Section 3.4).

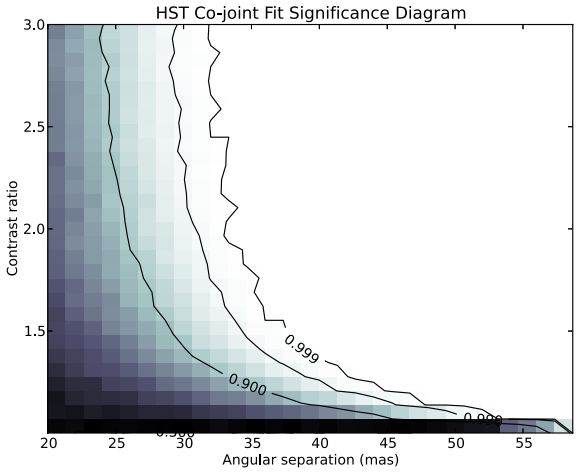


Figure 3. *HST* snapshot contrast-detection limits for co-joint fitting as a function of separation, averaged over position angle. The 90%, 99%, and 99.9% contours are overplotted labels. Paler regions indicate higher significance. This figure represents a simulation of the region near the origin of Figure 2 with a finer sample grid. Note the turnaround at low contrast: kernel phase performs poorly at detecting very low contrast companions.

(A color version of this figure is available in the online journal.)

3.1. Known Binaries

In all cases where companions were reported by Reid et al. (2006a) or Reid et al. (2008), our new analysis independently recovered strong systematic signals confirming binarity. We therefore confirm all previously reported detections, and we stress that our analysis was blind in the sense that no prior knowledge was employed in our search.

In common with closure phases, non-zero excursions in the kernel phases encode information about asymmetric structures, although the abstracted nature of the observable makes for significant challenges in intuitive data presentation. One approach to present the way a binary signal is encoded upon the kernel phases, and the fingerprint match of this complex function to the actual recorded data, is to simply plot the best-fit model quantities against the observed data in a correlation diagram.

Figures 5 and 6 present such a diagram of the binary model fit against extracted kernel phase data for two illustrative cases: previously known binaries 2M 0147−4954 and 2M 0700+3157. The one-to-one correspondence line is overplotted, delineating the locus of perfect fit.

Kernel phase analysis has yielded greatly improved astrometric precision on most previously known binary systems. Whereas previous studies, relying on visual analysis and PSF subtraction, quoted separations to the nearest 10 mas and position angle to ∼1°, kernel phase delivers about one order of magnitude better precision. All fitted binary parameters agree, typically to within 1σ, when separate fits to J- and H-band kernel phases are computed. Table 2 gives final best-fit parameters from the simultaneous J/H four-parameter fit. In many cases, the best fits differ from previously published estimates, sometimes very substantially (although formal errors were not quoted in Reid et al. 2006a or Reid et al. 2008).

Counterintuitively, for some of the most readily apparent binaries (e.g., 2M 0004−4044, 0025+4759, 0915−0422, and 2152+0937), kernel phase methods proved problematic and could fail to converge to a good fit. This was particularly true for well-separated, low-contrast systems which are most easily discerned by simple inspection. Wide binaries induce phase curvature not well sampled by our pupil model, and for these cases, a direct fit to the squared visibilities was performed. Visibility data were calibrated by dividing by the ensemble means over the sample, with dimensionless errors of ∼0.05 added in quadrature. As with the kernel phases, statistical analysis was based on nested sampling and the results also yielded overwhelming improvements in astrometric precision.

Table 2
Model Parameters for Known L-dwarf Binaries

2 MASS Number	Sep. (mas)	Pos. Angle (deg)	Contrast Ratio (J)	Contrast Ratio (H)
0004−4044[a]	84.6 ± 0.2	224.6 ± 0.1	1.04 ± 0.01	1.02 ± 0.02
0025+4759[a]	329.0 ± 0.3	233.04 ± 0.06	1.32 ± 0.04	1.03 ± 0.04
0147−4954	139.8 ± 0.1	72.66 ± 0.05	2.37 ± 0.09	2.06 ± 0.06
0429−3123	525.2 ± 1.2	285.3 ± 0.2	3.51 ± 0.1	2.82 ± 0.06
0700+3157	179.6 ± 0.5	105.8 ± 0.1	4.52 ± 0.07	3.81 ± 0.03
0915−0422[a]	738.6 ± 0.15	26.89 ± 0.01	1.114 ± 0.002	1.264 ± 0.002
1707−0558	1009.5 ± 1.0	34.9 ± 0.05	10.6 ± 0.15	7.5 ± 0.2
2152+0937[a]	253.7 ± 0.09	94.5 ± 0.02	1.09 ± 0.03	1.15 ± 0.03
2252−1730[b]	125.9 ± 0.4	353.5 ± 0.1	2.46 ± 0.01	3.395 ± 0.03
2255−5713	178.6 ± 0.4	172.7 ± 0.1	5.05 ± 0.08	4.33 ± 0.02

Notes.
[a] Low contrast: fit with visibilities. See Section 3.1.
[b] This object is the subject of Reid et al. (2006b).

The PSF of the individual target 2M 0004−4044 was truncated at the edge of the image, presumably due to spacecraft mispointing, and therefore did not permit a useful kernel phase fit in H band. Nevertheless, a J-band kernel fit was found to agree well with the parameters published in Reid et al. (2006a), and likewise a joint visibility fit agreed well in both bands. The results of the visibility fit are quoted in Table 2. It is unclear, however, what effect the data edge truncation may have on our interferometric observables and therefore the values given in Table 2 are likely to be subject to an additional unknown error for this system.

Example correlation diagrams and a corresponding NICMOS image are shown in Figures 4 and 5. Note that while the PSFs of the primary and companion overlap and are difficult to visually distinguish, they permit a clear kernel phase fit with very precise parameter estimates.

3.2. Discovery of New Binary Candidates

Table 3 reports five firm binary candidates not detected in the original Reid et al. (2006a) or Reid et al. (2008) studies, but recovered at very high 99.9% confidence from our kernel phase analysis (with the exception of 2M 0045+1634 for which the detection confidence was only 99%). Correlation plots for all of these are displayed as Figures 7–11.

Reid et al. (2006a) noted three targets which exhibited broad PSFs; however, they went on to report these stars had "... no evidence for the presence of a secondary component, and the broader profiles are probably an instrumental effect." These were 2M 1507−1627 and 2M 1936−5502, with a PSF FWHM of 2.47 pixels (106 mas) and 2M 0036+1821 with an FWHM of 2.56 pixels (110 mas), as opposed to the FWHM of 2.3–2.4 pixels found through the rest of the unresolved sample. Our kernel phase analysis identifies two of these as binary candidates: 2M 1936−5502 and 2M 0036+1821. We note that for both of these, the kernel phase signal-to-noise ratio is only of

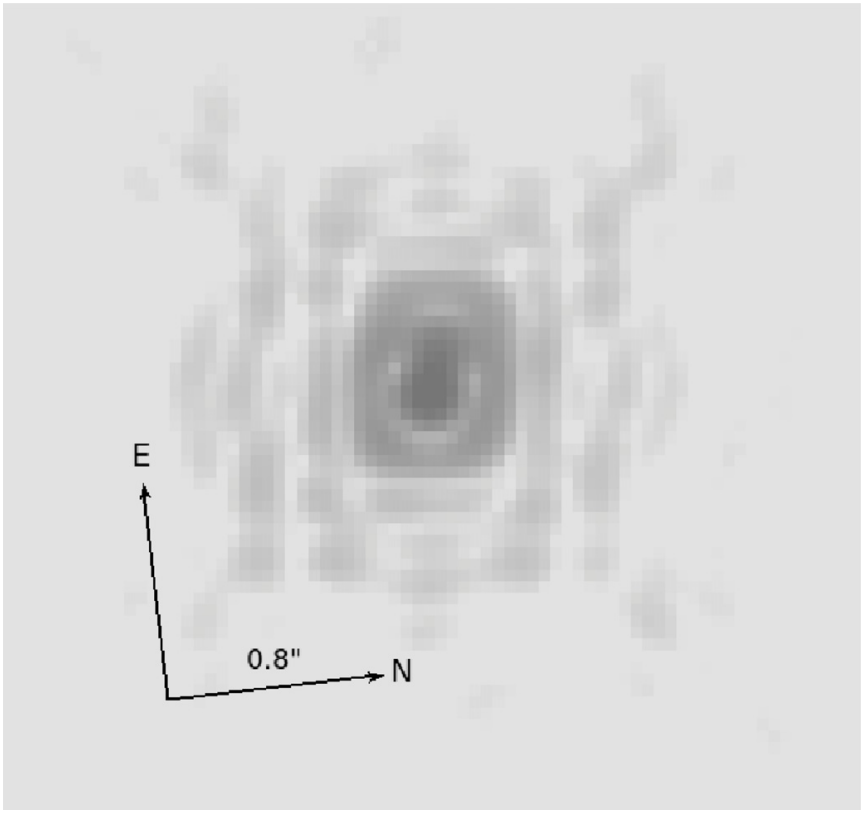


Figure 4. Log scale image of 2M 0147−4954 in *F*170*M* filter. The corresponding correlation diagrams are displayed in Figure 5. Note that while the PSFs are hard to distinguish visually, the kernel phase fit is excellent.

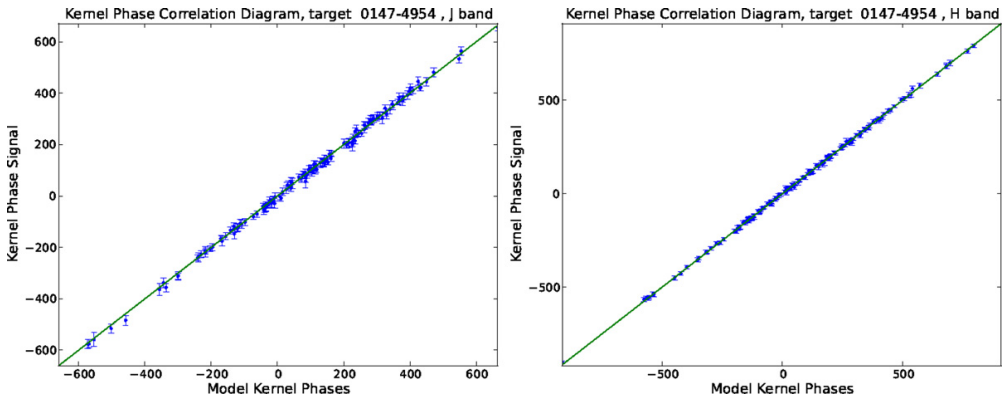


Figure 5. Correlation diagrams for the previously detected binary 2M 0147−4954 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

order two, and they exhibit a correspondingly noisy correlation plot.

2M 1936−5502 supported an alias fit at around 225° position angle, and the contrast in *H* band was very poorly constrained. This may well be considered the most marginal fit reported in this section. Nevertheless, both bands support overlapping position angle modes at 330°, and this object was considered for further study.

For 2M 0036+1821, there are two distinct χ^2 minima in *H* band, one of which overlaps precisely with the single distinct minimum from the *J*-band data. In assigning parameter estimates for this object, we have assumed some uncalibrated source of noise affected the *H* band and have therefore restricted the parameter space deliberately to contain only that region around the high-significance *J*-band χ^2 detection.

The third target with a reported broad PSF, 2M 1507−1627, shows marginally significant companion fits with parameters which are inconsistent between *J* and *H* bands. Until additional data can be recovered with higher signal to noise, we classify this object to be unresolved with no companion.

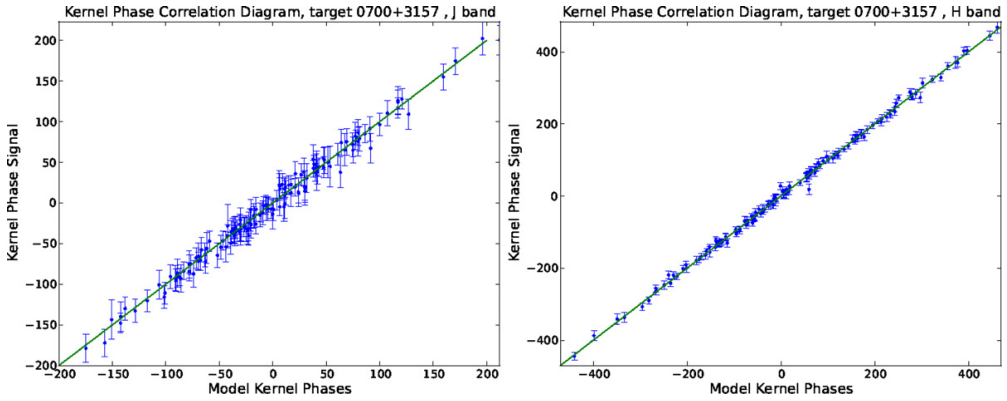


Figure 6. Correlation diagrams for the previously detected binary 2M 0700+3157 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

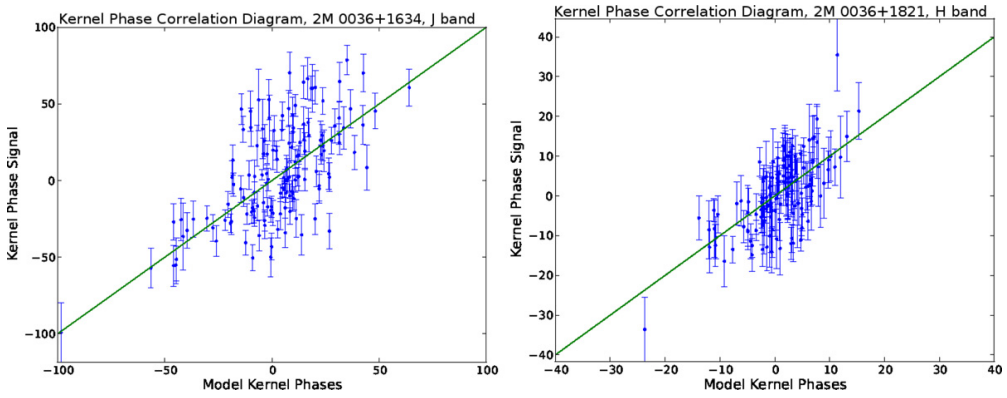


Figure 7. Kernel phase correlation diagrams for the newly confirmed binary 2M 0036+1821 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

Table 3
Model Parameters for New L-dwarf Binary Candidates

2 MASS Number	Spectral Type	Distance (pc)	Sep. (mas)	Pos. Angle (deg)	Contrast Ratio (*J*)	Contrast Ratio (*H*)	*J* (mag)	*H* (mag)
0036+1821[ab]	L3.5	8.77 ± 0.06[d]	44.5 ± 1.2	198.4 ± 1.3	1.85 ± 0.3	2.9 ± 1.1	12.47	11.59
..... A	L4						12.93	11.91
..... B	L5-6						13.6	13.06
0045+1634[c]	L0	26.8 ± 4.0	50.3 ± 0.7	300.6 ± 2.5	1.11 ± 0.02	1.12 ± 0.03	13.06	12.06
..... A	L0						13.75	12.75
..... B	L0						13.87	12.87
1936−5502[a]	L4	15.08 ± 1.2[e]	67.1 ± 6.4	330.9 ± 1.0	17.7 ± 3.9	35.5 ± 8.1	14.49	13.63
..... A	L4						14.54	13.66
..... B	(T-Y)						17.67	17.53
2028+0052	L3	26.1 ± 3.9	45.8 ± 1.2	107.7 ± 1.1	1.52 ± 0.1	3.1 ± 0.5	14.3	12.38
..... A	L3						14.85	12.7
..... B	L4					15.3	13.9	
2351−2537	L0.5	17.8 ± 2.7	63.3 ± 0.3	348.8 ± 0.3	2.4 ± 0.1	2.24 ± 0.2	12.46	11.73
..... A	L0						12.84	12.13
..... B	L1						13.79	13.006

Notes.
[a] Reid et al. (2006a) listed these objects as having a significant PSF abnormality. A third object showing such an abnormality, 2M 1507−1627, does not show a kernel phase binary fit.
[b] Bernat et al. (2010) detected this binary with Palomar AO aperture masking.
[c] This object was observed with spectral differential imaging by Stumpf et al. (2010). A PSF broadening was noticed but no other unambiguous signal was detected.
[d] This distance is known from trigonometric parallax to be 8.77 ± 0.06 pc (Dahn et al. 2002).
[e] Using the newer spectral class—absolute magnitude tables from M. Pecaut & E. Mamajek (in preparation), we revise the Reid spectroscopic distance from 15.4 pc to 18.6 ± 2.8 pc.

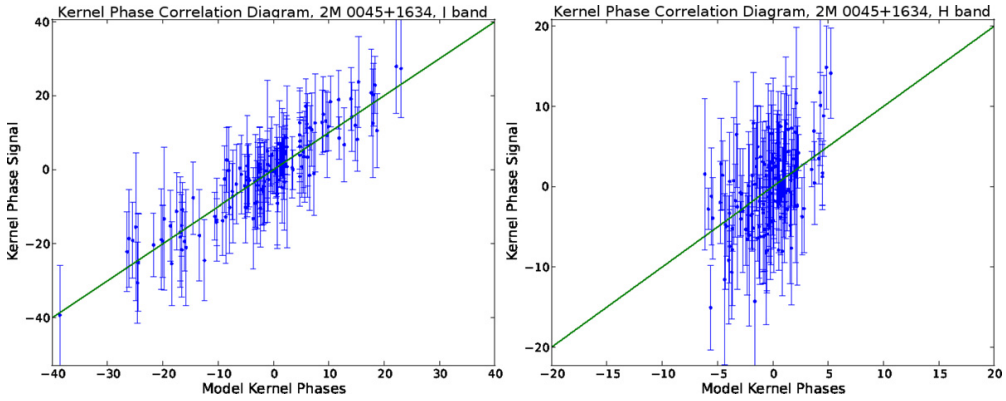


Figure 8. Kernel phase correlation diagrams for the newly detected binary 2M 0045+1634 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

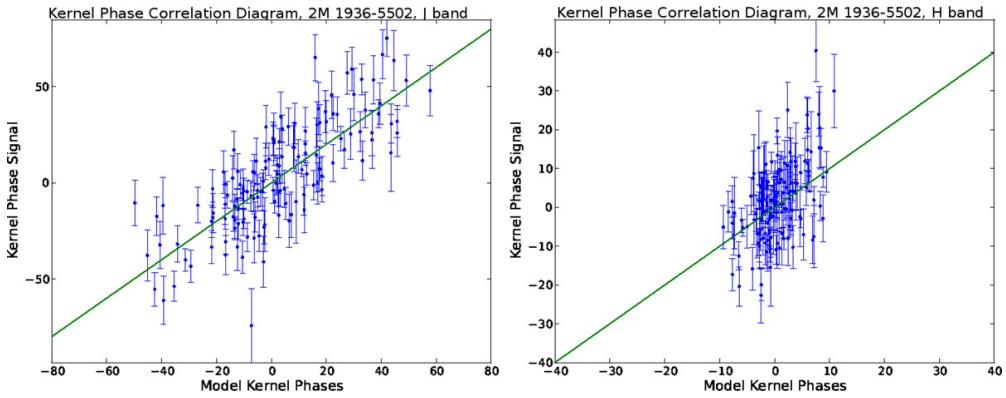


Figure 9. Kernel phase correlation diagrams for the newly detected binary 2M 1936−5502 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

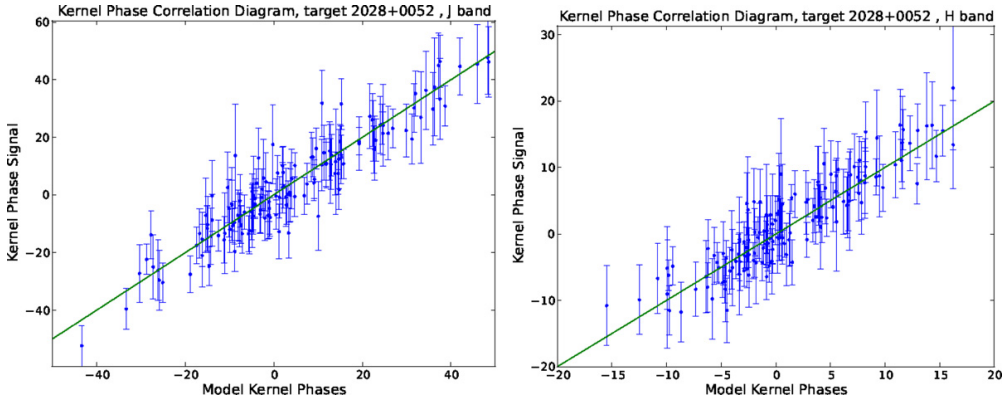


Figure 10. Correlation diagrams for the newly detected binary 2M 2028+0052 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

A companion to 2M 0036+1821 was also reported by Bernat et al. (2010) with NRM interferometry in *K* band with the LGS AO system on the Hale Telescope at Palomar Observatory. From data taken in 2008 September (some 3.5 years after the *HST* observations), these authors report a contrast ratio of 13.1 ± 3.1 and separation of 90 ± 11 mas at a position angle of $114^\circ \pm 5^\circ$. Although changes in the binary separation parameters are to be expected with progress in an orbit, the markedly different contrast ratio when compared to our fit in Table 3 is, at first glance, hard to reconcile. However, we note that both studies employ similar fundamental methodologies and are affected by the strong separation–contrast ratio degeneracy previously discussed. In particular, the work of Bernat et al. (2010) observed only in the K_s band and therefore enjoyed none of the advantages

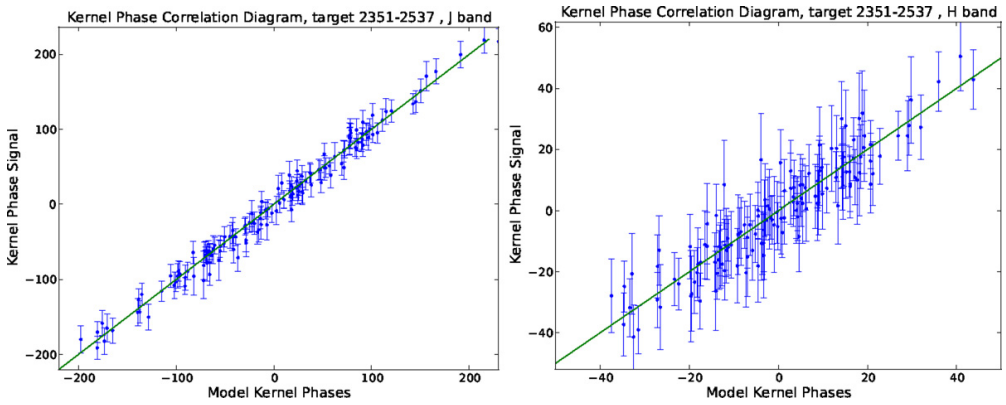


Figure 11. Correlation diagrams for the newly detected binary 2M 2351−2537 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

offered by dual wavelengths in lifting the ambiguity discussed in Section 2.3. Indeed, in their discussion an alternate detection at 25:1 contrast and 243 mas with equal probability is debated and ruled out based on the *HST* archival data. To test the hypothesis that both the Bernat et al. (2010) companion and the one reported here are consistent with the same degree of phase asymmetry, we re-fit our kernel phase data with an enforced higher contrast ratio above 10:1. This immediately resulted in much larger best-fit separations, approaching those from Bernat et al. (2010), and we therefore conclude that the two studies have probably identified the same companion.

All of these systems were originally assigned spectral classes in Reid et al. (2006a), which must be altered to reflect the discovery of new companions. A similar approach is taken here, using tabulated $J-H$ colors (M. Pecaut & E. Mamajek 2013, in preparation) to determine new spectral classes, taking into account the contrast ratio found at each wavelength. The class assigned to the secondary from its color is compared to the expected contrast with the primary as a function of spectral class, and found in each case to be relatively consistent. These classes, being based on contrasts which are themselves subject to error, are accurate only to within one division. Of special interest are 2M 2351−2537 and 2M 1936−5502. In the former case, Reid et al. (2008) reassigned the spectral type from L0.5 determined in Reid et al. (2006a) to M8, and excluded the system from their 20 pc catalogue of L dwarfs on these grounds. On the other hand, Andrei et al. (2011) classify it as L0.0 based on precision photometry. Revealing it as a binary system, its colors and luminosity imply that both primary and secondary must be early L dwarfs, with possible classes L0 and L1.

Given the uncertainty in contrast in the 2M 1936−5502 system, no accurate spectral type can be determined for the secondary component. If the high-contrast fit is confirmed, then the secondary must have a spectral type later than T9. A brown dwarf of Y class is possible in principle, but coeval companionship to an early-L primary seems unlikely if the primary is a brown dwarf. If this is the case, then the companion may be of planetary mass. On the other hand, if the primary is a star, then this could be an evolved companion of brown dwarf mass as discussed in Kirkpatrick et al. (2012). We note that Faherty et al. (2012) present a new parallax distance for this object of 15.08 ± 1.2 pc, which compares well with the spectroscopic distance estimate in Reid et al. (2006a) and comparatively poorly with new estimates we calculate with the revised tables of M. Pecaut & E. Mamajek (in preparation), from which we obtain a distance of 18.6 ± 2.8 pc, though we note that the measurement is still consistent with the quoted uncertainties.

For 2M 0036+1821, the distance is known from trigonometric parallax to be 8.77 ± 0.06 pc (Dahn et al. 2002). The angular separation of 44.5 mas therefore equates to 0.4 AU projected physical separation. This is therefore one of the closest projected separations of any resolved brown dwarf binary. The 90 mas separation from Bernat et al. (2010) gives typical binary orbital periods of ∼2–3 years, while on the other hand our 44.5 mas separation gives orbits of $\lesssim$1 year. Ignoring the changes in separation (due to degeneracy errors), the position angle change of 84° observed between the *HST* and Palomar data sets could be consistent with more than one orbit in the former case or several orbits in the latter: either is consistent with the data and a more rapid observing cadence is required to unambiguously follow the orbit in this particular system.

Excepting 2M 0036+1821 and 2M 1936−5502, the distances of other objects are known from spectroscopic parallax and are therefore subject to change depending on whether the system is found to be a binary. In general, discovery of a companion implies a higher total luminosity for the system, and therefore a greater distance. Table 3 lists recalculated spectroscopic distances, with standard 15% uncertainties. In particular, the reassignment of 2M 2351−2537 as an L0/L1 binary moves it back into the 20 pc L-dwarf sample, while 2M 0045+1634 and 2M 2028+0052 fall beyond 20 pc. The spectroscopic distance of 9.2 ± 1.4 pc for 2M 0036+1821 agrees well with the spectroscopic parallax of 8.77 pc.

Our kernel phase re-analysis of the *HST* archive does not detect a binary reported by Bernat et al. (2010) in 2M 0355+1133 at 90 mas separation. One explanation may be that at the *HST* epoch the projected separation was smaller and therefore unresolved. This is nevertheless included in the discussion of binarity fraction in Section 4.3. The object 2M 0045+1634 was observed by Stumpf et al. (2010) with differential spectral imaging using the *HST* NICMOS camera; a PSF variation was noted but no other signal found in the vicinity of the brown dwarf.

Thus four of the five detections reported in Table 3 can be considered new, and we confirm the previously reported detection in 2M 0036+1821 albeit with discrepant best-fit contrast. All detections are significantly below the formal diffraction limit and at separations of the order of the detector plate scale, which is

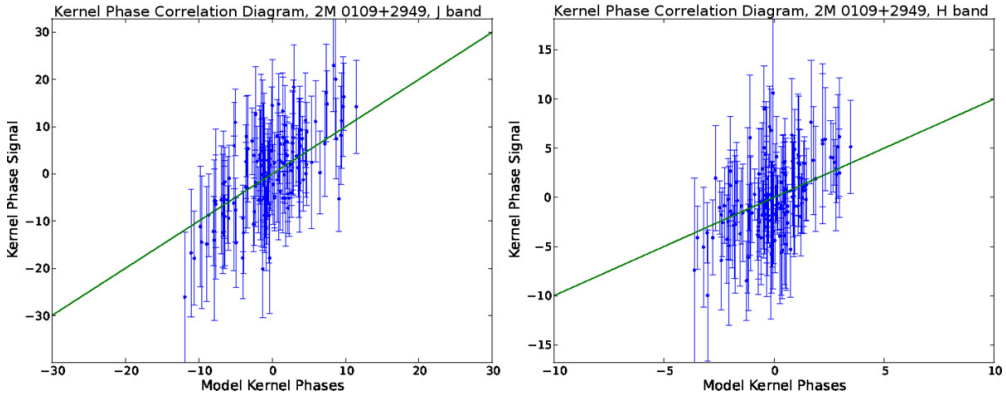


Figure 12. Kernel phase correlation diagrams for marginal detection 2M 0109+2949 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

43.1 mas pixel^{-1}. It is therefore unsurprising that the candidate binaries are subject to parameter correlation and substantial uncertainties. These may therefore require confirmation and accurate constraint of parameters found with follow-up observations employing a larger telescope.

3.3. Possible New Companions at Higher Contrasts

Estimates in Martinache (2011) suggested the possibility of detections at contrast ratios greater than 50:1 using kernel phases. The recent discovery of planetary-mass companions to brown dwarfs (Todorov et al. 2010) indicates that these objects exist and are at least common enough that in a sufficiently large data set candidates may be found. On the other hand, given the parameter degeneracy, it is possible that some high-contrast detections revealed by these techniques are in fact very close low-contrast binaries. Owing to this large uncertainty, no attempt has been made to assign separate magnitudes or spectral classes to the faint companions identified in this section.

A study of all the unresolved objects in our sample extending to a contrast of 100 revealed four objects with fits agreeing in position angle between bands: 2M 0314+1603, 2M 1539−0520, 2M 0830+4828, and 2M 0109+2949.

These objects were then subjected to a nested sampling joint fit to examine correlation at best fit. In the extreme contrast–separation regime probed here, the parameters show very pronounced degeneracy and it is impossible to reliably distinguish between high-contrast objects further out and lower contrast objects closer in, even using both *J* and *H* bands. No error was added in quadrature, as in each case the kernel phases were already overfit with reduced $\chi^2 < 1$.

2M 0314+1603 is the best of the candidates in Table 4 and exhibits excellent correlation in *J* band. Nevertheless, while *H*-band data favor a χ^2 minimum at this same position angle, it is unconstrained in contrast ratio and reliably runs off to high contrast in each attempted fit. Likewise 2M 1539−0520 supported good correlation diagrams but a surprisingly large discrepancy in contrast ratio between bands. This is likely to be the result of a strongly degenerate fit which constrains neither contrast well; the contrast ratios should therefore not be considered well determined.

With a *J* contrast >160, as a coeval companion to a brown dwarf this can only be a planetary mass object. On the other hand, it is possible that the L primary may be a star at the low-mass limit of the main sequence and its companion is itself an evolved brown dwarf. Such a system is still of considerable interest: as noted in Dieterich et al. (2012), low-mass binary systems tend to be of equal mass, increasingly so at lower primary masses. Accordingly, few brown dwarf secondaries to main-sequence primaries are known, especially at close separations. This is the low-mass end of the "brown dwarf desert" first identified by Marcy & Butler (2000), which has since become the subject of intense study (Grether & Lineweaver 2006; Allen et al. 2007; Kraus et al. 2008; Deleuil et al. 2008; Kraus et al. 2011; Dieterich et al. 2012; Evans et al. 2012). A confirmation of the status of this system, regardless of result, would therefore be a potentially significant finding.

Correlation plots for the marginal detections are shown in Figures 12–15.

3.4. Survey Confidence and Detection Limits

In order to quantify our survey detection threshold as a function of model parameters, we performed a Monte Carlo study simulating detection of a population of model binaries. We ran 100 simulations, adding a binary at each point on a grid in separation, position angle, and contrast ratio, in both *J* and *H* bands, as a co-joint fit. We then added Gaussian noise randomly to each of these, with the distribution given by our measured error distribution as discussed in Section 2.2. The detection rates as a function of separation and contrast are averaged over position angle (distributions were found to be azimuthally symmetric). Quantitative detection thresholds were formulated from comparison of the binary model χ^2 to the null hypothesis. This method is similar to that used in Martinache (2010). We also examined the low-contrast, low separation domain in which we have found many of our binaries with a finer sample grid. The contrast-detection limit curves turn around at very low contrasts, as kernel phase performs poorly in discerning very low contrast companions, which give rise to highly symmetrical images with a weak Fourier phase signal. This is true of any phase-based method, as the Fourier phases encode information about spatial asymmetries in a source image, and this problem therefore affects closure phases in NRM interferometry as well. This therefore places a lower limit on the contrasts detectable at very small separations. The contrast thresholds obtained from these studies are shown in Figures 2 and 3, respectively. Our estimates here compare favorably with the kernel phase contrast-detection limits predicted for the Keck Telescopes in Martinache (2011), which are at best closer to 50:1. This very high sensitivity

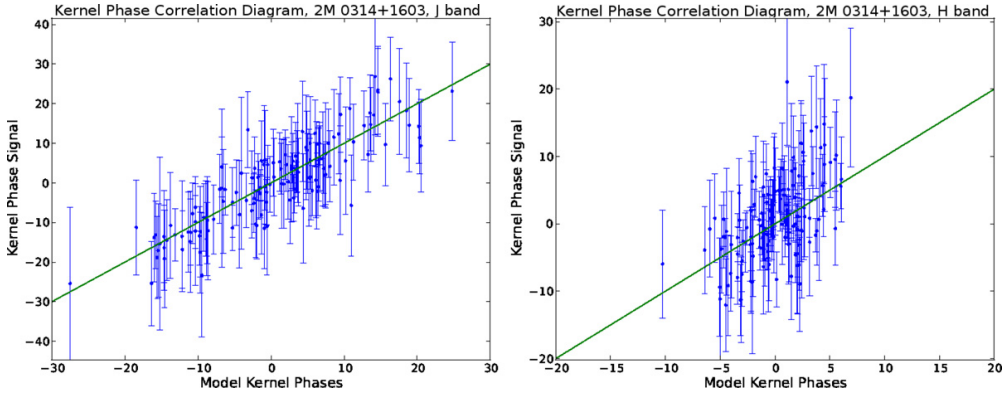


Figure 13. Kernel phase correlation diagrams for marginal detection 2M 0314+1603 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

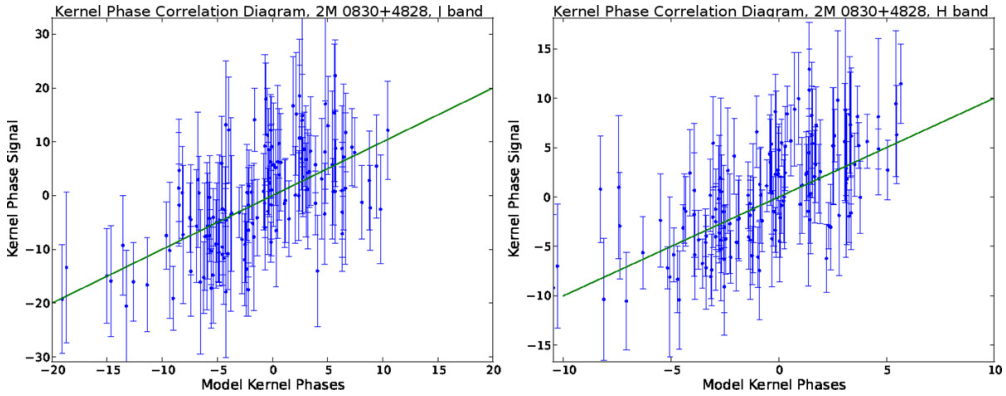


Figure 14. Kernel phase correlation diagrams for marginal detection 2M 0830+4828 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

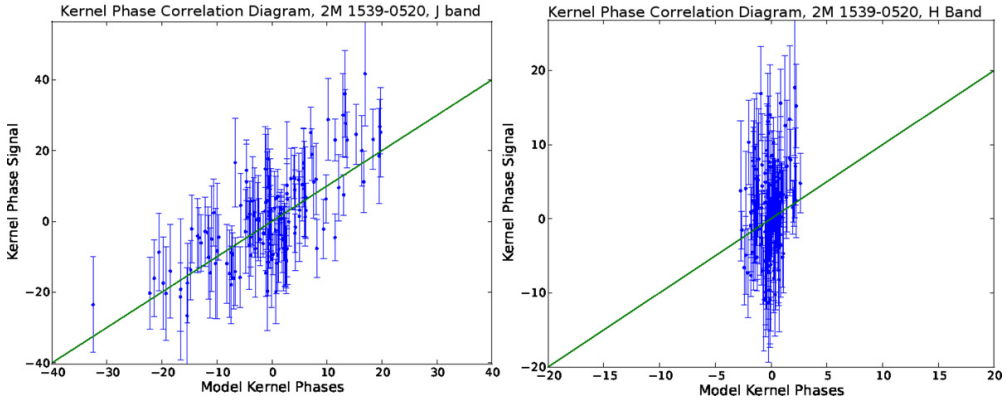


Figure 15. Kernel phase correlation diagrams for marginal detection 2M 1539−0520 in *J* band (left) and *H* band (right).
(A color version of this figure is available in the online journal.)

with the *HST* is unsurprising, given the high quality of the wavefront, and this technique therefore holds great promise in its application to similar snapshot samples.

Over any companion detection hangs the question as to whether the pair of stars imaged are physically associated, or whether they merely happen to lie along the same line of sight but are otherwise unrelated. It is therefore important to establish the expected count of background stars in the direction of each star, which will in general differ according to the star's position, owing to shape of the Galaxy and the distribution of dust.

The *Galaxia* software package (Sharma et al. 2011) is a synthetic survey tool, which calculates the expected density of

Table 4
Model Parameters for Marginal High-contrast Companions

2 MASS Number	Separation (mas)	Position Angle (deg)	Contrast Ratio (J)	Contrast Ratio (H)
0109+2949	49 ± 13	268 ± 3	40 ± 25	43 ± 29
0314+1603	124 ± 16	227 ± 2	70 ± 10	>160
0830+4828	48 ± 9	120 ± 3	29 ± 16	14 ± 11
1539−0520	35 ± 5	332 ± 5	(4 ± 3)	(28 ± 16)

Galactic stars at a given magnitude in a given band along a given line of sight. This was applied in 0.1 deg^2 regions around the coordinates of each binary candidate, searching for background stars between 12th and 18th apparent magnitude in J and H bands. For a canonical 12th magnitude primary, these span the contrast range from 1 to 100, and therefore cover both the brown dwarf and planetary-mass companion regimes. This was then scaled down to the expected counts in a 200 mas circle around the target.

In summary, no field in either band shows significantly more than 2% probability of finding a background star within 200 mas, and in the overwhelming majority of cases, including almost all new binary candidates and marginal detections, this figure was an order of magnitude lower. It is therefore exceedingly likely that background stars do not contribute in any way to the population of companion candidates in this survey.

4. DISCUSSION

4.1. Kernel Phase Performance

From our retrieval of existing and new binaries, we have demonstrated that kernel phase interferometry performs well in medium and wide band filters. For shorter wavelengths the PSF quality degrades increasing the errors; however, in all cases space telescope data are firmly in the regime where wavefront quality is excellent and easily sufficient for the purposes of the algorithm. This method has demonstrated the delivery of very precise astrometry for medium-separation systems. If employed over multiple epochs, this will permit correspondingly accurate dynamical mass measurements for most detected binaries.

For the very closest companions, we note that the signal-to-noise ratio is lower and accordingly kernel phase yields weaker constraints on the binary parameters, which nevertheless should still permit the determination of dynamical masses. For systems whose close separation puts them beyond the diffraction limit, errors are dominated by covariance between the separation and contrast creating model ambiguity between close, bright companions and distant, fainter companions. We have shown that when multi-wavelength observations are employed, this degeneracy can be partially lifted.

4.2. Opportunities for Dynamical Mass and Radius Measurement

Adopting distance and mass estimates from Reid et al. (2006a) and Reid et al. (2008), and our own parameter estimates, typical binary objects reported here such as 2M 2351−2537 and 2M 2028+0052 would have a binary separation of ∼1 AU and therefore orbital periods of 3–4 years. A follow-up campaign with LGS AO could be used to track the orbits of these binaries and dynamically determine their mass in the near term, as has already been achieved for GJ 802B in Ireland et al. (2008).

With the already observed epochs from Reid et al. (2006a, 2008) and Bernat et al. (2010), two more epochs should be sufficient to permit a first fit to the binary parameters of each system. Targets observed in 2006 or 2008 will have completed a substantial fraction of their orbit since they were first observed; on the other hand, over a two-month period they will have rotated by 6° in mean anomaly, which is substantially greater than the ∼ degree uncertainties in position angle obtained with kernel phase or aperture masking. These are pessimistic figures, in that the systems are early L dwarfs and include several systems substantially closer than 2 AU, which leads to correspondingly shorter orbital periods for targets of interest.

Table 5
Mass Estimates for New L-dwarf Binary Candidates

2 MASS Number	$M_{0.5\,\mathrm{Gyr}}$ ($M_\odot$)	$M_{1\,\mathrm{Gyr}}$ ($M_\odot$)	$M_{5\,\mathrm{Gyr}}$ ($M_\odot$)
0036+1821 A	0.049	0.064	0.073
..... B	0.045	0.060	0.071
0045+1634 A	0.07	0.078	0.081
..... B	0.07	0.078	0.081
1936−5502 A	0.049	0.064	0.073
..... B			
2028+0052 A	0.052	0.066	0.074
..... B	0.049	0.064	0.073
2351−2537 A	0.07	0.078	0.081
..... B	0.059	0.072	0.075

In Table 5, we present approximate masses computed using the methods of Reid et al. (2006a). These were calculated by taking the absolute J magnitude of the new spectral class as listed in M. Pecaut & E. Mamajek (in preparation), applying J magnitude bolometric corrections from Golimowski et al. (2004) and comparing bolometric magnitudes with the 0.5, 1 and 5 Gyr isochrones from Chabrier et al. (2000). The calculated values are subject to direct uncertainties of order of ∼0.03 $M_\odot$: despite the significant uncertainty in photometry, bolometric magnitude is a very strong function of mass and age and purely photometric uncertainties are small. On the other hand, the models used come with significant caveats, especially because brown dwarfs cool as they age: the mass estimates, given photometry, therefore depend on the assumed age of the system. Obtaining dynamical masses is key to calibrating these models brown dwarfs in general, as discussed in Section 1. These close new binaries therefore present a significant addition to the population of targets which can be followed up on short timescales.

In addition to this, many brown dwarfs are known to exhibit periodic radio or Hα emission, modulated by the body's rotation period, cataloged in detail in Antonova et al. (2013). The radio emission is believed to be from electron cyclotron maser instability in the brown dwarf magnetosphere (Hallinan et al. 2008; Kuznetsov et al. 2012). Knowing this rotation period and the projected rotational velocity $v \sin i$ from spectroscopic observations, it is possible then to determine the radius of the object if its spin is assumed to be perpendicular to the orbital plane of the system. This has been used by Berger et al. (2009) to determine the radius of a radio- and Hα-variable component of the brown dwarf binary system 2MASSW J0746425+200032. This new sample of binary systems therefore also presents the opportunity to systematically study the radii of any brown dwarfs found to have radio emission. Notably, 2M 0036+1821 has a rotation period of 3.08 ± 0.05 hr as determined in Hallinan et al. (2008), and therefore presents the clearest new opportunity for radius measurement in this data set.

4.3. Binarity Fraction in L-dwarf Sample

We now calculate the revised L-dwarf binary fraction in the 20 pc sample. Following Reid et al. (2008), correcting for Malmquist bias requires that for statistics in a 20 pc sample, we should consider only systems within 19 pc. The unbiased estimator is then simply the binary fraction observed. For uncertainties, ordinarily Poisson statistics would be used; in this case, however, the sample is too small ($N < 100$), and so binomial statistics are required (Burgasser et al. 2003).

Where Reid et al. (2008) had 8 binaries out of 64 systems observed in the nearest 20 pc, giving $\epsilon_b = 12.5^{+5.3}_{-3.0}\%$, five of the previously unresolved systems in the sample now support binary detections. Of these, the spectral class of 2M 2351−2537 is reclassified to L0/L1, while the increased spectroscopic distance to 2M 0045+1634 leaves it outside of 20 pc. Given that 2M 2028+0052 was already known to lie beyond 20 pc, three of these are admissible in the 20 pc sample. Furthermore, 2M 1936−5502 supports a high-contrast fit, indicating a potential sub-L-dwarf classification, so among the new discoveries, there are two firm detections: 2M 0036+1821 and 2M 2351−2537. In addition to these, we admit a detection of a companion to 2M 0355+1133 in Bernat et al. (2010), for a total of three new binaries. This yields a new binary fraction of $\epsilon_b = 17.2^{+5.7}_{-3.7}\%$. If 2M 1936−5502 is included, the binary fraction is then $18.75^{+5.8}_{-3.7}\%$.

4.4. Formation of Brown Dwarfs

A typical brown dwarf has a mass significantly below the Jeans mass of a collapsing protostellar gas cloud (Reid & Hawley 2005), and it is therefore difficult to explain the observed abundance of field brown dwarfs based on the standard Jeans collapse theory of star formation. This leads to two widely discussed hypotheses (Basu 2012). The first, namely gravoturbulent collapse, posits that turbulent gas dynamics allows for the collapse of smaller clouds than naive analysis suggests, and therefore brown dwarfs may form in essentially the same manner as other stars. The alternative is embryo ejection, whereby gravitational interactions in a protostellar system may eject low-mass companion embryos before they accrete enough gas to achieve fusion in their cores (Bate et al. 2002). Bate & Bonnell (2005) propose that protostellar cores start out at the opacity limit for turbulent fragmentation (of order a few M_J) and proceed to accrete gas until they are ejected dynamically from the cloud. This behavior differs particularly strongly from direct turbulent collapse in the substellar regime, so that brown dwarfs are thrown out very early in the accretion process. Therefore by testing model predictions for these ultracool stars we provide one of the most stringent tests of the general validity of the Bate & Bonnell (2005) and Bate (2012) models of star formation for all stellar types.

The chief objection to direct gravitational collapse models had previously been that the collapse of such a low-mass cloud would require it to start out very dense and very cold compared to well-studied star-forming regions (Reid & Hawley 2005). Recent millimeter interferometry by André et al. (2012) has caught just such a gravoturbulent collapse in the act, observing a gravitationally bound cloud of mass $\sim$0.02–0.03 $M_\odot$. In addition to this, Monin et al. (2013) report the detection of a molecular outflow from the brown dwarf binary FU Tau, which is the third young brown dwarf system found undergoing a formation process analogous to low-mass stars. With at least one example of a brown dwarf forming directly, models that do not account for this behavior must therefore be refined.

On the other hand, while embryo ejection can easily account for the formation of isolated field brown dwarfs, binaries are problematic. The binarity fraction of very low mass stars can discriminate between these two models (Burgasser et al. 2007). Embryo ejection predicts a low binarity fraction in field brown dwarfs, as the gravitational interaction required would disrupt all but the most tightly bound binaries. The predicted binarity fraction is thus <5% in the oldest models for embryo ejection (Bate et al. 2002), rising only to 8% ± 5% for systems with a primary mass in the range of 0.08–0.1 $M_\odot$ in the most recent models incorporating radiative feedback (Bate 2012). The observed binarity fraction of ultracool stars seems closer to 15% (Bouy et al. 2003; Reid et al. 2006a, 2008) which puts it into conflict with the models of Bate et al. (2002), Bate & Bonnell (2005), and Bate (2012), although the authors noted that this prediction is subject to significant uncertainty.

The high binarity fraction reported in this paper is still further evidence in favor of a higher binarity rate than predicted by embryo ejection. Moreover, with the additional detections this prediction now lies several σ away from the observed value. The hybrid ejection model of Basu & Vorobyov (2012), by which still-collapsing clouds can be ejected from a protostellar disk, makes few clear predictions regarding binarity, other than that there should be few large-separation binaries. Given the results of André et al. (2012) in observing a gravitationally collapsing brown dwarf mass object, it is not implausible to suggest that direct gravitational collapse may account for most or all field brown dwarfs.

The effect of radiative feedback in ameliorating the difficulties of modeling brown dwarf formation in Bate (2012) has been suggested by Dieterich et al. (2012) as an explanation for why the initial mass function and companion mass function change at the hydrogen burning limit; while this idea is highly speculative, it seems likely that the identification of systems with a late main-sequence primary and a brown dwarf secondary as reported in Section 3.3 will be of interest in testing this hypothesis.

5. CONCLUSIONS

We have shown that extraction and fitting to the self-calibrating kernel phase interferometric observables allow for a significant increase in the robust detection of companions and in the accuracy of best-fit parameters recovered when compared to simple inspection of images. The *HST*, or any instrument delivering good wavefront quality, may benefit dramatically from such an approach.

Using the kernel phase technique on all 79 ultracool dwarfs in the combined *HST* samples of Reid et al. (2006a) and Reid et al. (2008), we independently recover all 10 prior detections and improve on the precision of fitted parameters such as position angle and separation by a factor of ∼10. Furthermore, we report five new binary detections missed by the original authors, four of which are presented here for the first time, and additionally four marginal detections of close or high-contrast companions. This population forms an excellent base for dynamical studies to establish masses; a prospect particularly favored by the improved precision in parameter estimation. As well as these confident new detections, our kernel phase analysis identifies up to four more marginally resolved close or faint companions, which may be of planetary mass.

The finding of a larger binary fraction for this sample helps shed light on the formation mechanisms of very low mass

objects. The five additional ultracool binaries, if confirmed, lend further support to gravoturbulent collapse models for the formation of low-mass stars in the field.

We thank I. Neill Reid, Olivier Guyon, Paul Stewart, Barnaby Norris, Anthony Cheetham, and James Allison for their interest, helpful feedback, and consistent encouragement, as well as Sanjib Sharma for his assistance. We also thank the anonymous reviewer, for drawing attention to several items of recent work and whose comments helped us significantly improve this paper. Support for program number HST-AR-12849.01-A was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. This research has made use of NASA's Astrophysics Data System.

REFERENCES

Allen, P. R., Koerner, D. W., McElwain, M. W., Cruz, K. L., & Reid, I. N. 2007, AJ, 133, 971
André, P., Ward-Thompson, D., & Greaves, J. 2012, Sci, 337, 69
Andrei, A. H., Smart, R. L., Penna, J. L., et al. 2011, AJ, 141, 54
Antonova, A., Hallinan, G., Doyle, J. G., et al. 2013, A&A, 549, A131
Aylott, B., Veitch, J., & Vecchio, A. 2009, CQGra, 26, 114011
Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Natur, 320, 595
Basu, S. 2012, Sci, 337, 43
Basu, S., & Vorobyov, E. I. 2012, ApJ, 750, 30
Bate, M. R. 2012, MNRAS, 419, 3115
Bate, M. R., & Bonnell, I. A. 2005, MNRAS, 356, 1201
Bate, M. R., Bonnell, I. A., & Bromm, V. 2002, MNRAS, 332, L65
Berger, E., Rutledge, R. E., Phan-Bao, N., et al. 2009, ApJ, 695, 310
Bernat, D., Bouchez, A. H., Ireland, M., et al. 2010, ApJ, 715, 724
Bouy, H., Brandner, W., Martín, E. L., et al. 2003, AJ, 126, 1526
Burgasser, A. J., Kirkpatrick, J. D., Reid, I. N., et al. 2003, ApJ, 586, 512
Burgasser, A. J., Reid, I. N., Siegler, N., et al. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 427
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Dahn, C. C., Harris, H. C., Vrba, F. J., et al. 2002, AJ, 124, 1170
Deleuil, M., Deeg, H. J., Alonso, R., et al. 2008, A&A, 491, 889
Dieterich, S. B., Henry, T. J., Golimowski, D. A., Krist, J. E., & Tanner, A. M. 2012, AJ, 144, 64
Evans, T. M., Ireland, M. J., Kraus, A. L., et al. 2012, ApJ, 744, 120
Faherty, J. K., Burgasser, A. J., Walter, F. M., et al. 2012, ApJ, 752, 56
Feroz, F., Gair, J. R., Hobson, M. P., & Porter, E. K. 2009, CQGra, 26, 215003
Golimowski, D. A., Leggett, S. K., Marley, M. S., et al. 2004, AJ, 127, 3516
Grether, D., & Lineweaver, C. H. 2006, ApJ, 640, 1051
Hallinan, G., Antonova, A., Doyle, J. G., et al. 2008, ApJ, 684, 644
Huélamo, N., Lacour, S., Tuthill, P., et al. 2011, A&A, 528, L7
Ireland, M. J., Kraus, A., Martinache, F., Lloyd, J. P., & Tuthill, P. G. 2008, ApJ, 678, 463
Jennison, R. C. 1958, MNRAS, 118, 276
Kirkpatrick, J. D., Gelino, C. R., Cushing, M. C., et al. 2012, ApJ, 753, 156
Kraus, A. L., & Ireland, M. J. 2012, ApJ, 745, 5
Kraus, A. L., Ireland, M. J., Martinache, F., & Hillenbrand, L. A. 2011, ApJ, 731, 8
Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762
Krist, J. E., Golimowski, D. A., Schroeder, D. J., & Henry, T. J. 1998, PASP, 110, 1046
Krist, J. E., Hook, R. N., & Stoehr, F. 2011, Proc. SPIE, 8127, 81270J
Kuznetsov, A. A., Doyle, J. G., Yu, S., et al. 2012, ApJ, 746, 99
Lloyd, J. P., Martinache, F., Ireland, M. J., et al. 2006, ApJL, 650, L131
Marcy, G. W., & Butler, R. P. 2000, PASP, 112, 137
Martinache, F. 2010, ApJ, 724, 464
Martinache, F. 2011, Proc. SPIE, 8151, 815111
Martinache, F., Lloyd, J. P., Ireland, M. J., Yamada, R. S., & Tuthill, P. G. 2007, ApJ, 661, 496
Martinache, F., Rojas-Ayala, B., Ireland, M. J., Lloyd, J. P., & Tuthill, P. G. 2009, ApJ, 695, 1183
Monin, J.-L., Whelan, E., Lefloch, B., Dougados, C., & Alves de Oliveira, C. 2013, A&A, 551, L1
Mukherjee, P., & Parkinson, D. 2008, IJMPA, 23, 787
Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, ApJL, 638, L51
Pravdo, S. H., Shaklan, S. B., Henry, T., & Benedict, G. F. 2004, ApJ, 617, 1323
Reid, I. N., Cruz, K. L., Burgasser, A. J., & Liu, M. C. 2008, AJ, 135, 580
Reid, I. N., & Hawley, S. L. (ed.) 2005, New Light on Dark Stars Red Dwarfs, Low-mass Stars, Brown Dwarfs (Heidelberg: Springer)
Reid, I. N., Lewitus, E., Allen, P. R., Cruz, K. L., & Burgasser, A. J. 2006a, AJ, 132, 891
Reid, I. N., Lewitus, E., Burgasser, A. J., & Cruz, K. L. 2006b, ApJ, 639, 1114
Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3
Sivaramakrishnan, A., Tuthill, P. G., Ireland, M. J., et al. 2009, Proc. SPIE, 7440, 74400Y
Sivia, D., & Skilling, J. 2006, Data Analysis: A Bayesian Tutorial (Oxford: Oxford Univ. Press)
Skilling, J. 2004, in AIP Conf. Proc. 735, ed. R. Fischer, R. Preuss, & U. V. Toussaint (Melville, NY: AIP), 395
Stumpf, M. B., Brandner, W., Joergens, V., et al. 2010, ApJ, 724, 1
Todorov, K., Luhman, K. L., & McLeod, K. K. 2010, ApJL, 714, L84
Tuthill, P., Lloyd, J., Ireland, M., et al. 2006, Proc. SPIE, 6272, 62723A
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555

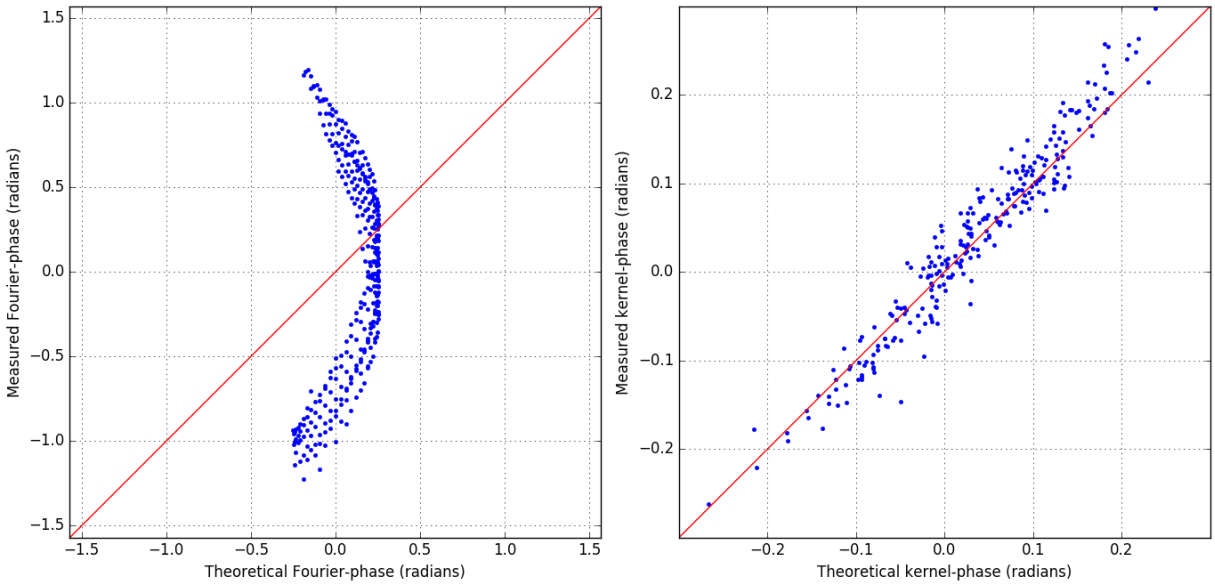


FIGURE 64: Autre représentation de la comparaison entre phase de Fourier brute et kernels de phase, cette fois ci, sous forme de diagramme de corrélation.

Limites de l'approche

Bien que l'erreur de reconstruction ne soit pas optimale, le gain de filtrage de la phase apporté par les kernels est appréciable. Etant donné la longueur d'onde d'observation, on peut traduire l'erreur RMS sur la phase de 0.03 radians en son équivalent en front d'onde, et voir qu'on passe d'une collection d'observables bruts (la phase dans le plan de Fourier) inutilisables à une collection d'observables justes avec un résidu de l'ordre de 7 nanomètres[77] seulement.

77. On utilise pour ça la simple relation $\delta = \phi * \lambda / 2\pi$

Au delà de l'observation du gain apporté par les kernels, on peut s'interroger sur les limites de cette approche et sur ce qui fait que la déviation standard reportée pour la simulation idéale qui vient d'être présentée, est de l'ordre de 0.03 radians et pas un ou plusieurs ordres de grandeur plus bas. Plusieurs tests sont possibles pour tenter de répondre à cette question.

On peut commencer par vérifier qu'en l'absence d'aberration, la phase brute extraite dans le plan de Fourier utilisant l'algorithme LDFT correspond effectivement à cette phase théorique attendue. Dans les conditions de la simulation mise en oeuvre, l'écart moyen entre la phase théorique et celle extraite dans des conditions idéales, pour une image de 256 pixels de côté, est inférieur à 4×10^{-6} radians, avec une incertitude de l'ordre de 10^{-4} radians. On remarque que cette incertitude est en bon accord avec le résidu d'extraction par la LDFT présenté dans le panneau du bas de la figure 60.

Cette erreur, multipliée par la matrice $\mathbf{K}$ donne un biais plancher de l'ordre de 10^{-7} radians pour les kernels, avec une incertitude également de l'ordre de 10^{-4}. On voit que les kernels de phase déduits des images, pourtant sans bruit ne sont pas au niveau du plancher attendu : la procédure de fabrication des kernels n'est visiblement

pas parfaite et il est légitime de s'interroger sur ce qui en affecte la performance. En plus d'erreurs numériques attribuables à la précision de la simulation, on peut au moins évaluer l'impact de deux effets susceptibles de contribuer à cette erreur.

L'HYPOTHÈSE LINÉAIRE, qui est une approximation valide pour les petites erreurs de phase, est la première idée qui vient pour expliquer l'origine de cette erreur. On peut en estimer l'impact en mesurant, pour un écran de phase instrumental d'allure générale stable (par exemple celui proposé dans le panneau de gauche de la figure 61) mais dont on peut faire évoluer l'amplitude.

LA REPRÉSENTATION DISCRÈTE de la pupille, est une représentation approximative de la réalité continue. La discrétisation peut également contribuer à l'erreur observée dans ce genre de simulation idéale. On peut répéter l'analyse proposée pour l'étude de la linéarité avec plusieurs modèles discrets, du plus grossier au plus fin.

La figure 66 présente la synthèse de cette petite étude et représente l'évolution de l'erreur RMS de reconstruction des kernels en radians en fonction de l'amplitude RMS de l'erreur sur le front d'onde pour quatre modèles discrets, suivant une grille régulière avec une maille de taille décroissante : le modèle SP-50 (courbe bleue) avec une maille de 50 cm, le SP-40 (courbe verte) avec une maille de 40 cm, le SP-30 (courbe rouge) avec une maille de 30 cm et enfin le SP-20 (courbe cyan) avec une maille de 20 cm (courbe cyan). L'allure générale de ces courbes confirme le comportement attendu :

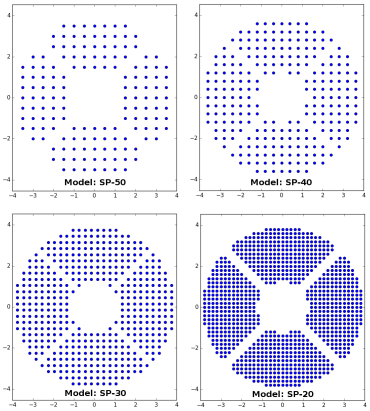


FIGURE 65: Les quatre modèles de pupille utilisés pour cette analyse.

— pour ces images à signal sur bruit infini, la fidélité de la reconstruction augmente avec la finesse du modèle discret. La courbe bleue, correspondant au modèle le plus grossier (148 échantillons répartis tous les 50 cm) présente l'erreur de reconstruction la plus importante, quelle que soit l'amplitude de l'aberration sur le domaine représenté. La courbe cyan (964 échantillons répartis tous les 20 cm) présente elle l'erreur la plus faible.

— pour tous les modèles, l'erreur de reconstruction augmente avec l'amplitude de l'aberration introduite. Sur le domaine exploré, l'effet de non-linéarité est négligeable. On a une erreur de reconstruction systématique dont l'amplitude semble proportionnelle à l'amplitude de l'aberration.

On peut confirmer cette observation en affichant sous forme d'image, l'évolution de l'erreur sur les kernels en fonction de l'amplitude de l'aberration d'entrée. Egalement proposée par la figure 66, la structure globale de cette représentation 2D, montre que les erreurs individuelles évoluent effectivement linéairement en fonction de l'amplitude. La non-linéarité ne semble vraiment dominer cette erreur que beaucoup plus tard, lorsque l'amplitude dépasse ~ 250 nm.

Ce que l'on constate cependant, c'est qu'augmenter la résolution du modèle semble effectivement conduire à une meilleure adéquation entre les données extraites de cette simulation idéale et le mo-

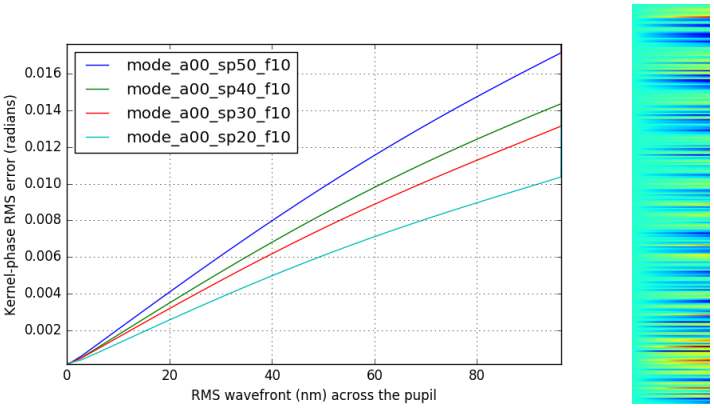


FIGURE 66: A gauche : Erreur de reconstruction des kernels pour quatre modèles discrets, en présence d'une aberration à géométrie fixe mais d'amplitude variable (ici entre 0 et 100 nm RMS). A droite : représentation sous forme d'image de l'évolution de l'erreur de reconstruction par kernel pour le modèle le plus grossier (correspondant à la courbe bleue). L'évolution linéaire de l'erreur sur les 262 kernels (affichés de haut en bas) en fonction de l'amplitude de l'aberration (de gauche à droite).

dèle théorique attendu. Avec cette métrique, le gain n'est cependant pas à la hauteur de l'effort supplémentaire et même pour le modèle le plus fin (avec un pas de 20 cm à l'échelle de la pupille), une erreur résiduelle subsiste. En pratique, d'autres sources d'incertitude comme le bruit de photon ou le bruit de lecture, vont contribuer au budget d'erreur et imposer une finesse de modèle au delà de laquelle il ne sera plus pertinent de travailler.

LA MODÉLISATION PROPOSÉE N'EST DONC PAS ENCORE IDÉALE. Une partie des erreurs observées semble attribuable à une mauvaise prise en compte des araignées. Pour s'en convaincre, on peut comparer les erreurs associées aux tests précédents à une nouvelle série de simulations, cette fois pour une pupille sans araignées et un modèle qui lui corresponde. Pour un même niveau d'aberration (140 nm RMS), quatre modèles avec les mêmes pas de 50, 40, 30 et 20 cm sont utilisés pour produire les courbes de kernels représentées dans la figure 67. Pour faciliter la lecture de cette figure, les valeurs des kernels extraits pour chaque modèle sont ordonnées par ordre croissant de la valeur des kernels théoriques. On voit que même pour le modèle le plus grossier, l'erreur est réduite d'un facteur 2. Pour aller vers de meilleures performances, le critère binaire qui détermine à l'heure d'aujourd'hui si un point de la grille discrète fait partie ou pas du modèle doit être modifié pour mieux prendre en compte les structures comme les araignées et les bords de la pupille des télescopes.

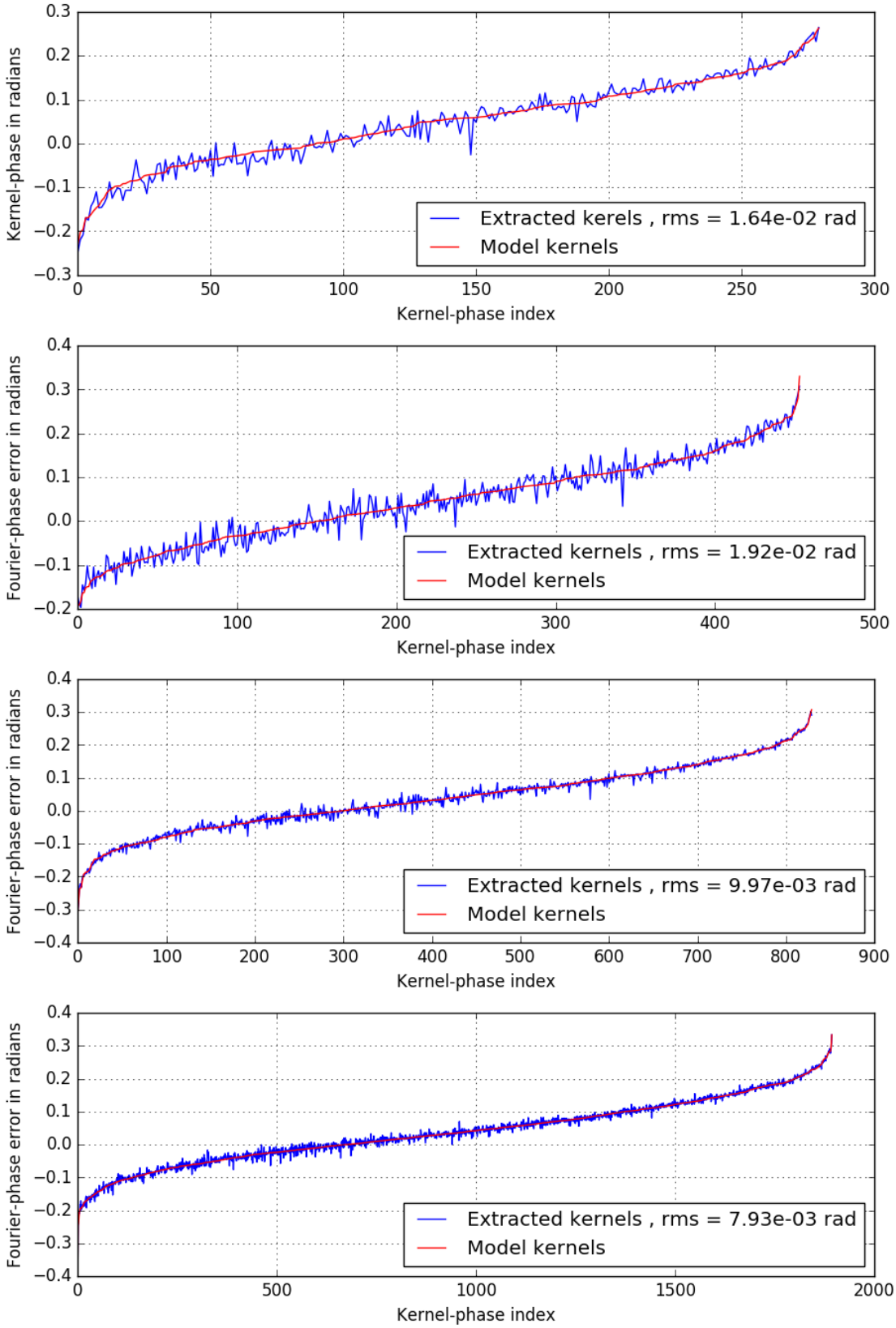


FIGURE 67: De haut en bas : courbes de kernels théoriques (en rouge) et observées (en bleu) pour quatre modèles appliqués à la même image, correspondant à la simulation d'une pupille sans araignées.

Une illustration de la limite du critère binaire est proposée dans la figure 68 : le point du modèle mis en évidence, éliminé par le critère binaire contribue en réalité en partie à un grand nombre de bases interférométriques du réseau virtuel : sa contribution est cependant atténuée par l'effet d'obstruction partielle par la structure de l'araignée. Pour prendre en compte cet effet, il faut modifier la construction de la matrice de transfert **A**. Pour la Figure 48, celà revient à prendre en compte la possibilité de sommer des vecteurs avec une norme variable.

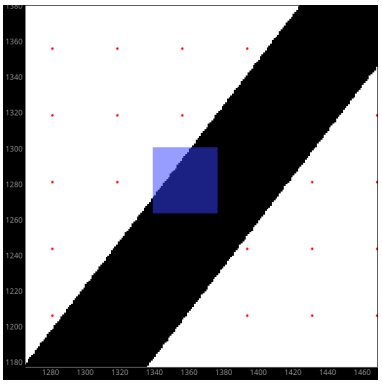


FIGURE 68: Exemple de mauvaise modélisation des araignées : le point d'échantillonnage, correspondant à la zone mise en évidence en bleu au centre de l'image est éliminé du modèle de base utilisé pour calculer les kernels. L'occultation par l'araignée n'est pourtant pas totale. Il est possible de mieux décrire cette pupille en associant à ce point pour le moment manquant une transmission τ dont la valeur sera comprise entre 0 et 1.

Cette norme est donnée par le produit normalisé des surfaces surfaces équivalentes des sous-ouvertures virtuelles contribuant à la base interférométrique considérée. Ainsi, pour chacune des lignes de **A**, au lieu d'ajouter ± 1, il faut rajouter ce produit normalisé.

Le code d'extraction des kernels disponible en ligne permet de prendre ce cas de figure en compte : si le modèle discret fourni en entrée inclut une colonne avec des valeurs de transmission locale ($0 < \tau < 1$), XARA construit en effet automatiquement un modèle dit "gris". L'utilisation de cette construction modifiée fait l'objet d'un des projets de thèse en cours dans le cadre du projet ERC KERNEL. C'est une fonctionnalité qui devrait jouer un rôle important lorsque l'on cherchera à extraire des kernels d'une image produite par une pupille apodisée (voir la conclusion de ce mémoire).

Etalonnage

J'ai tenté de fournir une description assez détaillée des défauts du modèle discret existant ainsi que de proposer quelques pistes qui permettent de réduire les biais qui leur sont associés. Même si on peut regretter que la méthode ne soit pas parfaite, la méthode proposée fonctionne et réussi à éliminer le bruit de phase. Dans son livre passionnant "The Measure of All Things" [78] décrivant le travail épique de mesures géodésiques par Pierre Méchain et Jean-Baptiste Delambre ayant mené à la définition du mètre étalon pendant la Révolution Française, l'auteur américain Ken Adler associe la compréhension des biais dans les mesures physiques, qui est une étape clé de l'entrée dans la science moderne, à l'épopée de Delambre et Méchain. Les biais, négligeables lorsque la précision d'une mesure en domine le budget d'erreur (comme c'était le cas avant la mise au point par Borda d'appareils de mesures de précision appelés cercles répétiteurs), sont désormais une composante majeure des mesures en astronomie, dont la précision ne cesse d'augmenter. L'histoire décrite par Adler révèle ainsi qu'une erreur systématique entre deux séries de mesures du méridien a été dissimulée pour préserver la réputation d'un des deux protagonistes de cette aventure.

78. K. Adler. *The Measure of All Things*. Free Press, 2002

HEUREUSEMENT, LE BIAIS D'EXTRACTION PEUT ÊTRE ÉTALONNÉ. [79] La procédure d'étalonnage des mesures, classiquement utilisée en interférométrie ou en imagerie haut contraste, est d'alterner l'obser-

79. Je fais attention ici à bien utiliser le terme français approprié, certains de mes collègues s'amusant souvent à me reprendre lorsque j'utilise le terme anglophone de calibration !

vation de sources de référence (ou calibrateurs !) à celle des sources du programme astrophysique justifiant le temps passé au télescope. Pour fournir un bon étalonnage, ces sources de référence doivent présenter des caractéristiques difficiles à réunir en pratique dans le cas d'observations faites depuis le sol :

— ces sources sont soit non-résolues soit parfaitement connues pour permettre d'estimer le biais de mesure associé à l'instrument

— elles sont de magnitude et type spectral comparable aux objets d'intérêt pour produire des données avec un rapport signal sur bruit semblable et aient bénéficié d'une correction adaptative similaire

— elles sont aussi proches que possible des sources d'intérêt, à la fois pour que la lumière traverse une atmosphère avec des propriétés comparables de seeing et de réfraction mais aussi pour que les déformations mécaniques du télescope restent aussi stables que possible

Les bonnes sources de calibration sont donc très rares et en pratique, on se contente de ce que l'on trouve. On peut améliorer les choses en synthétisant cet étalon idéal en combinant ensemble les signaux acquis sur plusieurs sources. Depuis l'espace (voir Figure 69), les critères sont plus faciles à réunir car on n'a pas besoin de prendre l'atmosphère en compte néanmoins, on voit en pratique que pour atteindre les hauts niveaux de performance (ie. contraste), la région du détecteur utilisée pour acquérir l'image a un impact sur le résultat l'étalonnage. La mise au point d'une stratégie idéale d'étalonnage fait partie des chantiers au programme de l'ERC KERNEL.

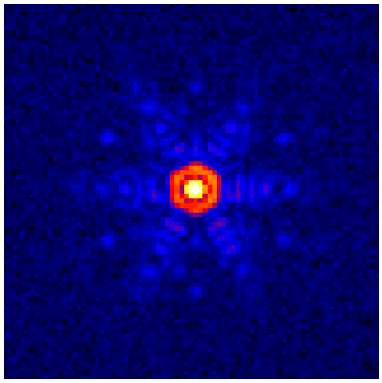


FIGURE 69: Exemple d'observation spatiale simulée (λ = 480 nm) d'une naine brune de magnitude 14, par l'instrument NIRISS du JWST.

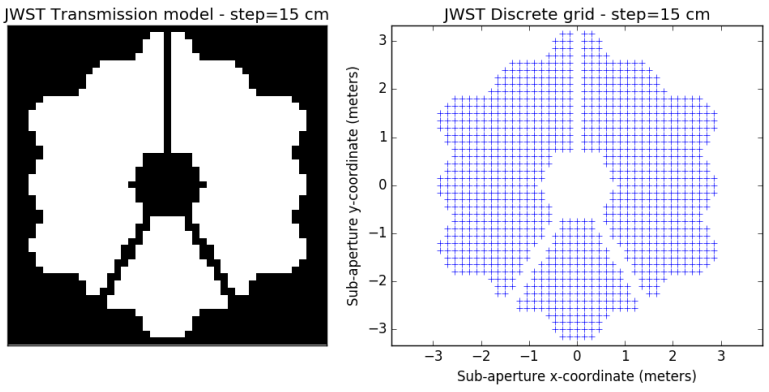


FIGURE 70: Construction d'une représentation discrète de la pupille du JWST, avec des sous-ouvertures virtuelles placées sur une grille régulière de pas 15 cm.

Dans le cas des kernels, comme pour la clôture de phase, les kernels de la source de référence sont simplement soustraits de ceux de la source d'intérêt. Je vais illustrer l'impact de cette étape d'étalonnage en utilisant des données simulées de l'instrument NIRISS du futur télescope spatial JWST, produites par Alexandra Greenbaum (University of Michigan) dans le cadre d'un projet de prédiction des performances de l'approche KERNEL appliquée à JWST dans un programme de recherche de companons autour de naines brunes

ultra-froides. Pour cette démonstration on utilise encore une matrice de transfert bâtie à partir d'un modèle discret binaire (ie. pas d'information de transmission), avec un pas 15 cm (voir Figure 70).

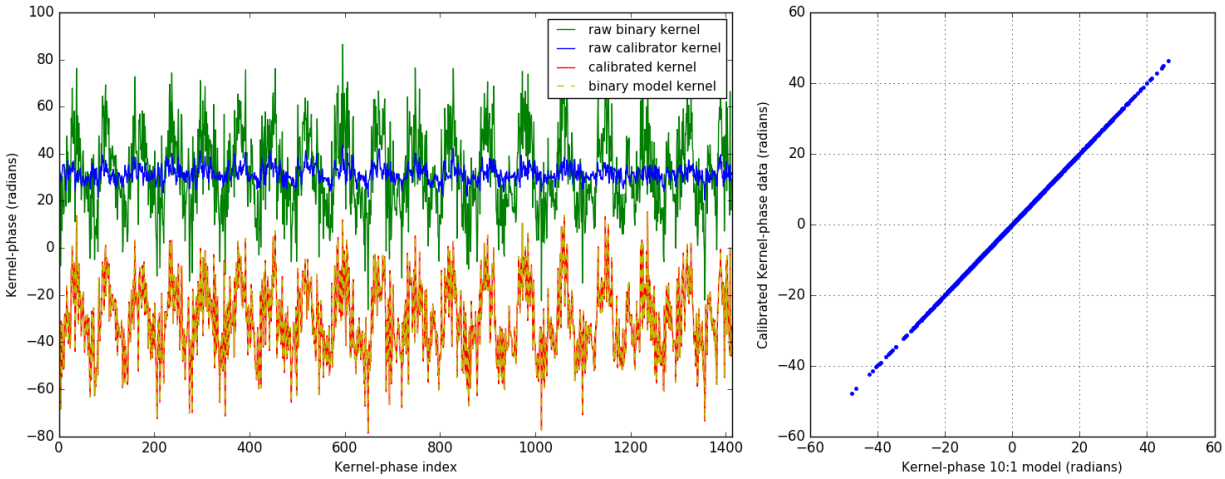


FIGURE 71: Illustration de l'impact de l'étalonnage des kernels (rapport signal sur bruit infini). Le graphique de gauche présente les valeurs contenues dans quatre vecteurs de kernels : les kernels bruts acquis sur un objet binaire sont en vert, les kernels bruts acquis sur une source de calibration sont en bleu. La différence entre les deux courbes, est représentée en rouge et est comparée à un modèle de l'objet (courbe jaune pointillée). A droite : les kernels étalonnés (courbe rouge à gauche) sont directement comparés au modèle attendu (courbe jaune pointillée à gauche).

Le biais de modélisation est commun à toutes les observations et peut être estimé sur une source de référence. Pour ce modèle particulier, le biais est représenté par la courbe bleue du panneau de gauche de la Figure 71. Pour des images simulées sans bruit de lecture et sans bruit de photon, on peut voir que la soustraction des kernels d'une source de calibration (courbe bleue) à ceux d'une source d'intérêt (courbe verte) conduit à une observation étalonnée (courbe rouge) en parfait accord avec ce que prédit modèle théorique (objet binaire). Le panneau de droite de la figure 71 montre en effet dans ce cas où le rapport signal sur bruit est infini, une corrélation parfaite entre les kernels étalonnés et un modèle de source présentant un compagnon à une séparation angulaire de 147 mas et un contraste de 10.

En présence de bruit dans l'image, l'estimation du signal contenu dans les kernels étalonnés reste non-biaisée mais le bruit de photon et le bruit de lecture du détecteur introduisent de la dispersion. La Figure 72 présente ce cas de figure et montre que malgré tout, la détection d'un compagnon avec contraste de 100 à la même séparation angulaire reste tout à fait faisable.

Recherche de compagnon

Indépendemment des notions de fidélité de modèle évoquées avant étalonnage, les caractéristiques de ce modèle vont également imposer la taille du champ de vue dans lequel la mesure des kernels a du sens. En effet, à l'instar de l'impact de la densité des actionneurs d'un miroir déformable imposant la taille de la région de contrôle de l'optique adaptative, le pas s du modèle discret de la pupille de diamètre D définit une fréquence maximale $f_c = 0.5 \times D/s$ et par

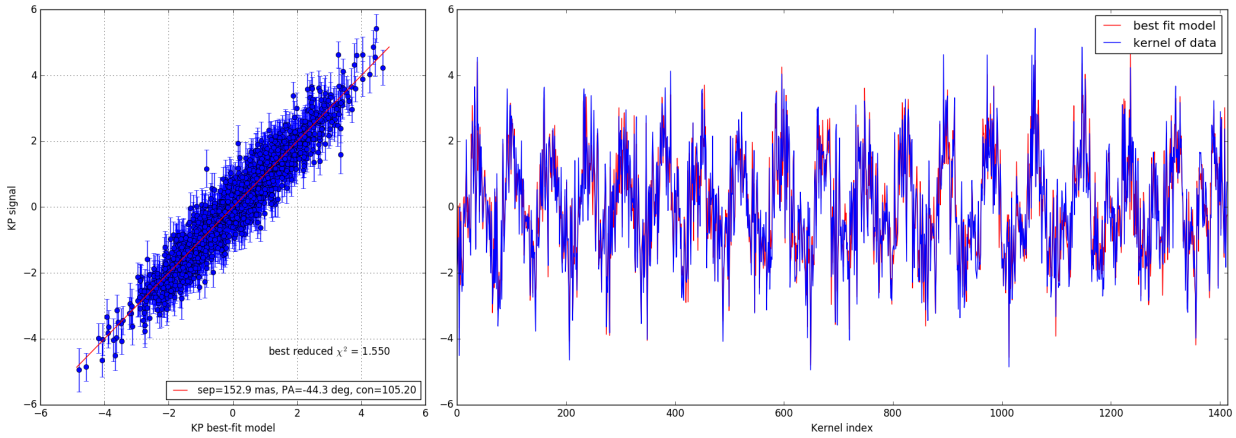


FIGURE 72: Résultat d'analyse de d'observations simulées réalistes par NIRISS/JWST d'un objet comparable à la naine brune ultra-froide WISE 1405+5534.

conséquent une taille de champ effective au delà de laquelle la mesure des kernels est affectée par de l'aliasing.

Un modèle avec un pas $s = 0.4$ m sur la pupille du JWST ($D = 6.5$ m) présente donc une fréquence de coupure $f_c \approx 8\,\lambda/D$, ce qui à la longueur d'onde des données considérées ici ($\lambda = 4.8\,\mu$m)[80], correspond à un champ de travail de rayon $r \approx 1200$ mas.

80. Encore un exemple d'utilisation de la formule 2

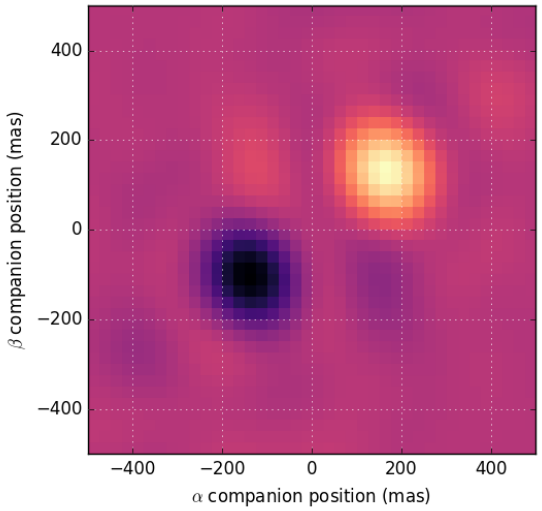


FIGURE 73: Exemple de carte de colinéarité entre le vecteur de kernels calibrés extrait des données et le signal kernel théorique résultant de la présence d'un compagnon pour une grille prédéfinie de positions.

Pour des compagnons suffisamment contrastés ($c < 0.01$ - on parlera de régime haut contraste), l'amplitude du signal kernel θ induit par la présence de ce compagnon est directement proportionnelle au contraste. La recherche initiale de compagnon peut donc être menée assez simplement à partir d'une grille de positions (α, β), rentrant dans la région contrainte par la fréquence de coupure du modèle pour lesquelles on calcule un vecteur de kernels théoriques $\theta^M(\alpha, \beta)$. La position de la grille qui maximise le produit scalaire :

$$c(\alpha, \beta) = \theta \cdot \theta^M(\alpha, \beta) \tag{53}$$

est la position pour laquelle le vecteur de kernels théoriques θ^M se

projette le mieux sur le signal réel θ. La figure 73 présente un exemple de carte de ce produit scalaire (on parlera de carte de colinéarité). A l'instar des kernels qui sont une mesure de l'asymétrie d'une source, cette carte de colinéarité est antisymétrique. La carte présente un maximum dans le quadrant en haut à droite, à 45°, indiquant pour la grille choisie, le point le plus probable de présence d'un compagnon.

Cette position, contrainte par le pas de la grille retenue (ici de 25 mas par pixel), sert ensuite de point de départ à un ajustement de modèle à trois paramètres (séparation angulaire ρ, angle de position ω et contraste c). Le modèle retenu par l'algorithme d'optimisation minimise le critère de variance :

$$\chi^2 = \sum_{k=1}^{N} \frac{(\theta_k - \theta_k^M(\rho, \omega, c))^2}{\sigma_k^2}, \tag{54}$$

où les valeurs σ_k sont les incertitudes associées aux différents kernels mesurés. Un exemple de résultat d'ajustement a été présenté dans la figure 72.

Limites de détection

Dans le régime haut-contraste, il a été mentionné que l'amplitude globale du signal kernel est directement proportionnelle au contraste. Cette propriété peut être exploitée pour expliciter le modèle local du vecteur de kernels :

$$\theta^M(\rho, \omega, c) = c \times \theta^0(\rho, \theta). \tag{55}$$

On peut alors calculer la dérivée du χ^2 (voir équation 54) par rapport au contraste et chercher la valeur de c qui minimise le χ^2 en identifiant celle qui annule cette dérivée :

$$\frac{\partial \chi^2}{\partial c} = 0 \iff \sum_{k=1}^{N} \frac{\theta_k \cdot \theta_k^0}{\sigma_k^2} = c \times \sum_{k=1}^{N} \left(\frac{\theta_k^0}{\sigma_k}\right)^2. \tag{56}$$

Pour déterminer la valeur locale de la limite de détection en contraste, il faut expliciter l'incertitude de cette quantitée, qui peut être évaluée en utilisant la formule classique de propagation des erreurs [81] :

81. Négligeant, on y reviendra plus tard, les termes de covariance.

$$\sigma_c^2 = \sum_{k=1}^{N} \left(\frac{\partial c}{\partial \theta_k}\right)^2 \times \sigma_k^2. \tag{57}$$

L'incertitude sur la valeur du contraste est alors donnée par :

$$\sigma_c = 1 / \sqrt{\sum_{k=1}^{N} \left(\frac{\theta_k^0}{\sigma_k}\right)^2}. \tag{58}$$

Ainsi dans le régime haut contraste, à partir des seules incertitudes contenues dans le vecteur σ et les valeurs théoriques des kernels θ^M, on peut déterminer les limites de détection pour toutes les valeurs de séparation angulaire et d'angle de position. La figure 74 présente les limites de détection associées aux données simulées de

NIRISS/JWST qui accompagnent ce texte depuis le chapitre sur l'étalonnage, pour des séparations angulaires allant de 0 à 1200 mas.

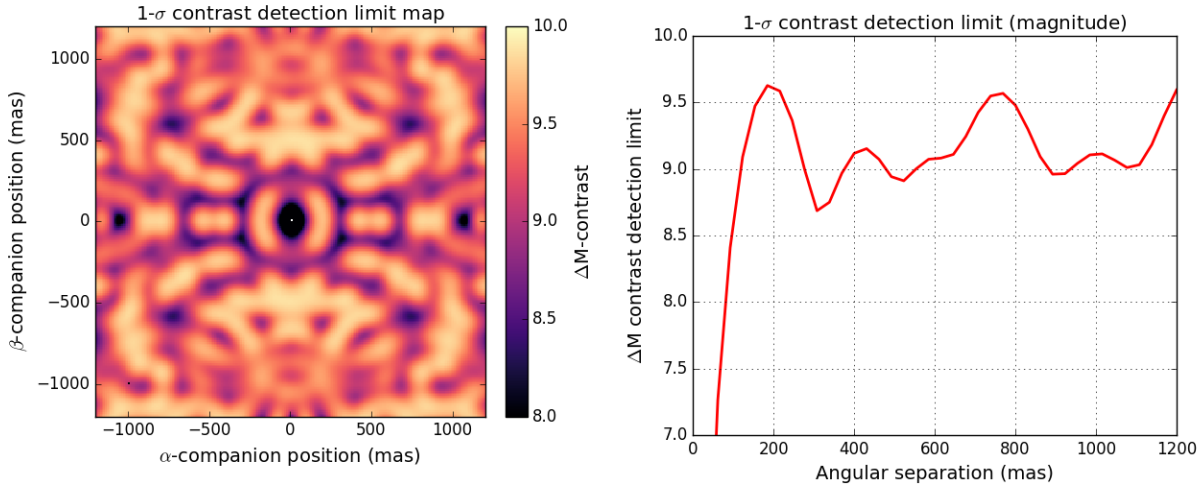


FIGURE 74: Limites de détection en contraste (différence de magnitude ΔM) obtenues entre 0 et 1200 mas de séparation angulaire à partir des données JWST/NIRISS présentées, et analysées avec un modèle discret avec un pas de 40 cm. A gauche : représentation sous forme de carte couvrant tout l'espace image accessible au modèle. A droite : profil dans la direction à 45°.

Pour le cas de figure représenté, la performance est assez remaquable. La source simulée est de magnitude 14 à la longueur d'onde considérée. Une intégration totale de 40 minutes sur la source d'intérêt résulte en un total de 1.2×10^7 photons (un temps identique doit être passé sur une source de calibration). Au delà de λ/D, la sensibilité est de l'ordre de $\Delta M \sim 9$ magnitudes (niveau de détection à 1-σ). La performance se dégrade de façon prédictible en observant des objets plus faibles, avec une limite de détection en flux directement proportionnelle à la luminosité de la source centrale. Le mode de réduction kernel des images prises par JWST présente un très fort potentiel : les délais malheureux retardant la mise en service de cet observatoire nous offrent le temps d'offrir kernel comme mode de réduction officiel des données du JWST. Cette étude du potentiel de kernel pour JWST fait l'objet d'une publication en cours de préparation dans le cadre du projet KERNEL.

Evolutions de l'approche

On constate qu'à part aux très faibles séparations angulaires, c'est à dire jusqu'à $\sim$ 1 λ/D qui correspond à 150 mas, les limites de détection sont assez uniformes, ce qui ne répond pas tout à fait à l'expérience de la recherche de compagnons dans des images à haut contraste. En effet, en s'éloignant du coeur de la PSF dominant une image, la contribution locale de la diffraction et donc celle du bruit de photon de la source brillante diminue, ce qui devrait rendre les détection plus facile et se traduire par une augmentation des limites de détection.

Par construction, les signaux contenus dans les kernels (issus d'un calcul de SVD) sont orthogonaux les uns aux autres. Les kernels ne sont théoriquement plus affectés par le bruit de phase instrumentale mais restent affectés par d'autres bruits de mesure affectant l'image,

en particulier le bruit de photon et le bruit de lecture du détecteur utilisé. Pour mieux prendre ces bruits en compte et réfléter leur impact sur la mesure de kernels, il convient de pouvoir acquérir une matrice de covariance des kernels, incluant sur sa diagonale les erreurs précédemment utilisées pour estimer les limites de détection de l'équation 58 mais contenant en sus les termes non-diagonaux de covariance.

Dans son article discutant de l'origine des erreurs sur les mesures de clôtures et kernels de phase, mon collègue M. Ireland [82] propose d'utiliser cette covariance $\mathbf{C_K}$ des kernels pour construire une nouvelle combinaison linéaire de kernels pour lesquels les bruits originalement corrélés sont rendus orthogonaux. La matrice de covariance $\mathbf{C_K}$ étant définie positive peut se décomposer :

82. M. J. Ireland. Phase errors in diffraction-limited imaging: contrast limits for sparse aperture masking. **MNRAS**, 433:1718–1728, Aug. 2013

$$\mathbf{C_K} = \mathbf{S}^T\mathbf{D}\mathbf{S}, \tag{59}$$

où $\mathbf{S}$ est une matrice unitaire et $\mathbf{D}$ une matrice diagonale contenant les incertitudes liées à ces nouvelles combinaisons linéaires de kernels pour lesquels les effets du bruit sont maintenant diagonalisés (M. Ireland parle de kernels statistiquement indépendants). Une nouvelle utilisation de l'équation 58 dans ce nouveau cas de figure révèle des limites de détection en contraste qui augmentent avec la séparation angulaire comme attendu en présence de bruit de photon.

Avec les méthodes d'extraction et d'analyse telles que je les ai décrites, a t'on atteint les limites de ce qui est possible ou peut on améliorer les limites de détection ?

Cette question fondamentale est au coeur du premier sujet de thèse du projet ERC KERNEL. Il semble qu'en plus d'une meilleure prise en compte de la covariance, la formulation de différents tests statistiques permet d'améliorer le potentiel de la méthode et de mieux préciser la véritable probabilité de détection d'un objet pour une probabilité de fause alarme choisie à l'avance. Cette thèse fait l'objet d'une co-direction offrant une passerelle entre les communautés du traitement formel du signal et de l'interférométrie à un télescope (que ce soit avec un masque non-redondant ou en pleine pupille). En explorant les limites actuelles de la méthode, le projet permet également de définir des pistes permettant de les transcender. Je vais cependant devoir m'arrêter là sous peine de commencer à empiéter sur des travaux en cours de réalisation par Alban Ceau, doctorant du projet KERNEL.

Analyse de surface d'onde dans l'espace de Fourier

Les noyaux de phase sont construits à partir des vecteurs propres de la matrice de transfert de phase **associés à des valeurs singulières nulles**. Une analyse de surface d'onde à partir de l'information de phase extraite dans le plan de Fourier va se passer **dans l'espace orthogonal à celui des kernels**, correspondant aux vecteurs propres associés aux valeurs singulières non nulles.

Si l'on jette un oeil en arrière sur la figure 53, représentant les différentes valeurs singulières associées à la matrice de transfert d'un modèle discret de pupille comme celui de la figure 52, on voit que sur les N_A valeurs propres attendues de la matrice de transfert, exactement $N_A/2$ sont non-nulles. Ainsi, même pour une pupille de télescope typique, une partie de l'information sur la qualité de la surface d'onde dans la pupille peut être déduite de l'analyse directe d'une PSF, du moins dans le régime des faibles aberrations. Comprendre pourquoi exactement la moitié de l'information est directement récupérable requiert de rappeler quelques notions concernant la parité des aberrations.

Pair ou impair ?

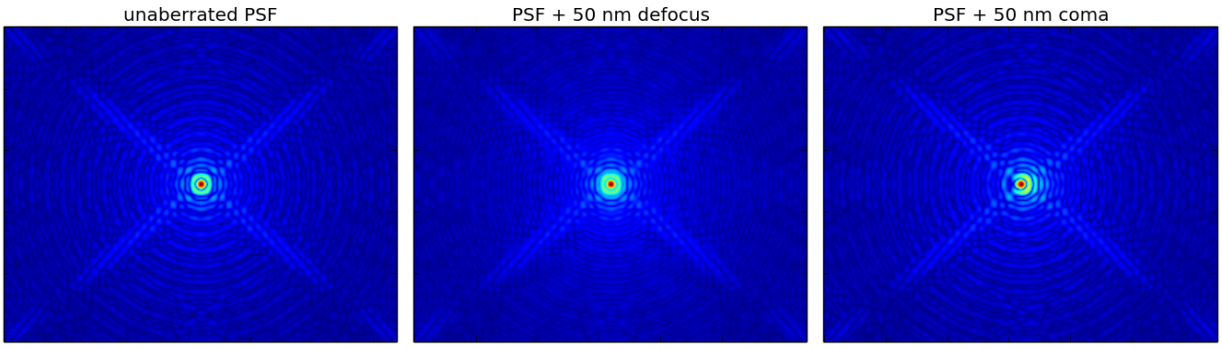


FIGURE 75: PSF simulées pour l'instrument SCExAO pour la longueur d'onde 1.6 μm. De gauche à droite : la PSF idéale, en l'absence d'aberration ; la PSF affectée par 50 nm de défocus ; la PSF affectée par 50 nm de coma.

Sur le domaine défini par le contour d'une pupille, un front d'onde aberré peut être décomposé de façon unique comme la somme de deux termes : une partie paire et une partie impaire. Un cas particulier d'une telle décomposition est celle en séries de Fourier. Tous les termes peuvent être regroupés en deux termes : une première partie paire et une deuxième partie impaire, correspondant respectivement à la somme des termes en cosinus et en sinus de cette décomposition.

Pour une pupille symétrique, le terme d'ordre pair ne présente aucune signature dans le plan de Fourier. Une démonstration formelle de cette propritété est possible[83]. Cependant il est possible d'intuiter cette propriété en observant ce qui se passe sur quelques cas particuliers. Dans ce but, la figure 75 compare l'effet sur la PSF instrumentale de SCExAO d'une aberration de type paire (ici du dé-

83. F. Martinache. The Asymmetric Pupil Fourier Wavefront Sensor. **PASP**, 125:422–430, Apr. 2013

focus) à celle d'une aberration impaire (de la coma). Un moment de contemplation de cette image convaincra le lecteur que bien que la signature du défocus soit visible et reconnaissable (si tant est qu'on en ait l'habitude) dans le plan image, il est impossible de savoir si ce dernier est positif ou négatif, les deux conduisant en effet à la production d'une PSF identique. Une PSF affectée par un défocus pur reste parfaitement centro-symétrique. Pour la coma, qui est impaire, la situation est différente : non seulement la signature est visible et reconnaissable dans l'image, le signe de l'aberration peut être identifié par la direction dans laquelle les anneaux de diffraction sont plus prononcés (ici du côté droit de l'image).

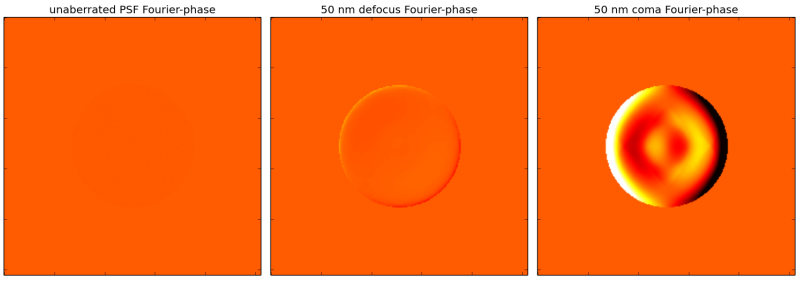


FIGURE 76: Phase de la transformée des trois PSFs précédentes. Pour aider à la comparaison, les trois images partagent la même échelle verticale. Seule la signature de la coma (troisième rangée) est visible dans cet espace.

Les $N_A/2$ valeurs singulières non-nulles produites par la SVD de la matrice de transfert de phase d'un modèle discret symétrique décrivent la partie impaire du front d'onde. Les valeurs nulles correspondent à tous les modes contribuant à la partie paire. On peut vérifier en comparant la phase de la transformée de Fourier des trois PSFs précédentes (voir la figure 76) : seule la coma présente une signature claire dans l'expression de la phase de la transformée de Fourier des images.

Diversité de phase ?

L'utilisation d'une image focale à des fins de métrologie fait traditionnellement appel à la technique de la diversité de phase[84]. Pour résoudre le problème des solutions dégénérées, on introduit une diversité de phase, c'est à dire, une modulation du front d'onde (classiquement, du défocus) d'amplitude connue, qui mélangée avec les autres aberrations déjà présentes dans la pupille, produit une seconde image permettant de trouver une solution unique au front d'onde.

84. R. A. Gonsalves. Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21:829–832, Oct. 1982

Peut-on envisager de mettre cette approche en oeuvre dans le cadre de KERNEL ? Non. En ne se basant que sur l'information contenue dans la phase dans le plan de Fourier qui est au coeur de l'approche KERNEL, la diversité de phase n'apporte rien : le défocus introduit s'ajoute simplement (dans la limite linéaire) aux aberrations déjà présentes. Si elle est d'ordre pair, elle ne se traduit par aucune signature dans le plan de Fourier. On ne bénéficie pas de l'effet de mélange.

Une telle approche est-elle impensable pour autant ? Peut-être pas. Mais elle requiert de réviser un peu la méthode : en plus de s'intéresser en détails à la phase de la transformée de Fourier de l'image, il faudrait également s'intéresser à son module. Une telle approche n'a pas encore été tentée. Contrairement au travail sur la phase de Fourier qui bénéficie de relations linéaires claires, une technique utilisant en plus l'information contenue dans le module de la transformée de Fourier des images implique sans doute un algorithme non-linéaire. Sinon, une approche hybride en deux étapes semble également possible : une première analyse de la phase dans le plan de Fourier pour déterminer la partie impaire du front d'onde, complétée par une analyse dans l'image (ou exploitant également le module de la transformée de Fourier) pour contraindre les modes d'aberration paire en présence d'une diversité. Ces approches feront peut être l'objet de travaux dans le futur. J'ai pour le moment choisi d'aborder le problème de façon différente.

Asymétriser la pupille

L'approche que j'ai mis en oeuvre dans SCExAO propose de lever cette dégénérescence en rendant la pupille elle même asymétrique. Cette asymétrie est introduite par un masque conjugué avec la pupille de l'instrument qui obstrue une partie de la pupille de façon non centro-symétrique. Une image de la pupille de l'instrument ainsi asymétrisée et de la PSF qui en résulte sont visibles dans la figure 77. Le design du masque n'a encore fait l'objet d'aucune optimisation : l'épaisseur de l'asymétrie a été choisie de façon à introduire une asymétrie clairement identifiable dans la pupille mais avec un impact cosmétique mineur dans l'image. La méthode n'est donc certainement pas encore optimalement sensible et la géométrie de ce masque mériterait sans aucun doute d'être revue. Néanmoins, tel qu'il est décrit, il répond déjà aux besoins de l'instrument et a permis de mettre en oeuvre des méthodes de contrôle de front d'onde inédites.

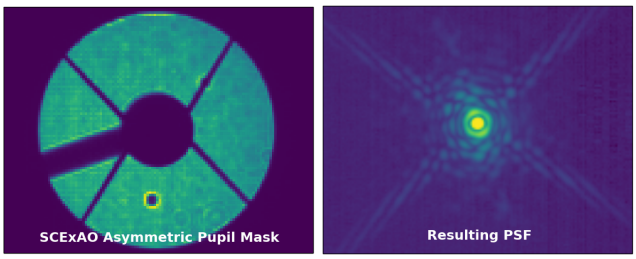


FIGURE 77: A gauche : pupille de l'instrument SCExAO. Conjugué avec la pupille du télescope, le masque rajoute une barre d'épaisseur égale à la moitié du diamètre de l'obstruction centrale, dont l'azimuth exact n'a pas été contrôlé au moment de la mise en place. A droite : la PSF résultante. La barre épaisse introduit une aigrette de diffraction supplémentaire dans la direction perpendiculaire à l'azimuth de la barre.

A cette pupille asymétrique est associé un nouveau modèle discret, qui inclut la connaissance de cette asymétrie (voir la figure 78). La SVD de ce nouveau modèle révèle autant de valeurs singulières non-nulles que de points dans la pupille moins un. Alors qu'on parle de N_K noyaux de phase pour les kernels, on parlera des N_E phases propres de la matrice de transfert de phase (*eigen-phases* en anglais) :

$$N_E = N_A - 1. \tag{60}$$

La différence de -1 entre le nombre de sous-pupilles virtuelles et le nombre de phases propres provient de l'impossibilité de connaître la référence absolue de la phase du front d'onde. En pratique, cela veut dire qu'une sous-ouverture virtuelle du modèle de la pupille est choisie comme référence. La mesure du front d'onde est une mesure du piston de chaque sous-ouverture virtuelle par rapport à cette référence. On dispose de suffisamment de phases propres (N_E) pour inverser la matrice de tranfert de phase (en pratique, un pseudo-inverse de Moore-Penrose $\mathbf{A}^+$ est calculé) et produire une mesure de la surface d'onde pour la partie non-occultée de la pupille à partir d'une seule image :

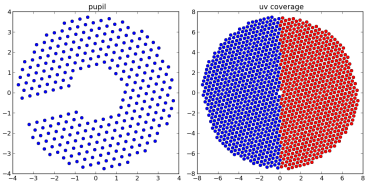


FIGURE 78: Modèle discret incluant la barre asymétrique

$$\varphi = \mathbf{A}^+ \cdot \Phi. \tag{61}$$

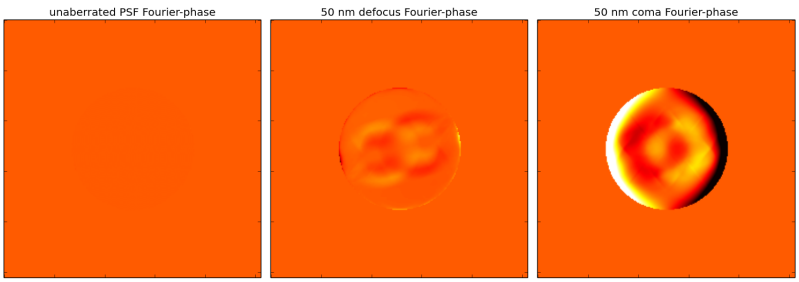


FIGURE 79: Phase de la transformée de Fourier des trois cas de figure précédents (sans aberration, avec défocus et avec coma), après insertion du masque asymétrique. Encore une fois, les trois images partagent la même échelle verticale. Avec l'asymétrie, la signature du défocus est maintenant visible.

La figure 79 répète l'analyse présentée dans la figure 76, seulement cette fois, en présence du masque asymétrique. Une fois le masque en place, la signature du défocus, initialement absente de la phase dans le plan de Fourier, est rendue visible - même si elle ne l'est pas autant que celle de la coma. Un critère d'optimisation de l'asymétrie pourrait être d'égaliser l'amplitude de la signature entre les modes pairs et impairs. Un optimum serait vraisemblablement atteint si 50 % de la pupille est masquée. La pertinence d'une technique occultant une large partie de la pupille serait discutable.

Mesure de front d'onde dans le plan focal

En pratique, le pseudo-inverse $\mathbf{A}^+$ de la matrice de transfert de phase est calculé (tout comme la matrice des kernels $\mathbf{K}$) à partir du résultat de la SVD de la matrice de transfert. Cette procédure permet de choisir le nombre de modes qui vont contribuer au pseudo-inverse et limiter la sensibilité aux erreurs de discrétisation qui ont été discutées dans le cas des kernels.

La figure 80 présente, dans le cas d'une simulation numérique, comment la qualité de la reconstruction évolue en fonction du nombre total de modes propres de la matrice de transfert de phase conservés pour contribuer au pseudo-inverse. Comme dans le cas des kernels,

on remarque qu'une bonne adéquation entre le modèle discret et la géométrie exacte de la pupille est importante si l'on veut reproduire le contenu haute fréquence de l'image, correspondant aux modes d'ordre élevé.

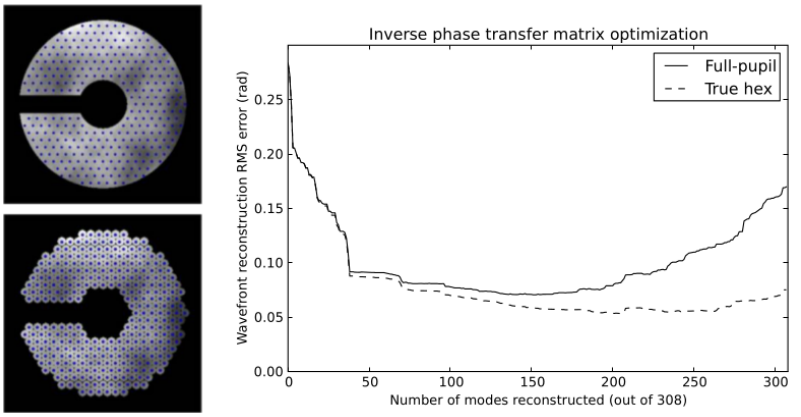


FIGURE 80: Evolution de la qualité de la reconstruction d'un front d'onde de type Kolmogorov, pour un nombre variable de modes préservés dans la construction du pseudo-inverse. Pour la pupille circulaire originale, la qualité de la reconstruction (courbe pleine) atteint un optimum lorsque environ 150 modes sont conservés sur les 309 possibles. Au delà, l'erreur de reconstruction augmente. On peut voir qu'une grande partie de ces erreurs sont attribuables à des erreurs de discrétisation : une pupille modifiée pour être plus proche du modèle discret présente une erreur qui n'augmente pas significativement en incluant un plus grand nombre de modes.

Au départ, deux masques identiques étaient installés dans SCExAO, la seule différence étant l'azimuth de la barre rendant la pupille asymétrique. L'analyse de front d'onde proposée étant aveugle à la partie du front d'onde occultée par la barre, l'idée était de combiner deux mesures utilisant les deux masques de façon à permettre une reconstruction complète du front d'onde sur l'intégralité de la pupille. Cette approche double masque n'a pas été exploitée en pratique. La mise en oeuvre sur SCExAO s'est limitée à une utilisation avec un seul masque, permettant un fonctionnement rapide en boucle fermée. Le but de cette boucle de contrôle est de mesurer et de compenser l'erreur de chemin optique non-commun attendue après chaque repointage du télescope. Ces erreurs d'alignement sont bien décrites par les aberrations classiques de Gauss-Seidel et correspondent aux premiers polynômes de Zernike (voir la figure 81).

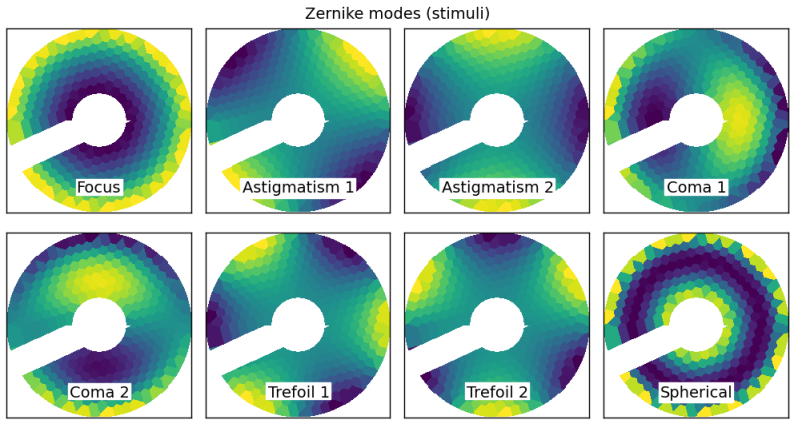


FIGURE 81: Modes de front d'onde décrits par les polynômes de Zernike allant de Z4 (défocus) à Z11 (asphéricité), interceptés par le masque asymétrique, utilisés pour compenser l'erreur de chemin optique non-commun présente en sortie de l'AO188.

UN FONCTIONNEMENT EN BOUCLE FERMÉE REQUIERT, en plus de la mesure de front d'onde, d'inclure les propriétés du miroir déformable (DM) : la correspondance géométrique exacte entre les positions des actionneurs et celles des sous-ouvertures virtuelles du modèle discret et la fonction d'incluence des actionneurs. Ces propriétés peuvent être modélisées et intégrées à une boucle de contrôle. Un tel modèle est en pratique difficile à maintenir sur le long terme. Les systèmes d'optique adaptative emploient un mode de fonctionnement pragmatique permettant d'inclure automatiquement ces informations en utilisant une matrice de réponse, acquise dans une première phase d'étalonnage.

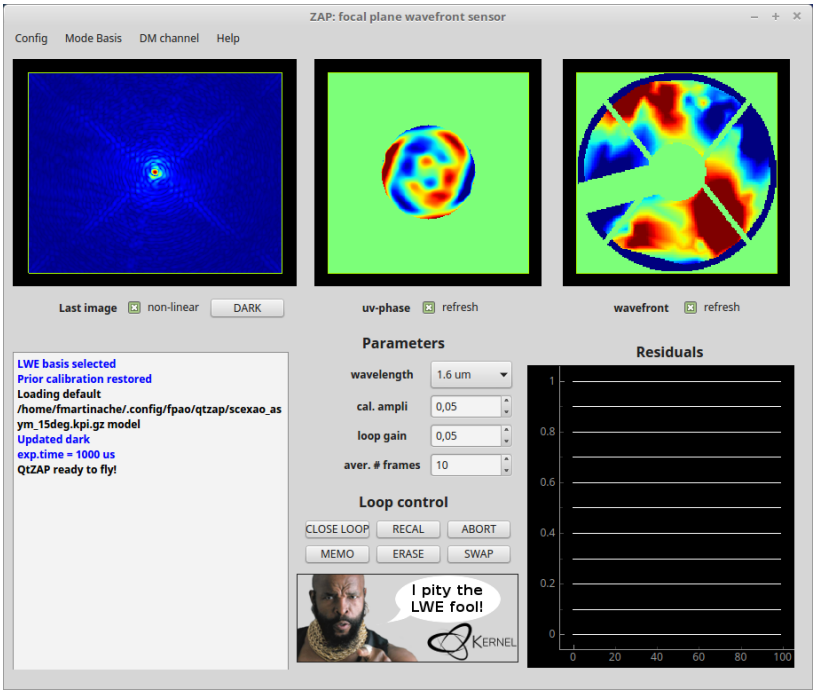


FIGURE 82: Capture d'écran du logiciel de contrôle de front d'onde plan focal de SCExAO : ZAP (*Zernike Asymmetric Pupil*) permet après chaque pointage du télescope et après fermeture de la boucle d'optique adaptative par l'AO188 du télescope, de compenser le chemin optique non-commun.

Cet étalonnage se fait en excitant de façon séquentielle les actionneurs du DM, pour générer (en ajoutant à la position de référence du DM) les modes de front d'onde que l'on souhaite contrôler pour une amplitude fixe (amplitude typique de quelques dizaines de nanomètres). La mesure de front d'onde faite par l'analyse dans plan focal pour chaque stimulus est intégrée à une matrice de réponse, notée $\mathbf{Z}$. Le modèle discret utilisé étant constitué de 292 sous-ouvertures virtuelles (voir la figure 78), la matrice de réponse contient huit vecteurs-colonnes de 291 éléments. En boucle fermée, la mesure instantanée du front d'onde φ est projetée sur cette base de modes pour trouver la solution $\hat{\alpha}$ au sens des moindres carrés de l'équation : $\mathbf{Z} \cdot \alpha = \varphi$:

$$\hat{\alpha} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\varphi. \tag{62}$$

Cette solution est utilisée pour fabriquer un terme de compensation : une combinaison linéaire de modes utilisés pendant la phase d'étalonnage, multipliée par un gain global (typiquement compris

entre 5 et 30 %), qui est rajouté sur un canal dédié du DM. Des corrections répétitives font converger la forme du DM vers un état qui, du point de vue de la caméra focale, fournit une PSF dont les aberrations de bas ordre sont absentes. Ce mode de contrôle est assuré par un outil appelé ZAP (*Zernike Asymmetric Pupil*) dont une capture d'écran est proposée dans la figure 82. En pratique, des corrections de quelques dizaines de nanomètres sur chacun des modes sont nécessaires après chaque réacquisition d'une étoile par l'optique adaptative.

En bande H ($\lambda = 1.6\,\mu$m) l'analyse fonctionne de façon satisfaisante sur une plage d'amplitude de $\pm$200 nm sur le front d'onde : poussée au delà de cette plage de fonctionnement, l'amplitude totale P2V (*peak-to-valley*) enregistrée dans la phase de Fourier peut dépasser 2π : la méthode devient inutilisable, à moins peut être d'utiliser un algorithme de dépliement de la phase, ce qui n'a pas été tenté.

La figure 83 montre comment la réponse expérimentale enregistrée sur une source de calibration évolue en fonction de l'amplitude du stimulus [85]. Le comportement est différent selon la structure des modes et de l'amplitude P2V de la phase de Fourier qui leur est associée. La réponse de la mesure est linéaire sur une plage de $\pm$100 nm (sur le front d'onde) pour tous les modes et peut être étendue un peu au delà car même si elle devient non-lineaire, la courbe de réponse reste monotone et la convergence en boucle fermée possible.

85. Notez juste que les valeurs des amplitudes sont données sur le DM et doivent donc être doublées pour correspondre aux valeurs sur le front d'onde

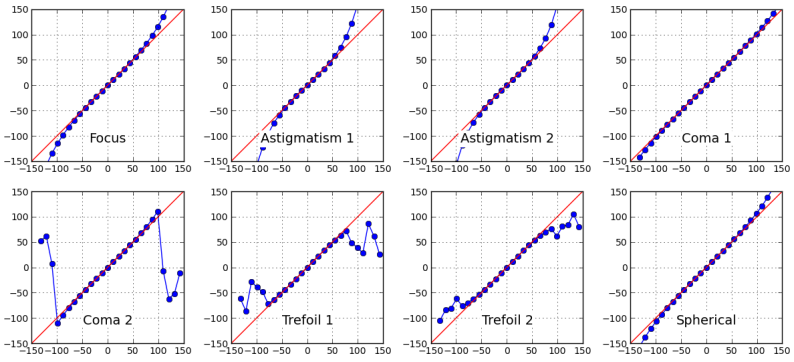


FIGURE 83: Réponse expérimentale de l'analyse de surface d'onde à pupille asymétrique pour des stimulus de Zernike d'amplitude variable. Les valeurs indiquées sont données pour le DM et doivent par conséquent être doublées pour correspondre au front d'onde.

Deux publications sont incluses : la première présentant l'article qui a introduit la méthode à pupille asymétrique et la seconde présentant son implémentation pratique dans le cadre de l'instrument SCExAO. Dans le cadre du projet KERNEL, nous avons pu démontrer [86] que la base des modes contrôlés pouvait être étendue et inclure des modes correspondant à un phénomène inattendu affectant tous les instruments XAO appelé "effet de faible vent". La présentation de ce résultat récent dépasse le cadre de cette HDR. Il montre cependant que la méthode proposée est polyvalente et que des améliorations sont possibles pour permettre un contrôle d'une plus grande famille de modes, en particulier si un meilleur détecteur est utilisé.

86. M. N'Diaye, F. Martinache, N. Jovanovic, J. Lozi, O. Guyon, B. Norris, A. Ceau, and D. Mary. Calibration of the island effect: Experimental validation of closed-loop focal plane wavefront control on Subaru/SCExAO. *ArXiv e-prints*, Dec. 2017

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **125**:422–430, 2013 April
© 2013. The Astronomical Society of the Pacific. All rights reserved. Printed in U.S.A.

The Asymmetric Pupil Fourier Wavefront Sensor

FRANTZ MARTINACHE

National Astronomical Observatory of Japan, Subaru Telescope, Hilo, HI 96720; frantz@naoj.org

Received 2012 October 05; accepted 2013 March 12; published 2013 March 27

ABSTRACT. This article introduces a novel wavefront sensing approach that relies on the Fourier analysis of a single conventional direct image. In the high Strehl ratio regime, the relation between the phase measured in the Fourier plane and the wavefront errors in the pupil can be linearized, as was shown in a previous work that introduced the notion of generalized closure-phase, or kernel-phase. The technique, to be usable as presented requires two conditions to be met: (1) the wavefront errors must be kept small (of the order of one radian or less), and (2) the pupil must include some asymmetry, which can be introduced with a mask, for the problem to become solvable. Simulations show that this asymmetric pupil Fourier wavefront sensing or APF-WFS technique can improve the Strehl ratio from 50% to over 90% in just a few iterations, with excellent photon noise sensitivity properties, suggesting that on-sky close loop APF-WFS is possible with an extreme adaptive optics system.

Online material: color figures

1. NON-COMMON PATH ERROR AND EXTREME AO

Contrast limits for the direct imaging of extrasolar planets from ground based adaptive optics (AO) observations are currently set by the presence of static and slow-varying aberrations in the optical path that leads to the science instrument (Marois et al. 2003). Because some of these aberrations are not sensed by the wavefront sensor, something called the non-common path error, they are responsible for the presence of long lasting speckles in the image. Since extrasolar planets are faint unresolved sources, it is impossible to discriminate them among these speckles in one single frame. The family of differential imaging techniques is aimed at calibrating out some of these static aberrations, in postprocessing by using either sky rotation (angular differential imaging, or ADI), polarization differential imaging (PDI), or wavelength dependence of the speckles (spectral differential imaging or SDI). Of these, ADI (Marois et al. 2006) has been successful in most notably producing the image of the planetary system around HR 8799 (Marois et al. 2008).

A new generation of high-contrast instruments using an approach of improved wavefront control called extreme adaptive optics (XAO), aims at producing higher contrast raw images, by including additional high-density wavefront control devices, fed by advanced wavefront sensing techniques to actively control the wavefront during the observations. In all cases, the primary goal of the high-order deformable mirror is to calibrate the non-common path error between the conventional AO system and the science camera, and then to try and help creating a high-contrast region in the image, by actively modulating the speckles in the field, for instance using speckle nulling (Martinache et al. 2012).

The control of the non-common path error is an important element of any XAO system, which requires the implementation of dedicated hardware or special acquisition procedures: the Gemini Planet Imager (Wallace et al. 2010) and the Project P1640 (Zhai et al. 2012) for instance, both include a dedicated interferometric calibration unit to characterize and compensate the non-common path error. Other options often involve some form of phase diversity, using multiple acquisitions with the science camera while moving an internal calibration source along the optical path with ESO's Spectro-Polarimetric High-Contrast Exoplanet Research (SPHERE; Sauvage et al. [2010]) and JPL's PALM-3000 (Bouchez et al. 2010) or even the science camera itself with the Subaru Telescope Extreme Adaptive Optics (SCExAO) project (Martinache et al. 2011).

This article introduces an alternate approach to the calibration of the non-common path error in XAO systems, using a noninvasive conventional hard-stop mask located in the pupil plane, and the direct analysis of the Fourier properties of the resulting images acquired with the focal plane camera. The simulations presented in this work demonstrate remarkable performance of the approach: combined with proper control of a deformable mirror (DM) that is present in all XAO systems, it can bring the wavefront quality from a starting Strehl ratio of 50% up to over 95% in less than five iterations. Because of the mask's very limited impact on the overall morphology of the point-spread function (PSF) and its high sensitivity, the technique seems compatible with a general purpose imaging (non-coronagraphic) instrument, and would enable continuous

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

tracking of the non-common path aberrations, to maintain very high Strehl during observation.

2. WAVEFRONT SENSING IN THE FOURIER PLANE

Even in near-perfect observing conditions, diffraction is a significant hindrance to the interpretation of images at the highest angular resolution. In this diffraction-limited regime, it is advantageous to adopt an interferometric point of view of image formation, and work not from images directly, but from their Fourier transform counterpart instead. Working on the Fourier transform of images is often refered to as working in the Fourier plane, the spatial-frequency plane or the u-v plane, the latter expression being widely used in sparse aperture interferometry. The Fourier transform being a complex function, encodes information both in terms of amplitude (or modulus) and phase: this article focuses solely on the phase.

To introduce the notion of kernel phase, Martinache (2010) proposes a simple but powerful description of how pupil phase errors propagate into phase signal measured in the Fourier plane, in terms of linear algebra. This model, when it applies (when wavefront errors are small, i.e., $\varphi < 1$ radian for a conventional imaging telescope), offers an interesting alternative to the traditional object–image relation, written in terms of convolution. When this model applies, the (usually unknown) instrumental phase in the pupil φ can be related to the phases Φ measured in the Fourier plane with a single, fully determined, linear operator called the phase transfer matrix (noted $\mathbf{A}$). The discrete model, presented in several references (Martinache 2010, 2011, 2012) is briefly reintroduced here for convenience.

The aperture of any instrument, usually referred to as the pupil, can be discretized into a finite collection of n elementary subapertures, using a regular grid (for instance square or hexagonal) adapted to the actual pupil shape. One of these elementary subapertures is taken as a zero-phase reference: the pupil phase of a coherent point source is therefore written as a $(n-1)$-component vector φ measured relative to this reference. Assuming that the image produced by the system is at least Nyquist sampled, in the Fourier plane, one is able to sample up to m phases, organized in a vector Φ.

The complete model (Martinache 2010), also includes one additional term Φ_0 that encodes the true phase information about the target of interest. The Fourier phase Φ therefore is written as:

$$\Phi = \Phi_0 + \mathbf{A} \cdot \varphi. \tag{1}$$

The motivation for this article is to use this formalism not to learn something about the target, which was shown to be recoverable using kernel phases, but to solve the wavefront sensing problem. This article will therefore from now on consider the target to be a nonresolved source, so that $\Phi_0 = 0$. XAO systems typically include a laser-fed single-mode fiber calibration source, for offline tests of wavefront control loops that can be used in this purpose.

Wavefront sensing in the Fourier plane is routinely achieved in long baseline interferometry. However, to be able to handle the large optical path differences (OPD) expected along a long baseline that can be up to several hundred meters long in the optical, resulting in phase shifts larger than 2π, the strategy is to disperse the light so as to increase the coherence length and solve the 2π wrapping of the phase (Michelson & Pease 1921; Koechlin et al. 1996). This strategy, so far only used with a small number of apertures was generalized by Martinache (2004), for a all in one Fizeau-type interferometer: an approach called "dispersed-speckles". This approach appears as an relevant solution for the calibration of OPD inside AO fed integral field spectrographs (Peters et al. 2012). In systems where the OPD is small (~1 radian), dispersion is not necessary, and the use of a single wavelength may suffice in solving the problem.

This approach however has one major flaw: for a pupil like the one shown in Figure 1, it is insensitive to structures of the wavefront that are even, that is for which $\varphi(-r) = \varphi(r)$. While an analytical demonstration of this fundamental property is provided in an Appendix to the article, one can build a more intuitive understanding of wavefront sensing in the Fourier plane, by considering the following scenarios, that illustrate how instrumental (pupil) phase errors φ propagate into phase measured in the Fourier plane:

1. If the phase is constant across the entire pupil, then none of the baselines formed by any pair of elementary subapertures records a phase difference, and the phase in the Fourier plane is

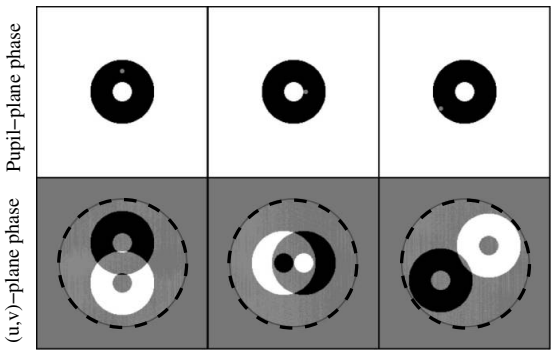


FIG. 1.—Illustration of the pupil phase propagation mechanism into phase measured in the Fourier plane. The *top row* shows a pupil phase "poke" applied to three different subapertures of the full two-dimensional pupil, in this case, a circular aperture with a 30% central obscuration. The *bottom row* shows for each case the corresponding distribution of phase that is observed in the Fourier plane. The overlaid *dashed-line circle* in the *bottom row* marks the cutoff spatial frequency of the transfer function.

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

zero everywhere; the non-common path error is perfectly calibrated.

2. If a phase offset δ_0 is added to one single subelement of the aperture, something we will refer to a "poke", then each baseline involving this subelement records a phase difference, which is exactly $\pm\delta_0$. Figure 1 presents several such scenarios.

3. If the pupil–plane phase error φ is completely random, each of the samples in the Fourier plane is then the average of **R** phase differences on the pupil, where **R** is a vector summarizing the redundancy of the baselines identified in the system.

To see that this approach to wavefront sensing would be insensitive to an even aberration, the reader can consider either of the scenarios shown in Figure 1, and observe that the phase pattern recorded in the u-v plane for either of these "pokes" is perfectly odd: $\Phi(-r) = -\Phi(r)$, a property expected to the fact that it results from the Fourier Transform of a real image.

If the same amplitude "poke" δ_0 is applied to the diametrically opposed subaperture in the same direction (therefore creating an elementary even aberration), the two resulting phase patterns in the Fourier plane will simply add to zero, and give the illusion that that the system is perfectly in phase, no matter what the amplitude of the "poke". One equivalent way to say it is that such a Fourier-based wavefront sensor is only sensitive to the odd component of the input wavefront.

One simple way to circumvent this severe limitation is to introduce some asymmetry in the system, and the simplest thing to alter is the pupil. High-contrast instruments are usually equipped with a multislot pupil wheel (i.e., for coronagraphic Lyot-stops), and for this work, we propose to modify the pupil used in Figure 1 by adding one single fairly thick spider arm that connects the outer edge of the pupil to the edge of the central obstruction. The actual pupil used for the results presented in this work is shown in Figure 2.

3. ASYMMETRIC PUPIL FOURIER WAVEFRONT SENSOR

We need to go back to the instrumental phase propagation model given summarized by equation (1), to see how the asymmetry introduced in the pupil affects the properties of the phase transfer matrix **A**. To build the actual matrix, the pupil is modeled into a discrete grid of subapertures that follow a regular hexagonal grid. Other options (e.g., square grid) remain possible, and have been used so far for kernel-phase analysis (Martinache 2010; Pope et al. 2012). The density of the grid should in practice be matched to the density of actuators the deformable mirror that controls the incoming wavefront. The image is expected to be at least Nyquist sampled.

Without the missing subapertures along the thick arm, the model would contain $n_a = 330$ pupil samples, propagating into $n_{uv} = 708$ distinct samples in the Fourier plane. Just like for the kernel phase, the proposed analysis relies on the spectral

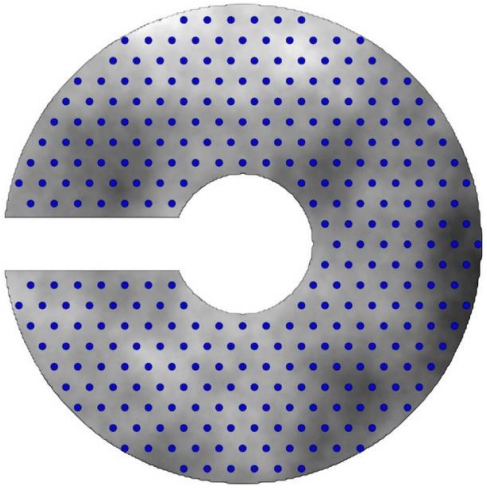


FIG. 2.—Discrete hexagonal model used to construct the phase transfer matrix **A**. Each of the *thick dots* overlaying the "true" image of the input pupil phase inside the asymmetric pupil, marks the location of a discrete pupil sample. This model contains 309 distinct pupil sample points. See the electronic edition of the *PASP* for a color version of this figure.

properties of the phase transfer matrix **A**. The singular value decomposition (SVD) of the 708×330 resulting phase transfer matrix **A**, given by:

$$\mathbf{A} = \mathbf{U} \cdot \Sigma \cdot \mathbf{V^T}, \tag{2}$$

where **U** and $\mathbf{V^T}$ are unitary matrices and Σ is a diagonal matrix containing the singular values of **A**, reveals that for this symmetrical geometry, 165 (that is exactly $n_a/2$) of the singular values of **A** are nonzero. Using this knowledge, it therefore appears possible to build a pseudoinverse for **A** that will enable the recovery of one half of the wavefront phase in the input pupil. This observation made after examination of the singular values of the phase transfer matrix, echoes the conclusion made in § 2, that a Fourier-based WFS can only sense the odd component of the input wavefront, in terms of linear algebra.

With the asymmetric pupil presented in Figure 2, things apparently change very little: the number of apertures in the pupil may drop to 309 (a loss of 21 apertures), they still project onto the same space of 708 samples in the Fourier plane, because of the redundance. However, a major change occurs in terms of spectral properties of this new phase transfer matrix. The SVD of the now 708×308 linear operator reveals that all the required 308 singular values are now nonzero. A pseudoinverse for **A** for such a configuration therefore enables full recovery of both odd and even parts of the wavefront.

In practice, the approach is used as follows: assuming that the residual aberrations on the wavefront are of the order of

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

one radian or less (rms), one asymmetric mask such as the one shown in Figure 2 is introduced in the pupil of the instrument, before the final focal plane, while shining a calibration source (often a laser-fed single-mode fiber). With nothing else being actuated, one (or more) image is acquired with the science detector, which will be assumed to be Nyquist sampled.

The data reduction procedure is very similar to what has been described for non-redundant masking interferometry data (Kraus et al. 2008) and kernel-phase data analysis (Martinache 2010). Each image is simply recentered, and multiplied by a window function to limit sensitivity to readout noise and filter out the spatial frequency content of the image that is beyond the control region of the wavefront control device. A Fourier transform for each image is then calculated, and the phase of this Fourier transform is sampled according to the discrete model described in § 2. The resulting vector Φ is then simply projected back onto pupil phase φ inside the region unocculted by the mask, using the pseudoinverse of $\mathbf{A}$.

4. APF-WFS PERFORMANCE

4.1. SVD Modes for the APF-WFS

The wavefront reconstruction by the APF-WFS relies upon the determination of singular value modes of the phase transfer matrix that in turn solely depend on the geometry of the pupil sampling. The modes, which form an orthonormal basis for the wavefront, are also sorted in order of decreasing level of significance, proportional to their associated singular value.

Figure 3 shows the first 30 modes (out of the total 309) the SVD of the phase transfer matrix produces. One will observe that although they do not closely match a familiar pattern of modes like the Zernike polynomials or the series of sines and cosines, the lower spatial frequencies (slope and curvature) are naturally reconstructed first, and higher spatial frequency content appears as the order of the modes increases, guided by the hexagonal geometry chosen to model the pupil.

4.2. Optimization of the Pseudoinverse for Wavefront Reconstruction

While fairly dense, the 309-element discrete representation of the pupil used to establish the linear relation for this APF-WFS remains an approximate representation of the continuous circular aperture that is the real pupil. While not significant for low-order modes of the wavefront, the impact of this discrete sampling is expected to increase for the high-order modes. This, combined with the fact that the same high-order modes are associated with small singular values (for this geometry, the singular values span over three orders of magnitude from $\sim$0.1 for the lowest to $\sim$150 for the highest), suggests that the wavefront reconstruction fidelity by APF-WFS will depend on the total fraction of modes one tries to reconstruct. The result of a series of simulations, presented here, involving the reconstruction of

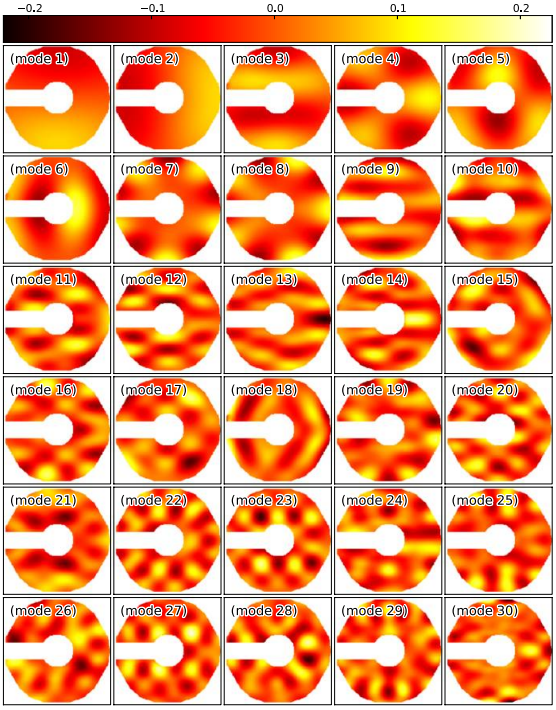


FIG. 3.—Representation of the first 30 SVD modes for the asymmetric pupil geometry shown in Figure 2, sorted in order of decreasing significance (proportional to their associated singular value). All modes are represented using a common color scale, that spans the $\pm$0.2 radian range. See the electronic edition of the *PASP* for a color version of this figure.

an identical input wavefront using different reconstructions illustrates this property.

The input wavefront is a Kolmogorov phase screen, scaled down to a Strehl ratio $\sim$85%. One asymmetric pupil mask, such as the one shown in Figure 2 is used to simulate the acquisition of one single, photon noise-free PSF, processed following the procedure highlighted in § 3. We look at the result of a reconstruction of the input wavefront, based on the direct inversion of equation (1):

$$\varphi = \mathbf{A}_\mathbf{k}^+ \cdot \Phi, \tag{3}$$

where $\mathbf{A}_\mathbf{k}^+$ is one possible pseudoinverse of the phase transfer matrix $\mathbf{A}$, calculated only after keeping the first k singular values contained in Σ:

$$\mathbf{A}_\mathbf{k}^+ = \mathbf{V} \cdot \Sigma_\mathbf{k}^+ \cdot \mathbf{U}^\mathbf{T}. \tag{4}$$

Each of these pseudoinverses then acts upon the same sampled Fourier phase vector Φ to produce an estimate of the wavefront. Figure 4 plots the reconstruction error rms as a function of

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

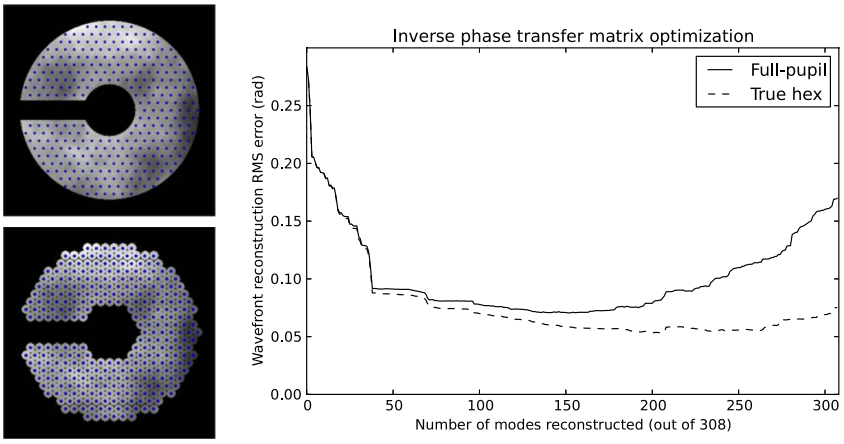


FIG. 4.—Curve of the APF-WFS wavefront reconstruction error (rms) as a function of the number of modes included in the reconstruction. Two curves are shown to highlight the impact of the discrete representation of the wavefront. The *solid line* curve, shows the reconstruction rms for the circular pupil shown in Figure 2, and reproduced in the *top left* panel. The figure shows that including the first 40 modes exhibits the most important increase in the wavefront reconstruction fidelity, which then slowly increases until reaching an optimum around 150 modes. Including more modes however increases the reconstruction error. The *dashed line* curve shows how the reconstruction rms evolves for the labeled "true hex" pupil (*bottom left*) that more closely matches the discrete sampling used to model the optical system. See the electronic edition of the *PASP* for a color version of this figure.

the number of modes used for the pseudoinverse of the phase transfer matrix, using anything between one and 309 modes (the maximum for this pupil model). The solid-line curve shows the evolution of the reconstruction error for the a simulated circular aperture like shown in Figure 2. One will observe that the error exhibits its most rapid improvement when the number of modes included in the reconstruction increases from 1 to about 40. This is somewhat expected, given that a Kolmogorov wavefront exhibits more power in the low spatial frequencies. The reconstruction keeps slowly improving and reaches an optimum when 150 modes are used. Beyond that, the wavefront reconstruction residuals eventually increase, and the performance of the reconstruction degrades. With noise-free data, even the low-significance singular values corresponding to high spatial frequency modes should be well reconstructed, so this reduction of the wavefront reconstruction fidelity requires another explanation: the unperfect match between the hexagonal grid sampling and the actual pupil geometry, to explain this behavior. To confirm this hypothesis, Figure 4 also shows the residuals obtained under identical conditions of wavefront and signal to noise, with a pupil mask labeled "true hex", that more closely matches the sampling used in the linear model, while the inclusion of the first 40 modes shows the same increase as with the circular pupil, the reconstruction fidelity keeps improving as the number of included modes increases. While an optimum is nevertheless observed (around 200 modes), the inclusion of all 309 available modes in the reconstruction does not significantly degrade the performance. The rms can in fact be brought arbitrarily close to zero by reducing the size of the subapertures of the "true hex" model or equivalently, by filtering out more aggressively the spatial frequency content beyond the sampling.

While a pupil mask with this "true hex" geometry can easily be manufactured, one may just as easily acknowledge this characteristic as a feature of the APF-WFS, and deliberately choose to adjust the total number of modes used in the reconstruction to the optimum found for a simulation in the noise-free case. The close loop results shown in the next section will therefore be limited to the reconstruction of the first 150 modes.

4.3. Steady State Close-Loop Simulation

One anticipated application of this technique is the characterization of the small non-common path error between the fast wavefront sensor arm and the science camera in an XAO system. This simulation illustrates this use case of the technique. After the initial close loop on the high-order wavefront sensor, the Strehl ratio on the science focal image is of the order of 50%, that corresponds to a wavefront rms error of 0.8 radians.

For each iteration, a photon noise limited (10^5 photons for this series of images) single PSF is acquired, and the wavefront reconstructed by the APF-WFS is then fed back to the upstream high order wavefront sensor to offset its "flat" shape.

The simulation does not include any modeling of the deformable mirror, and simply assumes that the wavefront determined by the APF-WFS can be perfectly accounted for by the wavefront control system. Figure 5 shows the result of five consecutive iterations of one such close-loop simulation, reconstructing the first 150 SVD modes of the wavefront each time. Three iterations are sufficient to bring the original 50% Strehl ratio

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

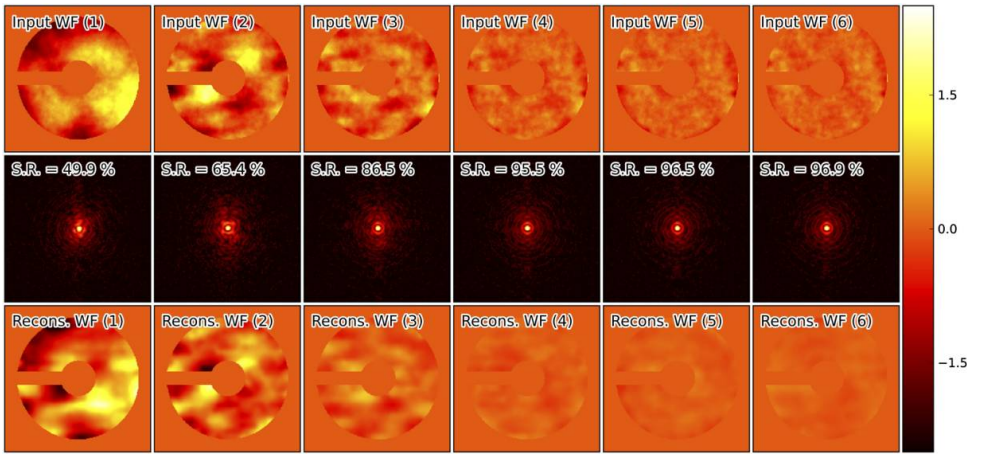


FIG. 5.—Sequence of wavefronts and resulting PSFs in a close-loop simulation of the APF-WFS, for 150 modes. The starting point of the simulation is a static Kolmogorov-type wavefront with an rms of 0.8 radians, resulting in a 50% Strehl PSF. From *left* to *right*, the figure shows the impact of successive iterations of the APF-WFS. The *top row* shows the simulated wavefront used to produce the (photon-noise limited) PSFs shown in the *middle row*. The *bottom row* shows the wavefront reconstructed from the analysis of the PSFs by the APF-WFS. All the wavefronts are represented using a common color scale, that spans the ± 2 radian range. See the electronic edition of the *PASP* for a color version of this figure.

PSF to a steady Strehl level better than 95%. Given that the linear model holds as long as $\varphi < 1$ radian, the method will converge with an initial Strehl as low as ~35%.

4.4. Sensitivity

In his overview of the limits of adaptive optics, Guyon (2005) defines a parameter noted β_p that characterizes the sensitivity to photon noise of a wavefront sensor. The rms reconstruction error by a wavefront sensor σ_m for a single mode due to photon noise is equal to:

$$\sigma_m = \beta_p \sqrt{\frac{1}{n_{ph}}}, \tag{5}$$

where n_{ph} represents the total number of photons used by the sensor to reconstruct this mode. According to Guyon (2005), the ideal wavefront sensor exhibits a photon noise sensitivity criterion $\beta_p = 1$, for each mode sensed. For reference, one of the best known and widely used wavefront sensing techniques: the Shack–Hartmann, exhibits a very nonuniform sensitivity over the modes it covers, with a peak sensitivity location that depends on the number of elements in the Hartmann screen. It is especially poorly sensitive ($\beta_p \gg 1$) for lower modes.

To illustrate how close to the ideal wavefront sensor the APF-WFS approach is, Figure 6 plots three curves that show the evolution of the rms wavefront reconstruction error for different level of illumination (10^4, 10^5 and 10^6 photons), for a random wavefront with a Kolmogorov-type structure and a Strehl ratio of 85%. The three curves, in this linear-log plot, equidistant in the vertical direction, show that the overall performance follows the expected $1/\sqrt{n_{ph}}$ trend (cf. eq. (5)).

For each of these series of simulations, Figure 6 also plots a dashed line, that represents the behavior of the ideal wavefront sensor described by Guyon (2005), with constant and optimal sensitivity $\beta_p = 1$ for all modes:

$$\sigma = \sqrt{\frac{n_m}{n_{ph}}}, \tag{6}$$

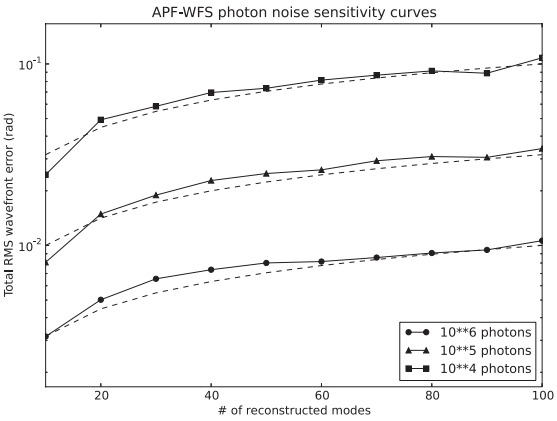


FIG. 6.—Sensitivity study of the APF-WFS. For three levels of total illumination (10^4, 10^5 and 10^6 photons), the figure shows the evolution of the total rms wavefront reconstruction error (in radians) due to photon noise.

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

where n_m and σ now respectively represent the total number of modes included in the reconstruction, and the total RMS wavefront error in radians. The close match between the three solid lines and the dashed lines over the wide range of number of modes covered (from 10 to 100) is remarkable and demonstrates that the APF-WFS, in the high-Strehl regime where it is expected to operate, exhibits almost optimal sensitivity, making it a very attractive option, in comparison to more classical wavefront sensors.

5. CONCLUSION

This presentation of the asymmetric pupil Fourier wavefront sensor completes the theoretical study of the linear model first used to introduce the notion of kernel-phase. The linear model, although only valid in the high-Strehl regime, is very powerful. The eigenvectors corresponding to zero singular values lead to kernel phases, which in turn encode information about the target of interest, and the eigenvectors corresponding to nonzero singular values provide information about the wavefront.

To recover the wavefront, some asymmetry needs to be introduced in the pupil, and one (or more) mask such as the one used for this work is a simple addition with a minimal impact, that most XAO instruments in the making should be able to afford and accommodate. Using this mask alone and a calibration source indeed enables the characterization of the wavefront from the final science detector in an accurate and very efficient manner. The region of the pupil hidden by the asymmetric mask is obviously not accounted for, and for a complete characterization of the wavefront over the entire pupil, serial introduction of a pair of masks with asymmetric arms at different azimuths is required.

Some improvements to this approach can already be anticipated: the discrete grid model may be refined by simply (1) increasing its density, and (2) including coefficients to take into account the fact that some sample points are on the edge of the pupil. The latter point may become an essential addition if one wants to make the technique compatible with apodized pupils, which are required for high contrast coronagraphy.

Given its appealing photon noise sensitivity properties, and the relatively weak impact on the overall PSF morphology, a close loop on-sky operation appears as a viable possibility, for non-coronagraphic observations. In rich fields, well separated PSFs could simultaneously be used and the corresponding estimates of the wavefront averaged to increase the signal to noise ratio. The technique as it is, focuses on the phase: wavefront amplitude errors which would require a dedicated simultaneous image of the pupil, are going to be interpreted as phase aberrations. If such amplitude errors are significant, or if the field of the instrument extends beyond the anisoplanatic patch, the technique, looking at multiple sources may however be able to do a more complete tomographic reconstruction of the wavefront in a high-Strehl MCAO system, and identify the origin of amplitude aberrations.

All simulations done so far were monochromatic, which is acceptable given that the primary application is the characterization of non-common path error in an XAO system using an internal calibration source. But an on-sky close-loop system will require operation in a broader band, and future work should explore the impact of broad band on general performance of the wavefront sensor. While one expects it to degrade somewhat, positive experience with kernel-phase analysis of data acquired in broad band filters (up to 40%) demonstrates that the linear model still holds, suggesting that the proposed approach should remain valid.

The technique is currently being implemented on the SCExAO system, with two masks with the asymmetric arm at different position angles, so as to be able to recover the wavefront over the entire pupil. A more detailed description of this implementation along with an experimental characterization of the sensor's performance will be the object of a future publication.

The author thanks Olivier Guyon for the useful discussions of the ideas presented in this work. This work is partially funded by NASA through a grant from the Space Science Telescope Institute (grant number: HST-AR-12849.01-A) which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

APPENDIX

This section demonstrates how in the absence of pupil asymmetry, even modes (for which $\varphi(-r) = \varphi(r)$) cannot be sensed in the Fourier plane. Classical diffraction theory tells us that the instant PSF is the square modulus of the Fourier transform of the complex amplitude in the pupil:

$$I = ||\mathcal{F}(Ae^{i\varphi})||^2, \tag{A1}$$

where A and φ respectively stand for the amplitude and phase of the wavefront in the pupil, and $\mathcal{F}$ symbolizes the Fourier transform. The so-called Fourier, or u-v, plane, as it is used in the article, is the Fourier transform of the image, and equivalent to the autocorrelation of the complex amplitude in the pupil plane, by virtue of the convolution theorem.

$$\mathcal{A} = Ae^{i\varphi} \otimes Ae^{i\varphi}. \tag{A2}$$

In general, kernelphase and the wavefront sensing idea presented in this article rely on the assumption that wavefront errors are small, so that the expression for the complex amplitude

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

can be linearized: $Ae^{i\varphi} \approx A(1+i\varphi)$. To simplify the notations, the functions describing the amplitude and phase in the pupil will be written as functions of a single variable x. Using the linearization, and eliminating second order terms, the convolution product of equation (2) can be explicided as follows:

$$\mathcal{A}(u) \approx \int_{-\infty}^{+\infty} A(x)A(x+u)(1+i\varphi(x)+i\varphi(x+u))dx \tag{A3}$$

$$\begin{aligned} &= \int_{-\infty}^{+\infty} A(x)A(x+u)dx \\ &\quad + i\int_{-\infty}^{+\infty} A(x)A(x+u)\varphi(x+u)dx \\ &\quad + i\int_{-\infty}^{+\infty} A(x)A(x+u)\varphi(x)dx. \end{aligned} \tag{A4}$$

The first (real) term of equation (4) only depends on the function describing the pupil amplitude $A(x)$, and not on the wavefront error $\varphi(x)$. We will therefore from now on only consider the imaginary part of this autocorrelation. To further simplify the equations, we introduce one new function $B(x) = A(x)\varphi(x)$. The imaginary part of this autocorrelation can be written as:

$$\Im(\mathcal{A}(u)) = \int_{-\infty}^{+\infty} A(x)B(x+u)dx + \int_{-\infty}^{+\infty} A(x+u)B(x)dx. \tag{A5}$$

A variable change $x' = x+u$ for the second integral term of equation (5) allows to rewrite the imaginary part of the autocorrelation as:

$$\Im(\mathcal{A}(u)) = \int_{-\infty}^{+\infty} A(x)B(x+u)dx + \int_{-\infty}^{+\infty} A(x)B(x-u)dx. \tag{A6}$$

The infinite integration domain can be split into two semi-infinite domains:

$$\begin{aligned} \Im(\mathcal{A}(u)) = &\int_{-\infty}^{0} A(x)B(x+u)dx + \int_{0}^{+\infty} A(x)B(x+u)dx \\ &+ \int_{-\infty}^{0} A(x)B(x-u)dx + \int_{0}^{+\infty} A(x)B(x-u)dx, \end{aligned} \tag{A7}$$

and one additional variable change $x' = -x$ is used for the integral terms over the negative domains. Equation (7) can be rewritten as:

$$\begin{aligned} \Im(\mathcal{A}(u)) = &-\int_{0}^{+\infty} A(-x)B(-x+u)dx \\ &+ \int_{0}^{+\infty} A(x)B(x+u)dx \\ &- \int_{0}^{+\infty} A(-x)B(-x-u)dx \\ &+ \int_{0}^{+\infty} A(x)B(x-u)dx. \end{aligned} \tag{A8}$$

The pupil is assumed to be symmetric, so that the amplitude function verifies $A(-x) = A(x)$. Similarly, the even component of the pupil phase also verifies $\varphi(-x) = \varphi(x)$. For equation (8), this translates into $B(-x) = B(x)$. For an even mode, the imaginary part of the autocorrelation becomes:

$$\begin{aligned} \Im(\mathcal{A}(u)) = &-\int_{0}^{+\infty} A(x)B(x-u)dx + \int_{0}^{+\infty} A(x)B(x+u)dx \\ &- \int_{0}^{+\infty} A(x)B(x+u)dx + \int_{0}^{+\infty} A(x)B(x-u)dx. \end{aligned} \tag{A9}$$

The reader will observe that all terms in equation (9) cancel out, which demonstrates that even modes of the pupil wavefront cannot be sensed in the Fourier plane. By contrast, the antisymmetric properties of an odd mode translates into $B(-x) = -B(x)$. Two of the four terms cancel out, leading to the simplified expression for the imaginary part of the autocorrelation:

$$\Im(\mathcal{A}(u)) = 2\int_{0}^{+\infty} A(x)B(x+u)dx. \tag{A10}$$

REFERENCES

Bouchez, A. H., Dekany, R. G., Roberts, J. E., et al. 2010, Proc. SPIE, 7736, 77361Q

Guyon, O. 2005, ApJ, 629, 592

Koechlin, L., Lawson, P. R., Mourard, D., et al. 1996, Appl. Opt., 35, 3002

Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762

Marois, C., Doyon, R., Nadeau, D., Racine, R., & Walker, G. A. H. 2003, in EAS Publications Series 8, Astronomy with High Contrast Imaging, EAS Publications Series, ed. C. Aime, & R. Soummer (Cambridge: Cambridge University Press), 233–243

Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

Martinache, F. 2004, Opt. A: Pure Appl. Opt., 6, 216
———. 2010, ApJ, 724, 464
———. 2011, Proc. SPIE, 8151, 815111
———. 2012, Proc. SPIE, 8445, 844504
Martinache, F., Guyon, O., Clergeon, C., & Blain, C. 2012, PASP, 124, 1288
Martinache, F., Guyon, O., Garrel, V., et al. 2011, Proc. SPIE, 8151, 81510Q
Michelson, A. A., & Pease, F. G. 1921, ApJ, 53, 249
Peters, M. A., Groff, T., Kasdin, N. J., et al. 2012, SPIE, 8446, 84462E
Pope, B., Martinache, F., & Tuthill, P. G. 2012, ApJ, arXiv:1302.6682
Sauvage, J.-F., Fusco, T., Petit, C., et al. 2010, Proc. SPIE, 7736, 77360F
Wallace, J. K., Burruss, R. S., Bartos, R. D., et al. 2010, Proc. SPIE, 7736, 77365D
Zhai, C., Vasisht, G., Shao, M., et al. 2012, Proc. SPIE, 8447, 84476W

This content downloaded from 192.54.176.192 on Wed, 23 Jul 2014 10:56:08 AM
All use subject to JSTOR Terms and Conditions

Astronomy & Astrophysics manuscript no. publisher_version
July 20, 2016
©ESO 2016

Closed-loop focal plane wavefront control with the SCExAO instrument

Frantz Martinache[1], Nemanja Jovanovic[2,3], Olivier Guyon[2,4,5].

[1] Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Parc Valrose, Bât. H. FIZEAU, 06108 Nice, France e-mail: frantz.martinache@oca.eu
[2] National Astronomical Observatory of Japan, Subaru Telescope, 650 North A'Ohoku Place, Hilo, HI, 96720, USA
[3] Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia
[4] Steward Observatory, University of Arizona, Tucson, AZ, 85721, USA
[5] College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

June 14, 2016

ABSTRACT

Aims. This article describes the implementation of a focal plane based wavefront control loop on the high-contrast imaging instrument SCExAO (Subaru Coronagraphic Extreme Adaptive Optics). The sensor relies on the Fourier analysis of conventional focal-plane images acquired after an asymmetric mask is introduced in the pupil of the instrument.
Methods. This absolute sensor is used here in a closed-loop to compensate for the non-common path errors that normally affects any imaging system relying on an upstream adaptive optics system.This specific implementation was used to control low-order modes corresponding to eight zernike modes (from focus to spherical).
Results. This loop was successfully run on-sky at the Subaru Telescope and is used to offset the SCExAO deformable mirror shape used as a zero-point by the high-order wavefront sensor. The paper details the range of errors this wavefront-sensing approach can operate within and explores the impact of saturation of the data and how it can be bypassed, at a cost in performance.
Conclusions. Beyond this application, because of its low hardware impact, the asymmetric pupil Fourier wavefront sensor (APF-WFS) can easily be ported in a wide variety of wavefront sensing contexts, for ground- as well space-borne telescopes, and for telescope pupils that can be continuous, segmented or even sparse. The technique is powerful because it measures the wavefront where it really matters, at the level of the science detector.

Key words. instrumendation – adaptive-optics

1. Introduction

Several approaches to high contrast imaging have now clearly demonstrated the power of focal-plane based image analysis. Most prominently, non-redundant aperture masking (NRM) interferometry (Tuthill et al. 2000), relying on interferometric calibration tricks in the focal plane has led to high contrast detections (of the order of 1000:1) in a regime of angular separation (typically between 0.5 and a few λ/D) that is still unmatched in practice by techniques like coronagraphy (Kraus & Ireland 2012; Sallum et al. 2015). As the generation of extreme adaptive optics (XAO) instruments is coming online, more advanced wavefront control schemes developed in the context of space-borne coronagraphy like speckle nulling (Bordé & Traub 2006) or the general framework of electric field conjugation (Give'On 2009) are being ported on-sky (Martinache et al. 2014; Cady et al. 2013). Nevertheless, it remains remarkable that such a venerable approach (the original masking idea by Fizeau was actually first tested in the 1870s), has remained relevant for well over a century. This is really a tribute to the deep understanding that interferometry has brought to the process of image formation.

More recently, it has been shown that the same self-calibrating tricks used in masking interferometry could in fact be applied to regular (i.e. unmasked) images, assuming AO-correction with residual wavefront errors $\leq$ 1 radian RMS. The notion of closure-phase (Jennison 1958), was indeed generalized and shown to be a special case of a wider family of self-calibrating observable quantities coined kernel-phases (Martinache 2010), since they form the basis for the null-space (or kernel) of a linear operator. This generalization also opened the way for a focal-plane-based wavefront sensing approach, relying this time on the eigen-phases of the same linear operator. While this problem is generally degenerate, one way to break this degeneracy proved to be simple, and involved masking a small but non-negligible fraction of the pupil to introduce some level of asymmetry. The principles of this asymmetric pupil Fourier wavefront sensor (APF-WFS) were described by Martinache (2013), and exploited by Pope et al. (2014) to show how it could be used, for instance, to cophase a segmented mirror. This paper further expands on the possible applications of this wavefront sensor, since it has now been implemented as part of the SCExAO instrument (Jovanovic et al. 2015b), to compensate for a non-common path error unseen by its upstream pyramid wavefront sensor.

2. Implementation of closed-loop wavefront control

2.1. Theoretical principle of APF-WFS

The APF-WFS method relies on the analysis of the Fourier properties of an AO-corrected image acquired after an asymmetric hard-stop mask has been placed in the pupil. The SCExAO instrument is equipped with two of these masks with the asymmetric feature at distinct position angles, so that every part of the in-

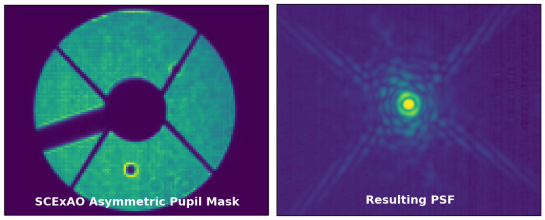


Fig. 1. Images of the pupil (left) and the focal plane (right) acquired by the SCExAO internal science camera. In addition to the four Subaru telescope spiders, the thick arm visible on the left hand side of the pupil introduces the asymmetry required for the wavefront sensing technique. The thick dot visible in the bottom pupil quadrant is induced by a dead actuator on the DM. In the focal plane, the presence of this asymmetry results in an additional set of diffraction spikes along a direction that is perpendicular to that of the arm and a lumpier first diffraction ring.

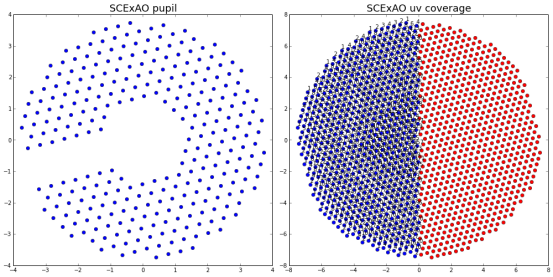


Fig. 2. Discrete model of the asymmetric pupil mask used for the calibration of the non-common path error in SCExAO. The pupil is discretized into a 292-element vector that projects onto a set of 675 equivalent interferometric baselines (or UV points) in the Fourier domain. The linear transformation that relates the wavefront to the phases measured in the Fourier transform of an image is entirely determined from this model. The presence of the asymmetry in the pupil ensures that an inverse relation for this phase -transfer matrix exists.

strumental pupil can be accounted for. A rotation wheel, located in a plane conjugated with the pupil of the instrument (see Figure 3 of Jovanovic et al. (2015b)) makes it possible to move the masks in and out of the beam as required by the observer.

Figure 1 shows an image of one of the asymmetric masks in the pupil and the point spread function (PSF) it produces. Combined with the use of the 2000-element deformable mirror (DM), this simple alteration of the pupil is a powerful tool used to control the low-order aberrations of the instrument PSF. All images featured in this paper were acquired using an H-band filter, centered on wavelength 1.65 μm and with an effective bandwidth of 0.3 μm. The pixel scale of the internal science camera is 12.1 mas per pixel which, for this wavelength, provides a sampling better than Nyquist.

It was shown that, in the low-aberration regime, typical of what is left over after a first layer of AO correction is applied, the phase Φ measured in the Fourier transform of an image $\mathbf{I}$ and the instrumental pupil phase φ are linearly related. On the internal calibration source, unaffected by atmospheric turbulence, the Strehl of images used in this study (such as during the calibration) is typically of the order of 80%. On-sky, since the results featured here were acquired before the XAO loop is closed, the Strehl is significantly lower, of the order of 50 %, which is a sufficiently good starting point for approximation to be valid.

The target phase information, associated with the spatial structures of the observed object Φ_0, is also present in the Fourier plane and simply adds to this instrumental Fourier phase. When wavefront aberrations are low (below $\sim$ 1 radian), the classical image-object convolution relation

$$\mathbf{I} = \mathbf{O} * \mathbf{PSF}, \tag{1}$$

can therefore be reformulated, if one works with the phase part of the Fourier transform of this image as follows:

$$\Phi = \Phi_0 + \mathbf{A} \times \varphi, \tag{2}$$

where $\mathbf{A}$ is an operator that describes the way the pupil phase φ propagates into the Fourier-plane.

When observing a point source, for which $\Phi_0 = 0$ (or if the object is known), this relation can be inverted if one introduces an asymmetry in the pupil (Martinache 2013). A direct focal plane image, with only a small amount of additional diffraction generated by the pupil asymmetry (see Figure 1), can therefore serve as a wavefront sensor.

To determine the structure of the operator $\mathbf{A}$, a discrete representation of the instrument pupil needs to be built - including the asymmetric mask - following a regular grid with a step such that the sampling density is reasonably representative of the original pupil. Then the way this discrete model projects into equivalent interferometric baselines in the Fourier plane needs to be studied. The model currently used on SCExAO is provided in Figure 2. It reduces the masked pupil to a 292-component vector that projects onto a 675-element vector in the Fourier domain. The phase transfer matrix $\mathbf{A}$ that establishes the mapping between the two spaces ($\Phi = \mathbf{A} \times \varphi$) is calculated using the `PYSCO` software, which is used for wavefront sensing, as well as for kernel-phase data analysis of diffraction limited images.

The presence of the asymmetry in the pupil ensures that an inverse relation for this phase transfer matrix exists, and can be used to infer the pupil-phase vector φ from the Fourier phase Φ, using the relation

$$\varphi = \mathbf{A}^+ \times \Phi, \tag{3}$$

where $\mathbf{A}^+$ is a Moore-Penrose pseudoinverse of the phase-transfer matrix $\mathbf{A}$, computed after rejecting modes associated with low singular values. The geometry of the asymmetric feature of the mask used for this work is not the result of an optimization and simply follows the shape used in the concept paper of Martinache (2013). A smaller asymmetric feature would be expected to result in a lower sensitivy but a systematic study of the sensitivity impact of the geometry of the asymmetry has yet to be done. In the mean time, the curious reader can check the experimental work of Pope et al. (2014), which shows that, in the case of a segmented aperture, the technique remains effective, even with a minimum of asymmetry (a single segment of the aperture) and suddenly breaks, if no asymmetry is present at all, validating the mathematical model this approach relies on.

2.2. Integrating a real system

The case featured in (Martinache 2013) was somewhat idealized. Working on monochromatic images, and with a perfect DM, it was able to exactly generate the wavefront correction determined

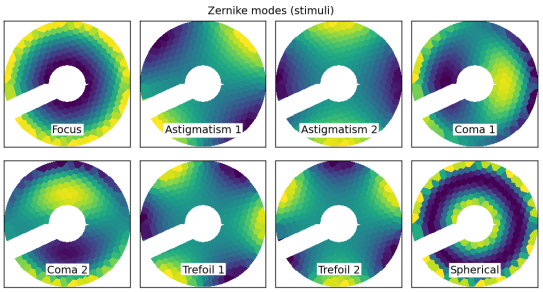


Fig. 3. The eight Zernike modes controlled by the SCExAO implementation of the APF-WFS. Sorted by a Noll index, these modes are (from left to right and top to bottom): Z4: focus, Z5: oblique astigmatism, Z6: vertical astigmatism, Z7: horizontal coma, Z8: vertical coma, Z9: vertical trefoil, Z10: oblique trefoil and Z11: primary spherical.

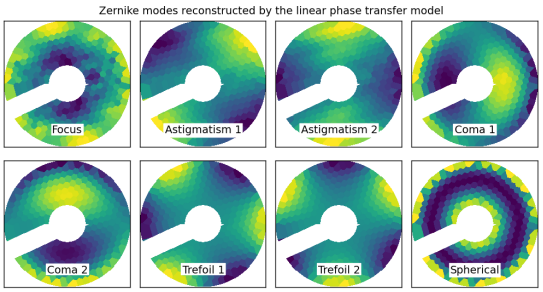


Fig. 4. The eight Zernike modes reconstructed by the linear model when 150 out of the 291 modes are kept in the computation of the pseudo-inverse of the phase transfer matrix $\mathbf{A}^+$.

by the analysis. To deploy this method on an actual closed-loop system is not as direct and requires us to take into account the actual properties of the DM, such as the response curve of the actuators, their influence functions, as well as a careful mapping of these DM actuators on the instrument pupil (Blain 2013). While possible, this type of model is very prone to errors and its maintenance is demanding because of small changes in the internal instrument alignment that are induced by temperature drifts or after a telescope slew.

AO systems in operation usually choose to rely on a more pragmatic approach that encapsulates this kind of model in a transparent manner: individual DM actuators or groups of actuators (pre-defined modes) are sequentially excited and the system response is recorded and assembled in a matrix. Filtering of the noisy modes before inversion (using SVD or similar procedures) leads to the obtention of a control matrix that can be used to directly multiply an input vector made of the wavefront sensor raw input data.

This pragmatic approach is the one that was retained for this implementation of APF-WFS. The control software was designed to sense and control eight low-order Zernike modes (Zernike 1934) that correspond to classical optical aberrations: focus, two terms of coma, astigmatism and trefoil, and spherical aberration. Figure 3 shows how these modes map on the discrete pupil model used to describe the instrument.

2.3. Properties of the phase transfer model

To better appreciate the impact of the phase-transfer model, one can look at the effect of the projection in the Fourier plane and then back in the pupil plane using the linear relations of Eq. 2 and Eq. 3. The original phase φ becomes

$$\varphi' = \mathbf{A}^+ \times \mathbf{A} \times \varphi. \quad (4)$$

Depending on the number of modes kept in the determination of the pseudo-inverse $\mathbf{A}^+$ of the phase-tranfer matrix, the reconstruction goes from perfect (if all modes are preserved) to very partial (if few modes are preserved). The modes discarded correspond to low singular values which, for a given level of signal-to-noise in actual data, would result in amplified noise. On SCExAO, for the control of these eight low-order modes, 150 out of the 291 available modes are maintained in the computation of the pseudo inverse. Under these conditions (see Figure 4), the reconstruction appears satisfactory, and confirms that the technique can indeed be used to control low-order modes, assuming that the linear model holds (and that aberrations are small).

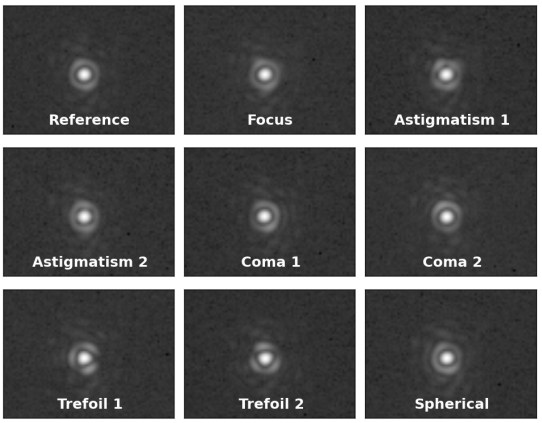


Fig. 5. Calibration data for the APF-WFS acquired by the SCExAO science camera. Top left: the reference PSF, acquired with the system in its starting state. From left to right and top to bottom: the PSF after the corresponding Zernike mode has been applied. A non-linear scale is used to better show the impact of a 30 nm RMS DM modulation.

2.4. Calibration

The calibration procedure for this implementation of the APF-WFS follows the linear control framework. After the asymmetric mask has been inserted, one image labeled as reference is acquired, followed by a sequence of images acquired after a Zernike mode of appropriate amplitude has been applied to the DM. Figure 4 features one of these calibration data-set, which was acquired with the focal camera of SCExAO on its internal calibration source (super-continuum laser) using a standard H-band filter for a 30 nm RMS deformation of the DM. We note that this displacement actually translates into a 60 nm wavefront amplitude modulation (the DM being a reflective system).

Each image is recentered and windowed by a super-gaussian function that filters out high-spatial frequencies, and reduces the impact of detector readout noise. It is then Fourier-transformed and the Fourier-phase is extracted according to the sampling

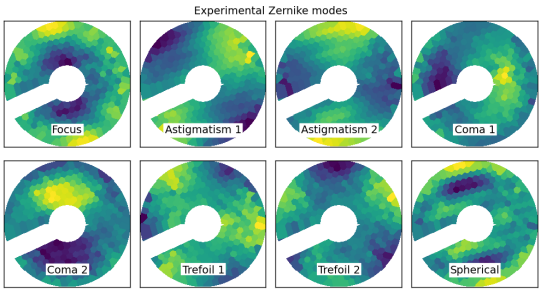


Fig. 6. Experimentally recovered Zernike modes. Save for the spherical aberration, it can be seen that the modes extracted from the analysis of the images of Figure 5 reproduce the features expected after looking at the theoretically reconstructed modes presented in Figure 4.

model featured in the right hand side of Figure 2 to populate a vector Ψ.

To each Zernike mode, a Fourier-phase vector Φ_i can be associated after subtracting the phase Ψ_{ref} measured in the initial or reference state:

$$\Phi_i = \Psi_i - \Psi_{ref}. \quad (5)$$

The wavefront associated with this Fourier-phase signature can be recovered using the pseudo-inverse $\mathbf{A}^+$ previously computed and applying Eq. 3. This wavefront in radians can, in turn, be converted into a DM displacement map (in microns) after being multiplied by the proper $\lambda/4\pi$ scaling factor; where λ is the wavelength expressed in microns and 4π contains the $\times 2$ factor caused by the reflection. Figure 6 features an example of experimentally recovered modes. One will observe that the reconstruction from the Fourier analysis of actual images appears visually satisfactory. Differences in the reconstruction with the modes plotted in Figure 4 can be attributed to imperfections of the pupil discretization model (see Figure 4 of Martinache (2013)), combined with practical subtleties like the fact that the dynamical range on the camera is limited and that classical noises, such as photon and readout, do apply.

To complete this description of the general aspect of the modes with a quantitative estimate of the quality of the reconstruction, Figure 7 directly compares the experimental reconstruction E to the theoretical Zernike modulation T, by plotting the local deduced DM displacement against its predicted value. One can confirm that all modes are not equally reproduced by the analysis and that, for this experimental setup, the sensor is most sensitive to astigmatism (for which the correlation coefficient is the strongest) and not particularly suited to the sensing of spherical aberration. Table 1 accompanies this figure and provides the value of the Pearson product-moment correlation coefficient for all modes:

$$r = \frac{\mathrm{cov}(E,T)}{\sigma_E \sigma_T}, . \quad (6)$$

Table 1 also shows that, with a correlation coefficient ~0.5, spherical aberration is significantly less well reconstructed than the other modes that exhibit a correlation coefficient >0.8. The specificity of the response to spherical aberration is, however, not an intrinsic limit of the sensing approach and can, in fact,

Zernike mode	quality
Z4 (focus)	0.855
Z5 (astigmatism 1)	0.949
Z6 (astigmatism 2)	0.925
Z7 (coma 1)	0.820
Z8 (coma 2)	0.862
Z9 (trefoil 1)	0.827
Z10 (trefoil 2)	0.854
Z11 (spherical)	0.522

Table 1. Pearson product-moment correlation coefficient of the experimentally reconstructed mode E with their theoretical counterpart T.

simply be explained by the 2D geometry of this aberration and how it fits within the footprint of the Subaru Telescope and its large (~30%) central obstruction. By looking for instance at Figure 3, we can see that the donut shape of the spherical aberration results in a pretty uniform distribution of the phase, which only varies near the inner and outer edges of the pupil. The basis of Zernike polynomials is defined for a complete circular aperture: quoted amplitudes correspond to a given wavefront RMS over the entire circular aperture. For the spherical aberration, the presence of the central obstruction naturally filters out a lot of the effect of the spherical aberration. This is further confirmed after a close examination of scatter plots of Figure 7: whereas all Zernike stimuli (along the horizontal axis of the plots) have the same amplitude, we can see that the resulting range of local DM displacement (corresponding to the horizontal spread of the data points) is appreciably shorter for the spherical aberration than it is for the other modes.

These experimentally obtained pupil-phase modes φ_i are stored in a 8×291 matrix $\mathbf{Z}$, referred to as the response matrix. In practice, unless the DM registration were to change in a dramatic manner, the calibration is quite robust: a response matrix acquired using the internal calibration source can be very well used during on-sky observations if the filter remains unchanged, and if the change of exposure time does not result in a saturated PSF core (see the discussion in Section 4).

On SCExAO, the acquisition of this response matrix only takes a few seconds, so it can easily be repeated if neessary after acquisition of a new target. In practice, it seems a response matrix that is acquired using the stable internal calibration source provides the best results.

2.5. Closed-loop operation

Just as during the calibration, focal-plane images acquired on-sky with the asymmetric mask are dark-subtracted, recentered, and windowed by a super-Gaussian function before being Fourier-transformed. After extraction of the Fourier phase, a wavefront is produced and directly projected onto the basis of modes (without subtracting the reference), to find the coefficients associated with all eight Zernike components. If the current wavefront sensor signal is φ, the instant Zernike coefficients (α) are the solution of $\mathbf{Z} \cdot \alpha = \varphi$. The least square solution $(\hat{\alpha}_i)_{i=4}^{11}$ of this system is the solution to

$$\mathbf{Z^T Z} \cdot \hat{\alpha} = \mathbf{Z^T} \varphi. \quad (7)$$

The solution $\hat{\alpha}$ to this well-behaved system of equations is used as an input for a control loop algorithm. The loop in operation on SCExAO implements a simple proportional controller,

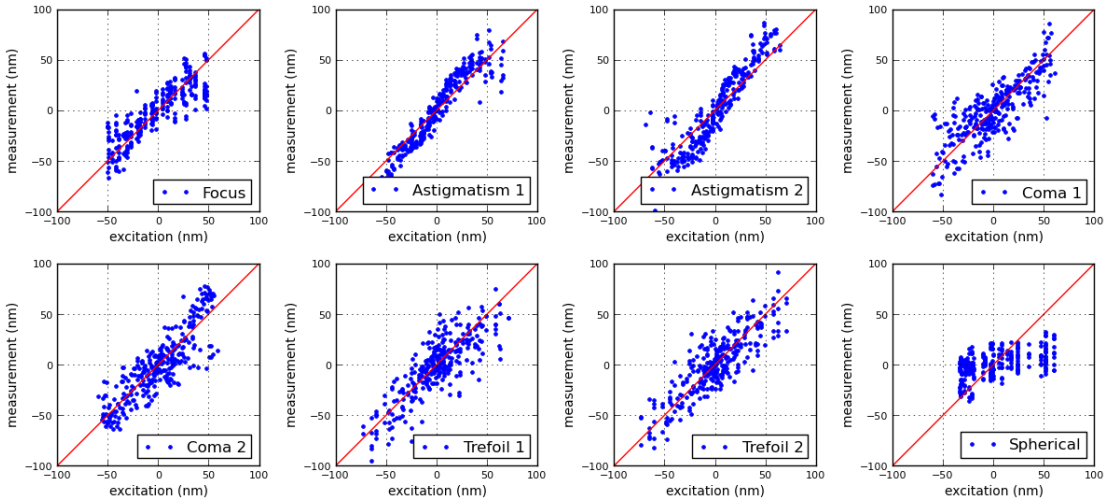


Fig. 7. Comparative quality of the modal reconstruction: for each mode, the local value of the DM displacement (in nm) for the theoretical Zernike mode of predefined amplitude (here 30 nm, labeled as the 'excitation') is plotted against its experimental measurement.

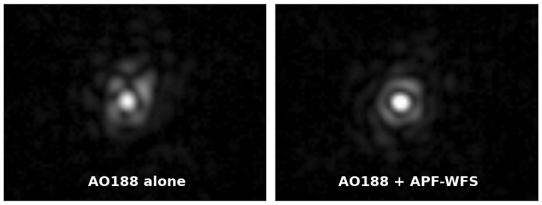


Fig. 8. Illustration of the impact of the APF-WFS. Left: 0.5 ms PSF acquired by SCExAO's internal science camera after the upstream AO loop has been closed. Right: identical exposure acquired 30 seconds after the APF-WFS loop has been closed. Despite residual imperfections owing to dynamic changes, the PSF quality is obviously improved.

with a gain common to all Zernike modes with value contained between 0.05 and 0.3, depending on the overall stability of the wavefront provided by the upstream AO. When looking at the internal calibration source, we can reliably use the highest gain. Since, for now, it is only used for a very short time (typically $\sim$ 15 seconds), at the time of target acquisition to flatten the static component of the wavefront, the current implementation of the algorithm proves satisfactory. Once the non-common path error is accounted for, the asymmetric mask is taken out of the optical path and the system is ready for observation using the full pupil of the telescope.

3. Performance

3.1. On-sky demonstration

The technique was successfully deployed and proved to be effective at reducing the non-common path error during on-sky observations behind Subaru Telescope's facility AO system AO188 (Minowa et al. 2010). Figure 8 illustrates the impact of the approach, with two 500 μs exposures of the target (Altair) acquired by SCExAO's internal science camera on UT 2015-10-30.

The first image shows the PSF after the AO188 loop has been closed on the target: although it features a well-defined diffraction core, the PSF clearly exhibits some static aberrations that can be attributed to the non-common path error between AO188 and SCExAO's focal plane. The second image shows the PSF about 30 seconds after the APF-WFS loop has been closed. The gain in Strehl is low (of the order of 5 %), but the PSF is improved at where it matters most for high contrast imaging and no longer features any obvious low-order aberration signature. Residual inhomogeneity of the first diffraction ring can be attributed to a combination of instantaneous AO residuals combined with the effect of the asymmetric arm.

SCExAO's internal science detector is a fast but low-sensitivity detector that can acquire images at up to 170 Hz full-frame rate whose specifications are given by Jovanovic et al. (2015b). APF-WFS seems to exhibit sufficient sensitivity to be used in a fast closed-loop that could very well track low-order aberrations with frequencies up to a tenth of the camera frame rate.

At the moment, the goal of the loop is to calibrate the quasi-static non-common path error. The control software keeps a rolling average of the 20 last wavefront estimations, and corrects for the average of these estimations at each iteration, thus filtering vibration-induced fast varying component. Combined with the acquisition, the (non-optimized) computation of the wavefront makes the loop run at a frequency of $\sim$8 Hz.

3.2. Cross-talk

Zernike polynomials (Zernike 1934) form a convenient basis to describe a wavefront within a circular aperture: designed to form an orthonormal basis, the first terms of the series happen to correspond to classical monochromatic optical aberrations like focus, astigmatism or coma. As previously seen, the presence of a central obstruction in the pupil makes this basis no longer perferctly orthogonal. Substitutes have been proposed (Mahajan 1981) to accomodate for the presence of this rather common

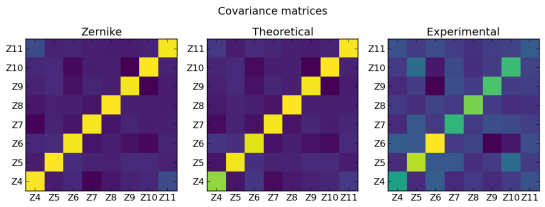


Fig. 9. Control modes inner product matrices characterizing the orthogonality properties of the implementation of the APF-WFS described in this paper on SCExAO. From left to right: 1. the mostly perfectly diagonal case of the Zernike polynomials basis, 2. the seemingly identical inner product matrix for the theoretical modes after reconstrucion by the linear model and filtering by the SVD, and 3. the inner product matrix for the experimentally reconstructed modes. The overall diagonal allure of the latter characterizes the sensor as suited to the control of the low-order Zernike modes. Numerical values for the experimental inner products are provided in Table 2.

	Z4	**Z5**	**Z6**	**Z7**	**Z8**	**Z9**	**Z10**	**Z11**
Z4	0.46	-0.09	0.12	-0.10	-0.04	-0.09	-0.03	0.06
Z5	-0.09	0.87	0.09	0.09	-0.08	-0.03	0.19	-0.03
Z6	0.12	0.09	1.00	-0.04	0.01	-0.25	-0.16	0.05
Z7	-0.10	0.09	-0.04	0.57	-0.04	0.05	0.03	-0.02
Z8	-0.04	-0.08	0.01	-0.04	0.75	-0.03	-0.07	-0.07
Z9	-0.09	-0.03	-0.25	0.05	-0.03	0.65	-0.03	-0.02
Z10	-0.03	0.19	-0.16	0.03	-0.07	-0.03	0.61	0.00
Z11	0.06	-0.03	0.05	-0.02	-0.07	-0.02	0.00	0.11

Table 2. Numerical values for the experimental modes' inner-product

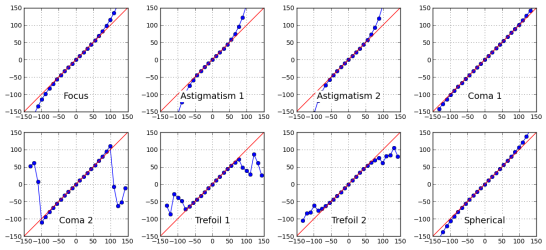


Fig. 10. Experimental response of the APF-WFS obtained on the SCExAO internal (super-continuum) source in the H-band. Each plot features (on the vertical axis) the reponse of the sensor to a Zernike mode of RMS amplitude that varies over a ±150 nm range (units for both axis are in nm). While nearly linear over the entire range for most modes, the sensor only exhibits a significant non-linear behavior for the coma 2, and the two trefoil modes when the DM Zernike amplitude is larger than 80-100 nm. We note that this limit is on the DM surface, which must be doubled if refering to aberrations on the wavefront.

feature of telescopes, but the asymmetric arm required for the wavefront sensing (see Section 1) would also require an adaptation.

Instead of trying to specify a new orthogonal basis perfectly adapted to our case, we have judged it more appropriate to stick to the conventional Zernike basis and verify a posteriori how orthogonal the different modes actually are. Figure 9 does this by plotting the 8×8 matrix of inner products between the eight control modes ($\mathbf{Z^T Z}$) for three cases: the input Zernike polynomials (given in Figure 3), the theoretical reconstruction of the linear model with 150 out of the 291 eigen modes kept in the phase transfer matrix inversion (given in Figure 4), and the experimentally acquired modes (given in Figure 6). An orthogonal basis will result in a perfectly diagonal inner-product matrix, whereas non-orthogonality would become manifest with strong non-diagonal components.

One can therefore observe, looking at the left hand side panel of Figure 9, that the Zernike modes form a satisfactory, nearly orthonormal basis, with a mostly uniform diagonal and with a limited amount of cross terms standing out (except in the case of the 16% cross-correlation between focus (**Z4**) and spherical (**Z11**). For the sake of consistency with the rest of the data presented in the paper, Figure 9 also shows in its central panel, that the modes reconstructed by the linear model reproduce most of these features, although one can observe ∼20% degradation of the relative strength of the focus signal. What we observe here is the effect of the filtering of low singular values in the construction of the pseudo-inverse $\mathbf{A}^+$ as used in Equation 4. With 150 out of the possible 291 modes kept in the construction of the pseudo-inverse, the inner product matrix for the experimentally recovered modes is also mostly diagonal.

Table 2 provides the numerical values for the experimental inner products, also graphically represented in the right hand side panel of Figure 9. Although some of the cross terms are non-neglibible, the terms along the diagonal still dominate, indicating that a control loop that relies on this calibration dataset will reliably converge toward a state that will cancel the non-common path aberration.

3.3. Range of linear response

As mentioned in Section 2.1, the APF-WFS relies on the assumption that an upstream AO correction is provided. The system is expected to deal with with small residual wavefront errors, and the calibration procedure described above, typically employs DM modulation amplitudes of ∼30-50 nm. To determine the amount of aberration the technique is able to deal with, we performed a systematic exploration of the response of the sensor to stimuli of variable amplitude. The instantaneous response of the sensor is projected onto the basis of modes following the procedure outlined in Section 2.5. Figure 10 summarizes the results of this systematic exploration of the response of the sensor, over a ±150 nm range of DM modulation amplitude.

Although not perfectly linear, for **Z4** (focus), **Z5**, **Z6** (astigmatism), **Z7** (coma 1), and **Z11** (spherical), the response remains monotonic over the entire ±150 nm range. For **Z8** (coma 2), **Z9** and **Z10** (trefoil 1 and 2), the response is only monotonic over the ±80 nm modulation range beyond which the sensor cannot be used reliably. The drastic difference of response between **Z7** and **Z8,** which corresponds with the same type of aberration (i.e. coma), can be explained by the azimuth of the asymmetric arm in the pupil stop, oriented such that it almost entirely masks out one of the two antisymmetric bumps that are characteristic of this aberration (see for instance the top left panel of Figure 3).

A strong non-linearity of the response is experienced when the pupil-phase peak-to-valley (P2V) wavefront becomes larger than 2π (which results in a phase wrap). The presence of the asymmetric stop at its current azimuth essentially divides the P2V by a factor of two in the case of **Z7** (i.e. **coma 1**), hence making the sensor able to handle twice as much coma along the horizontal axis than along the vertical axis. We note that the same effect (to a lesser extent) can also be observed when comparing **Z9** and **Z10**. We can nevertheless conclude that, in the H-band under normal operating conditions, the sensor is able to operate

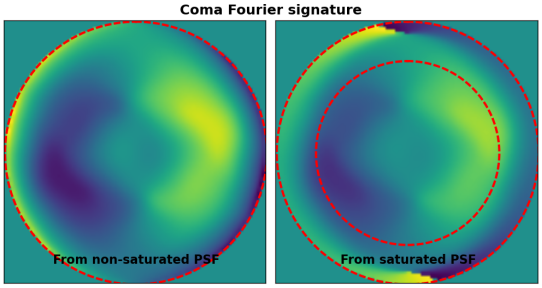


Fig. 11. Comparison of the UV-phase signature of the same amount of aberration (here coma) for a non-saturated PSF (on the left) and a saturated one (on the right). The saturation primarily affects the higher spatial frequencies of the image, corresponding to the outermost parts of the Fourier plane. In the inner region of the saturated case (within the highlighted smaller circle), the Fourier-phase signature of the aberration is similar to its non-saturated counterpart.

linearly as long as the RMS error on either mode is less than 200 nm on the wavefront.

4. Wavefront sensing from a saturated PSF

A well AO-corrected PSF is a highly contrasted object. The proper simultaneous sampling of the core of the PSF and its diffraction rings therefore requires a detector with a large dynamical range, which is rarely compatible with a fast readout. The SCExAO internal science camera has an effective dynamical range of ∼10000 counts so that, in practice, an exposure time has to be chosen that either gives access to a non-saturated PSF core (the normal operating mode of the wavefront sensing approach described thus far) or over-expose the PSF core to better see the fainter diffraction structures that surround it.

In its general form, the linear model of Eq. 2 only holds when working on non-saturated images that otherwise result in a non-translation invariant PSF. Pixels that are saturated by the bright core of the PSF can be treated as zeros, so that the effect of saturation can be modeled as a multiplication by a top-hat function that cuts off anything higher than a level imposed by the characteristics of the detector. This multiplication in the image space results in a convolution by an Airy-like function whose characteristic size is inversely proportional to the size of the saturated part of the PSF.

The effect of this convolution is expected to be most prominent in the outermost region of the Fourier plane where the phase will experience a change of sign. Figure 11 illustrates this effect by comparing the Fourier-phase signature of a specific aberration for a non-saturated PSF to that of a mildly saturated one. We can see that, while the outermost part of the Fourier-phase is obviously affected by the saturation, the innermost part of the PSF (highlighted by the smaller dashed circle) does resemble the original non-saturated case.

Figure 12 further makes this obvious by representing in a 1D plot the values of the Fourier-phase of the saturated image against the Fourier-phase of the non-saturated image. Whereas considered as a whole, the Fourier-phase of the saturated data appears as non-usable (the blue points are widely scattered), the inner part of this same saturated data set is strongly correlated (the red points) with the non-saturated data, suggesting that some of the wavefront information can be recovered in the saturated case,

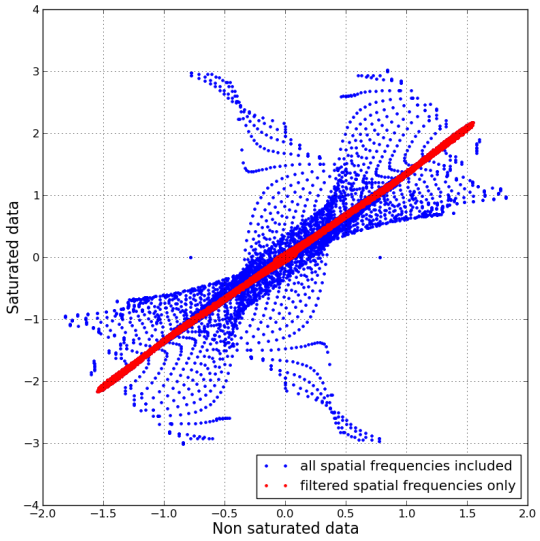


Fig. 12. 1D comparison of the phase (in radians) extracted from the Fourier plane, featured in Figure 11, in the non-saturated case (along the horizontal axis) and in the saturated case (along the vertical axis). The blue points include the data at all spatial frequencies while the red ones correspond to the inner part of the Fourier plane only. The strong correlation observed in the latter case suggests that some of the wavefront information can be recovered from the analysis of mildly saturated data.

assuming that we can filter out the information coming from the largest baselines.

To account for this filtering, the Fourier-phase model can be modified, and the parts of the phase transfer matrix $\mathbf{A}$ can be discarded along with the parts of the Fourier-phase vector Φ that are filtered out. The model we tested only preserves baselines that are 70% or less than the longest baseline in the model, which corresponds to the area inscribed within the inner circle plotted in the right panel of Figure 11. Out of the 675 original distinct UV-phase samples, 330 remain with this configuration, which is still of the order of the number of modes necessary to recover the full theoretical pupil phase information (291 modes).

For the computation of the pseudo-inverse $\mathbf{A}^+$ of this new system, 50 modes are kept. With less constraints from the UV-plane, the Zernike modes are less well reconstructed, but are nevertheless recognizable, as shown in Figure 13.

The calibration procedure introduced in Section 2.4 can be repeated with this new model, after which the APF-WFS is effectively able to operate in closed-loop from the analysis of saturated data, albeit with lower performance. To see how this saturation affects the sensor, the study presented in Section 3.3 was repeated in this new operating mode. The outcome of this study is presented in Figure 14.

In this peculiar saturated mode, the sensor is able to operate linearly over a limited range of aberrations. As previously observed (see Section 3.3), it is the modes whose geometry features localized bumps, such as coma of trefoil, that first limit the range of aberration that the APF-WFS can account for. The APF-WFS can nevertheless operate on images whose core is saturated, over a 100 nm wavefront RMS range that is roughly one half of what it can achieve in the non-saturated operating regime: the general

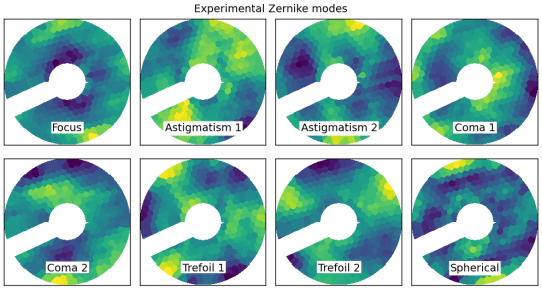


Fig. 13. Experimentally recovered Zernike modes after discarding the phase associated with the longest baselines, which are affected by saturation effects. This new series of experimental modes should be compared to the non-saturated case presented in Figure 6.

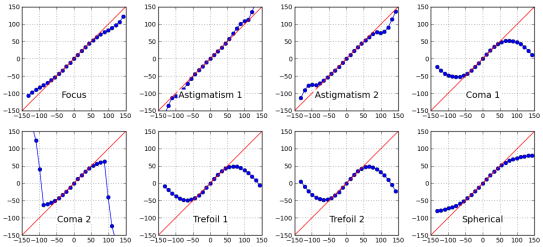


Fig. 14. Study of the response of the sensor in the saturated case to Zernike modes of varying amplitude. This series is to be compared to the case presented in Figure 10 in the non-saturated case.

APF-WFS approach is more robust than at first expected and can be used despite the less than ideal conditions.

5. Conclusion

Following on from the conceptual study proposed by Martinache (2013), this paper described the implementation of the asymmetric pupil Fourier wavefront sensor as one of the wavefront control loops of the SCExAO instrument. This approach has proven capable of repeatedly accounting for the non-common path error that affects the instrument after a new telescope pointing and provides an updated zero-point for the upstream pyramid wavefront sensor currently implemented inside SCExAO.

A surprisingly simple asymmetric hard stop mask introduced in the pupil of a diffraction-limited imaging instrument is thus proving to be a powerful diagnostic tool for the control of the non-common path aberrations. The reported capture range of the technique is currently limited to a fraction of a wave (RMS $\sim \lambda/8$). A combination of filters of decreasing wavelengths would provide a direct way to tolerate a cruder starting point. We are currently exploring the potential of an updated algorithm that simultaneously exploits the information sampled at multiple wavelengths to extend the capture range even more, this time within the coherence length. We note that other approaches using combinations of non-redundant aperture masks (Cheetham et al. 2012, 2014) also rely on this idea to extend their capture range.

The asymmetry results in slight cosmetic degradations of the PSF. While this does impact a coronagraphic instrument, it can be tolerated in a general purpose AO-corrected imaging instrument. A very interesting feature of this image-based wavefront sensing approach is that, if multiple sources are available in a given field, the APF-WFS algorithm can be used on all sources simultaneously. Depending on the complexity of the field, the same asymmetric mask, combined with the analysis of multiple sources in one image, can be used for multi-reference wavefront sensing, opening the way to a full 3D reconstruction of the wavefront from the analysis of a single focal plane image. This very property can also be put to use on artificially introduced incoherent replicas of an on-axis PSF of tunable intensity, as described by Jovanovic et al. (2015a), thus making the use of the technique compatible with that of a coronograph that otherwise destroys the interferometric reference required for a sensible Fourier-analysis of the image, as described in this work.

The use of this wavefront control technique extends well beyond the control of low-order modes on SCExAO: this paper provides experimental evidence that the technique is actually effective where the theory predicts it should be. In an exposure that simultaneously features an unsaturated PSF core and the diffraction features at large separation with sufficient SNR, APF-WFS can be used to control an arbitrary number of modes, as was shown in the concept paper. The APF-WFS can, in fact, be easily applied in a wide variety of wavefront sensing contexts, for ground- as well space-borne telescopes, and with a pupil that can be continuous, segmented, or even sparse. APF-WFS is powerful because it measures the wavefront where it really matters, at the level of the science detector. Given its low impact on the instrument hardware, it is an option that should be given some consideration as part of any high contrast imaging instrument with wavefront control capability.

References

Blain, C. 2013, PhD thesis, University of Victoria (Canada
Bordé, P. J. & Traub, W. A. 2006, ApJ, 638, 488
Cady, E., Baranec, C., Beichman, C., Brenner, D., Burruss, R., Crepp, J., Dekany, R., Hale, D., Hillenbrand, L., Hinkley, S., Ligon, E. R., Lockhart, T., Oppenheimer, B., Parry, I., Pueyo, L., Rice, E., Roberts, L. C., Roberts, J., Shao, M., Sivaramakrishnan, A., Soummer, R., Tang, H., Truong, T., Vasisht, G., Vescelus, F., Wallace, J. K., Zhai, C., & Zimmerman, N. 2013, in Proc. SPIE, Vol. 8864, Techniques and Instrumentation for Detection of Exoplanets VI, 88640K
Cheetham, A., Cvetojevic, N., Norris, B., Sivaramakrishnan, A., & Tuthill, P. 2014, Optics Express, 22, 12924
Cheetham, A. C., Tuthill, P. G., Sivaramakrishnan, A., & Lloyd, J. P. 2012, Optics Express, 20, 29457
Give'On, A. 2009, in Proc. SPIE, Vol. 7440, Techniques and Instrumentation for Detection of Exoplanets IV, 74400D
Jennison, R. C. 1958, MNRAS, 118, 276
Jovanovic, N., Guyon, O., Martinache, F., Pathak, P., Hagelberg, J., & Kudo, T. 2015a, ApJ, 813, L24
Jovanovic, N., Martinache, F., Guyon, O., Clergeon, C., Singh, G., Kudo, T., Garrel, V., Newman, K., Doughty, D., Lozi, J., Males, J., Minowa, Y., Hayano, Y., Takato, N., Morino, J., Kuhn, J., Serabyn, E., Norris, B., Tuthill, P., Schworer, G., Stewart, P., Close, L., Huby, E., Perrin, G., Lacour, S., Gauchet, L., Vievard, S., Murakami, N., Oshiyama, F., Baba, N., Matsuo, T., Nishikawa, J., Tamura, M., Lai, O., Marchis, F., Duchene, G., Kotani, T., & Woillez, J. 2015b, PASP, 127, 890
Kraus, A. L. & Ireland, M. J. 2012, ApJ, 745, 5
Mahajan, V. N. 1981, Journal of the Optical Society of America (1917-1983), 71, 75
Martinache, F. 2010, ApJ, 724, 464
—. 2013, PASP, 125, 422
Martinache, F., Guyon, O., Jovanovic, N., Clergeon, C., Singh, G., Kudo, T., Currie, T., Thalmann, C., McElwain, M., & Tamura, M. 2014, PASP, 126, 565
Minowa, Y., Hayano, Y., Oya, S., Watanabe, M., Hattori, M., Guyon, O., Egner, S., Saito, Y., Ito, M., Takami, H., Garrel, V., Colley, S., Golota, T., & Iye, M. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7736, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Pope, B., Cvetojevic, N., Cheetham, A., Martinache, F., Norris, B., & Tuthill, P. 2014, MNRAS, 440, 125
Sallum, S., Follette, K. B., Eisner, J. A., Close, L. M., Hinz, P., Kratter, K., Males, J., Skemer, A., Macintosh, B., Tuthill, P., Bailey, V., Defrère, D., Morzinski, K., Rodigas, T., Spalding, E., Vaz, A., & Weinberger, A. J. 2015, Nature, 527, 342
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555
Zernike, v. F. 1934, Physica, 1, 689

Mesure de front d'onde polychromatique

Les propriétés de la mesure de front d'onde présentée dans la section précédente montrent que le domaine de capture est limitée à de faibles aberrations : l'amplitude totale des aberrations doit être de l'ordre de 1 radian ou moins, ce qui se traduit par des erreurs sur le front d'onde de l'ordre de $\lambda/4$. Au delà, on a pu voir que la méthode est victime d'un effet de repliement de la phase : la mesure devient dégénérée.

On peut lever cette dégénérescence en exploitant les propriétés de dépendence chromatique de la phase instrumentale. Dans un milieu caractérisé par un indice de réfraction n, un chemin optique de longueur p (pour piston, exprimée en mètres) induit un retard de phase chromatique $\varphi(\lambda)$ suivant une loi simple :

$$\varphi(\lambda) = 2\pi n(\lambda) p/\lambda. \tag{63}$$

Pour simplifier les équations, on supposera que l'indice de réfraction vaut un sur le domaine de longueur d'onde considéré, soit $n(\lambda) = 1$.

Dans le cas général d'une ouverture redondante, un signal d'amplitude complexe s extrait du plan de Fourier pour une coordonnée (u, v) donnée de redondance r peut être explicité comme étant la somme de r phaseurs dont les différences de phases sont induites par des pistons aléatoires de valeurs inconnues (voir à nouveau la figure 15 illustrant le cas d'une base de redondance 5). Si on introduit l'inverse de la longueur d'onde : $\xi = 1/\lambda$, la dépendance chromatique de cette amplitude complexe s'écrit de la façon suivante :

$$s(\xi) = \sum_{k=1}^{r} \exp\left(i2\pi p_k \xi\right). \tag{64}$$

La transformée de Fourier de ce signal s'impose rapidement comme la meilleure façon d'interpréter son contenu :

$$\begin{aligned} S_j(p) &= \int_{-\infty}^{+\infty} s_j(\xi) \exp\left(-i2\pi\xi p\right) \mathrm{d}\xi & (65)\\ &= \sum_{k=1}^{r} \delta(p - p_k). & (66) \end{aligned}$$

où δ représente une distribution de Dirac. Le résultat de cette transformée de Fourier est une collection de pics de Dirac dont la position est proportionnelle aux différentes valeurs de pistons $(p_k)_{k=1}^{r}$.

Cette analyse est loin d'être nouvelle [87] : ce principe est ainsi routinement exploité, pour traquer les franges d'un interféromètre longue base à deux télescopes depuis maintenant deux décennies. Lorsqu'un interférogramme est dispersé, le domaine de capture de cette technique n'est plus limité à une fraction de la longueur d'onde $\sim \lambda/4$) mais étendu à une fraction de la longueur de cohérence $\Lambda_c = R \times \lambda$,

87. L. Koechlin, P. R. Lawson, D. Mourard, A. Blazit, D. Bonneau, F. Morand, P. Stee, I. Tallon-Bosc, and F. Vakili. Dispersed fringe tracking with the multi-r_o apertures of the Grand Interferometre a 2 Telescopes. **Appl. Opt.**, 35:3002–3009, June 1996

comme constaté dans un des articles fondateurs de l'interférométrie optique [88] :

"... the finding of the fringes is notably facilitated by a direct-vision prism placed in front of the eyepiece, which permits observation of interference bands with a path-difference of several hundred waves."

Un des premiers résultats de ma thèse de doctorat avait été de montrer que la méthode de cophasage par analyse des franges dispersées pouvait être généralisée au cas des ouvertures arbitrairement complexes [89]. La difficulté connue de ce cas de figure est que même si l'on peut effectivement directement mesurer les r valeurs de pistons des paires de sous-ouvertures contribuant à la même position du plan (u, v), on ne peut pas directement attribuer une valeur de piston donnée à une paire de sous-ouvertures particulière.

88. A. A. Michelson and F. G. Pease. Measurement of the diameter of alpha Orionis with the interferometer. **ApJ**, 53:249–259, May 1921

89. F. Martinache. Global wavefront sensing for interferometers and mosaic telescopes: the dispersed speckles principle. *Journal of Optics A: Pure and Applied Optics*, 6:216–220, Feb. 2004

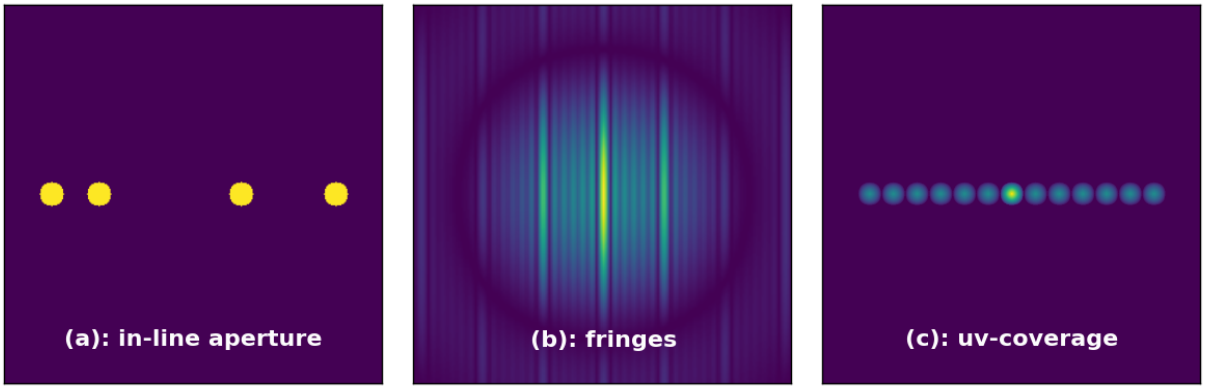


FIGURE 84: Configuration non-redondante à quatre sous-ouvertures à auto-corrélation compacte, utilisée dans le recombineur interférométrique MATISSE. La figure montre (a) la pupille, (b) des franges monochromatiques et (c) le module de leur transformée de Fourier. Au delà de quatre ouvertures, une ouverture à auto-corrélation compacte n'est plus possible.

L'attribution des pistons à des paires de sous-ouvertures, c'est à dire l'inversion qui permet de passer de la phase dans le plan de Fourier à la phase dans la pupille est un problème dégénéré lorsque les pupilles sont redondantes : c'est ce qui motive le design de recombineurs interférométriques qui remappent la pupille de façon non-redondante (voir l'exemple de la figure 84). Bien que relativement facile à mettre en oeuvre dans le cas d'un faible nombre d'ouvertures (comme les quatre faisceaux du VLTI dans le cas de MATISSE [90]), la condition de non-redondance stricte devient très contraignante à mettre en oeuvre pour une configuration faite d'un grand nombre de sous-ouvertures : le critère de Nyquist devant être vérifié par les franges les plus fines, les franges les plus épaisses sont sur-échantillonnées. Une telle recombinaison requiert un très grand nombre de pixels et est par conséquent plus susceptible au bruit de détecteurs.

UNE INTERPRÉTATION COLLECTIVE des informations collectées pour toutes les coordonnées accessibles du plan (u, v) permet de résoudre la question de l'attribution des pistons d'une pupille redondante. On dispose en effet du modèle linéaire qui a servi de base à la construction des kernels et des phases-propres : l'information contenue dans

90. B. Lopez, S. Lagarde, W. Jaffe, and collaborators. An Overview of the MATISSE Instrument - Science, Concept and Current Status. *The Messenger*, 157:5–12, Sept. 2014

l'opérateur linéaire **A** reste valide quelle que soit la distribution des pistons. Si l'analyse précédente permet de récupérer les valeurs individuelles des pistons p_k, il est possible de les combiner pour former un terme de phase pour chaque composante du plan (u, v) qui peut servir d'entrée au modèle de transfert de phase :

$$\phi^j = \phi_0^j + \frac{2\pi}{R} \sum_{k=1}^{R} p_k. \tag{67}$$

Lorsque le vecteur des M composantes de la phase de Fourier $(\phi_j)_{j=1}^{M}$ est ainsi reconstitué, le modèle linéaire s'applique à nouveau, cette fois ci, sans faire appel à l'approximation des faibles phases. Comme dans le cas précédent, si la pupille, offre une asymétrie, la matrice **A** devient inversible et chaque valeur de piston peut être assignée à une paire de sous-ouverture. Même si la matrice n'est pas inversible, il est toujours possible de former des kernels qui sont à nouveaux, des quantités insensibles aux pistons.

Cette méthodologie a pour le moment été seulement décrite dans un article soumis à SPIE [91] (incluse à la suite de cette section) et n'a pas encore fait l'objet d'une publication dans un journal à comité de lecture : c'est un des chantiers du projet KERNEL.

91. F. Martinache. Spectrally dispersed Fourier-phase analysis for redundant apertures. In *Optical and Infrared Interferometry and Imaging V*, volume 9907 of **Proc. SPIE**, page 990712, Aug. 2016

Spectrally Dispersed Fourier-Phase analysis for redundant apertures

Frantz Martinache[a]

[a]Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Parc Valrose, Bât. H. FIZEAU, 06108 Nice

ABSTRACT

One can process images acquired by single telescopes using adaptve optics (AO) in a manner similar to data usually acquired with a non-redundant aperture mask. Because it relies on a linearization of the phase transfer equation that describes the value of the phase in the Fourier-plane, the kernel- and eigen-phase analysis of images corrected by adaptive-optics is intrinsically limited to a regime of aberration where the rms error is less than the wavelength. This paper outlines the properties of an extension of the Fourier-phase analysis method, here applied to a set of spectrally dispersed data, similar to those produced by an intergral field spectrograph. Given sufficient wavelength coverage, the original baseline mapping model can be used to produce novel spectrally dispersed kernel-phase information for any aberration regime.

1. INTRODUCTION

It is possible to formally describe an astronomical image as the result of a convolution product between two functions: one that is representative of the true distribution of intensities of the source pointed at, aka the object, and an instrument function (including the telescope and the atmosphere), which in imaging is often referred to as the point spread function (PSF). This image-object convolution relation:

$$\mathbf{I} = \mathbf{O} \star \mathbf{PSF}, \tag{1}$$

is at the heart of diffraction-limited imaging with telescopes or with interferometers. Important efforts are put into building larger telescopes, using astute optical designs for the relay optics and the instruments, now boosted by advanced techniques like adaptive optics (AO), as means to reduce the contribution of the instrument function, so that a fair share of astrophysics can afford to neglect the contribution of the instrument PSF.

High angular resolution imaging stands out among general imaging in astrophysics because the images it deals with are dominated by diffraction effects. To conclude about the nature of a structure present in an image requires to be able to discriminate the contributions of the two terms of Eq. 1, that is to deconvolve the image. This general deconvolution problem is however an ill-posed problem with more unknowns than constraints.

The framework described in this communication proposes to process the Fourier counterpart of images. In a long baseline interferometry scenario, one location in the Fourier plane corresponds to a unique and well identified interferometric baseline. In this more general (potentially redundant) scenario, a given Fourier location encodes the combined contribution of a potentially large number of baselines, this time all contained within the geometric pupil of the single telescope used at the time of the acquisition. The extraction of pertinent information is made possible thanks to a simple model that keeps track of the mapping of interferometric baselines into spatial frequencies of the data acquired by an instrument.

In previous publications this baseline mapping model was put to use for two related, albeit orthogonal, applications. It was first showed[1] that it is possible to extract self-calibrating quantities from diffraction limited images that are robust against residual wavefront aberrations. These kernel-phases form a natural extension of the concept of closuse-phase used in optical long baseline interferometry[2] when three or more apertures are available. Unlike the special closure-phase case, kernel-phases can be extracted from data acquired at the focus of any telescope. This initial study was completed by one that explored wavefront metrology application[3] of the

email: frantz.martinache@oca.eu

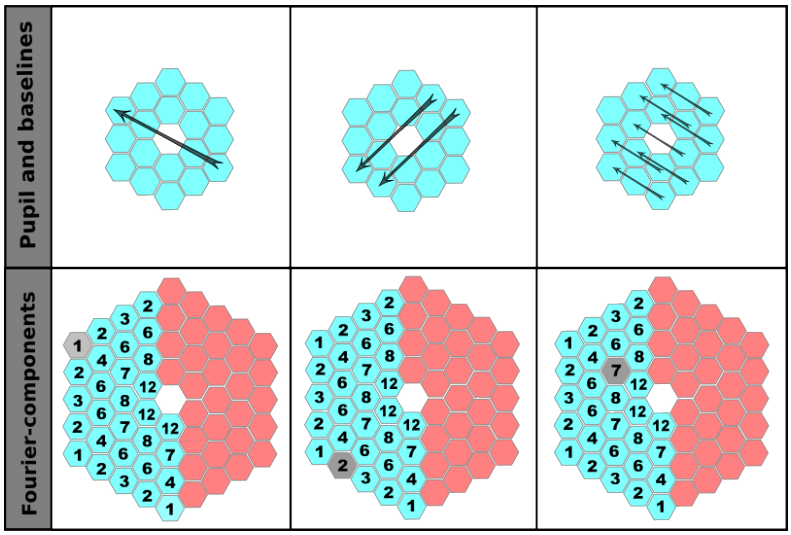


Figure 1. Illustration of the relation between baselines in the JWST pupil and the components found in the Fourier-transform of an image. The top row features the pupil, highlighting several groups of baselines. The top row features a schematic representation of the Fourier-plane: the number displayed inside each Fourier cell is the number of baselines contributing to it. From left to right, three cases are presented: a redundant baseline found twice in the array, a highly redundant baseline found seven times in the array and a non-redundant one. For each case, the upper row shows oriented segments materializing the baselines. The lower row highlights the corresponding region of the Fourier-plane excited by these baselines.

Fourier-phase analysis framework. It is indeed possible to gain access to the knowledge of the wavefront from the analysis of a single focal image, if the pupil presents some level of asymmetry.[?]

The baseline mapping model alone is however not sufficient to enable these applications: these applications only proved successful because of a linearization of the instrumental phase that is only valid when experiencing small residual wavefront aberrations.

This paper aims at further expanding the use of the baseline mapping model beyond this limitation, by looking at spectrally dispersed data, that is for instance available in a diffraction limited integral field spectrograph or in an multi-beam all-in-one spectrally dispersed interferometric combiner. This work is concerned with situations where the previously partially corrected PSF is degraded into a random speckle pattern by large aberrations.

2. THE LINEAR PHASE-TRANSFER MODEL

The presentation will look at the example of a segmented aperture, fairly representative of the pupil of the James Webb Space Telescope (JWST) with $N_A = 18$ hexagonal segments forming a fairly dense (hexagonal grid). This example also assumes that the only degree of freedom per segment is the piston, whereas in practice, one also has to deal with the tip-tilt. With a resulting $N_A - 1 = 17$ degrees of freedom, the example is sufficiently rich to appear generic and representative of more complex problems, but remains simple enough that one can actually keep track of the specifics of the method.

Figure 1 illustrates how for the JWST pupil model, the interferometric baselines formed by the segments in the pupil project into the Fourier-plane. The $N_A = 18$ segments of the pupil form $N_{UV} = 30$ distinct baselines. Some of these baselines appear only once in the array: like one example of featured in Fig. 1. Most baselines can be found multiple times in the array: the general trend is that the shorter the baseline, the more redundant it is. For the example retained, the shortest baseline, constituted by two adjacent segments can be found 12 times in three different directions.

To each data islet in the Fourier transform is associated a complex number, whose value depends on the properties of the observed source: the complex visibility sampled by the corresponding interferometric baseline, perturbed by residual instrumental piston errors. The kernel- and eigen-phase applications presented thus far take advantage of the stabilization of the instrumental phase offered by an AO system: when phase residuals are less than $\sim$1 radian, the expression of the Fourier-phase sampled at the uv-coordinate of index j can be linearized: the phase of the sum of a potentially large number R of phasors can be approximated as the average phase of these R same phasors.

$$\phi^j = \mathrm{Arg}(\sum_{k=1}^{R} v_0^j \exp i(\phi_0^j + \varphi_k)) \approx \phi_0^j + \frac{1}{R}\sum_{k=1}^{R} \varphi_k \tag{2}$$

where v_0^j and ϕ_0^j respectively are the intrinsic target visibility and phase for this baseline of index j. The knowledge of the census of baselines contributing to all Fourier components can be encapsulated in a single linear operator $\mathbf{A}$ such that the phase sampled at all relevant coordinates of the Fourier-plane results from:

$$\Phi = \Phi_0 + \mathbf{A} \cdot \varphi, \tag{3}$$

The geometric properties of the pupil will affect the characteristics of $\mathbf{A}$. In our scenario, this operator is a $N_A \times N_{UV}$ matrix, that describes the way the instrumental phase propagates into spurious phase in the Fourier-plane.

Given that Eq. 3 relise on the linearization of Eq. 2, the direct use of the linear model appears to be restricted to the AO-corrected regime: given the relation between an optical path difference (OPD) p and the phase $\varphi = 2\pi p/\lambda$: the requirement on the instrumental phase being less than $\sim$1 radian translates into wavefront error requirement smaller than $\sim \lambda/4$. When experiencing larger instrumental phase errors, the linearization scheme introduced in eq. 2 is no longer valid. And even in the case of the few possible non-redundant baselines, larger OPD values lead the 2π-periodic phase to wrap up rendering the actual phase degenerate.

3. SPECTRAL PROPERTIES OF THE FOURIER-PHASE

To solve for this difficulty, one can exploit the chromatic properties of instrumental phase: a piston p_k (in meters) on the baseline of index k induces a chromatic (i.e. wavelength-dependent) phase that follows this simple law:

$$\varphi_k(\lambda) = 2\pi p_k/\lambda. \tag{4}$$

Assuming that the index of refraction of the medium used to induce the piston does not vary significantly over the range of wavelength considered.

If this spectral dependence of the Fourier-phase can be monitored over a chosen bandpass at a given spectral resolution, the piston capture range of the technique is no longer bound by the wavelength ($p \ll \lambda$), but by the equivalent coherence length ($p \ll \Lambda_0$), that can be extended by increasing the spectral resolution of the observation.

We will assume for now that over this spectral bandpass, the source does not present a strong chromatic signature - and shall in fact neglect the intrinsic phase Φ_0 entirely for simplicity.

The complex amplitude signal $s_j(\lambda)$ sampled in the Fourier-plane at the relevant uv-coordinate of index j can be explicited as the sum of r phasors induced by unknown pistons. If one introduces the inverse of the wavelength $\xi = 1/\lambda$, one can write the equation for the complex amplitude signal in this convenient form:

$$s_j(\xi) = \sum_{k=1}^{r} \exp\left(i2\pi p_k \xi\right). \tag{5}$$

A Fourier transform of this signal quickly comes to mind as the natural way to process it:

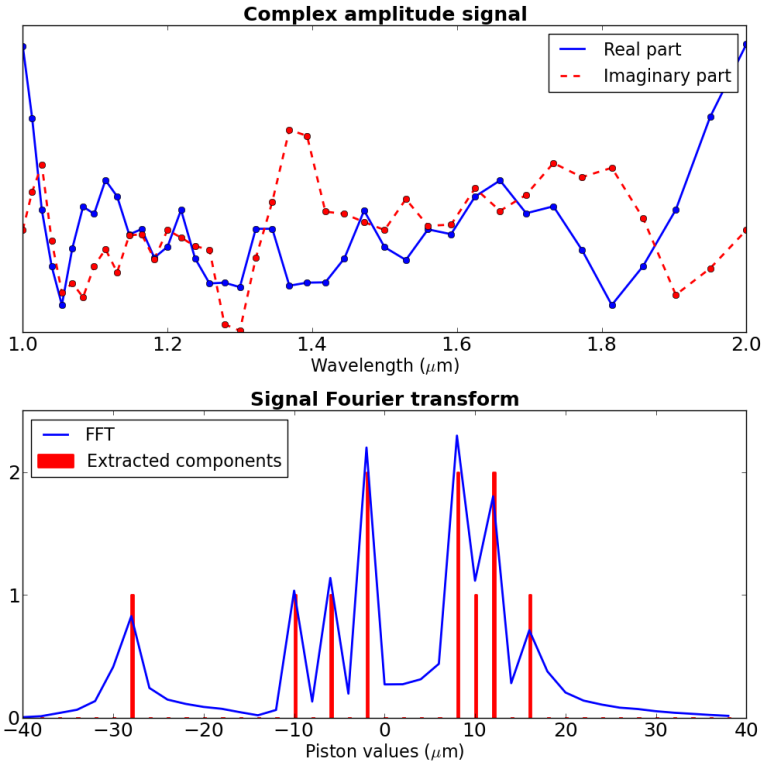


Figure 2. Example of piston extraction for a redundant ($R = 11$) baseline. The fluctuations of the complex amplitude, measured for a given spatial frequency as a function of wavelength (top panel), can be related to the different values of the corresponding 11 pistons, whose values are given by the position of a peak in the Fourier transform of the signal (bottom panel). In this example, 8 distinct peaks are visible. Several baselines indeed feature the same piston, resulting in several Fourier peaks of amplitude twice that of the others. Although the different pistons cannot directly be attributed to baselines, one can add and manipulate them like the linearized phases.

$$S_j(p) = \int_{-\infty}^{+\infty} s_j(\xi) \exp\left(-i2\pi\xi p\right) \mathrm{d}\xi \quad (6)$$

$$= \sum_{k=1}^{r} \delta(p - p_k). \quad (7)$$

where δ stands for the Dirac distribution. The result of this Fourier transform is therefore a collection of such Dirac peaks whose exact location directly gives the values of the different pistons $(p_k)_{k=1}^r$.

This type of analysis is not entirely novel: this very principle has been used[4] for instance for fringe tracking of a long baseline two-telescope interferometer for almost two decades, and the approach can very well be generalized to more complex apertures.[5]

Although in this last reference, it is suggested that an inverse Gerschberg-Saxton approach[6] is possible, the question of how to really solve the question of the redundancy hasn't however been fully addressed. With this approach, one may indeed be able to measure the r values of the pistons of the r baselines contributing to the uv-coordinate, however one is not able to directly attribute the different piston values to their corresponding baselines.

This constraint has been juged severe enough to enfore the rule of non-redundancy in the design of interferometric instruments made for a potentially large number of apertures like FIRST[7,8] and DragonFly.[9] The aperture is systematically remapped in a non-redundant manner, before imaging or spectral cross-dispersion, despite the resulting unefficient use of detector pixels. Although also valid, the situation is not as critical for the current generation of long baseline interferometry instruments that recombine a comparatively small number of apertures: four for MATISSE[10] and GRAVITY[11] at VLTI or six for MIRC at CHARA.[12] Strict adherence to the non-redundancy rule is however not sustainable for recombiners of a large number of apertures.

4. A GLOBAL APPROACH

The attribution of the different piston values is a degenerate problem, but only if the information collected for each uv-coordinate separately. The problem can actually be solved, if all the collected information for all uv-coordinates is processed as a whole.

This can be achieved thanks to the baseline projection model previously used for the contruction of kernel- and eigen-phase. The knowledge of the mapping of the baselines from the pupil to the Fourier-plane, as illustrated in Fig. 1 and encapsulated in the linear operator $\mathbf{A}$ used in Eq. 3 remains after all perferctly valid, regardless of the actual distribution of pistons.

For piston errors comparable or larger than the wavelength, one cannot invoke the linearization of the phase described in Eq. 2. We have however seen that the Fourier analysis of the dispersed Fourier complex amplitude in Eq. 7 gives access to the values of all pistons values $(p_k)_{k=1}^r$. Although they cannot be sorted out individually, these pistons can all be measured and added to each other. The result of this sum (multiplied by 2π), can serve as input to the phase transfer model of Eq. 3, in a manner identical to what has been done so far for the Fourier analysis of AO-corrected images:

$$\phi^j = \phi_0^j + \frac{2\pi}{R} \sum_{k=1}^{R} p_k. \quad (8)$$

In practice, one way to directly get access to the sum of all piston values is simply to compute the center of gravity of Eq. 7. More advanced model-fitting algorithms can also be used to single out the r expected components on the r times redundant baseline. Optimal extraction of this information is the object of ongoing work. At this point, the take home message of this work is that the phase transfer model of Eq. 8 is no longer an approximation but an exact description of the way the instrumental piston propagates into spurious Fourier-phase information.

The baseline mapping model whose use has so far been restricted to well corrected images (Strehl better than ∼40 %) remains perfectly valid regardless of the aberration regime. Depending on the specifics of the pupil, one can therefore again either use this composite observable for science (building up kernel-phases) or for wavefront sensing (recovering the full map of piston in the case of an asymmetric pupil), for a naturally redundant aperture.

Other studies[13] have proposed to exploit the spectral dependance of the piston-induced Fourier-phase information using non-redundant aperture masks to iteratively cophase a segmented telescope like the JWST used as an example. JWST is not equiped with an integral field spectrograph that would otherwise provide a continuous spectral coverage like the one described here. Its instruments however employ a collection of spectral filters that can be used to constrain the problem of phase retrieval even if the starting point features aberrations larger than their individual wavelength as demonstrated in prior work. Given the extension of the baseline mapping outlined in this work,[14] an adapted combination of non-redundant masking interferograms like can be acquired with NIRISS[15] and conventional (unmasked aperture) images in a finite set of spectral filters provides all that is required to first cophase and maintain the cophasing of a segmented telescope like the JWST.

5. CONCLUSION

Up until now, the Fourier analysis of images following the kernel- and eigen-phase framework have relied on the assumption of small wavefront residual errors resulting for instance from the action of an upstream adaptive wavefront control. Using the example of piston on the surface of a segmented aperture, this paper shows that the linear model at the heart of this Fourier approach remains valid and useful regardless of the experienced amplitude of pupil aberrations.

If a sufficiently rich multi-wavelength data set is acquired, one can use the chromatic properties of the instrumental phase, to extract the values of all the pistons along all the baselines that contribute to a single Fourier component. When one deals with these Fourier components separately, the problem appears degenerate. When treating the information collected along all baselines simultaneously, one can raise the degeneracy, using the baseline mapping model originally only used to handle small aberrations. The capture range of the method is no longer limited by the wavelength of operation but by the coherence length that is directly proportional to the spectral resolution of the data.

While the scenario outlined here relies on a continuous spectral coverage such as what is provided by an integral field spectrograph, provided with a sufficiently rich collection of spectral filters, that should be adapted to the range of aberration it is expected to account for, would also suffice to cophase a complex segmented aperture like that of the JWST.

REFERENCES

[1] Martinache, F., "Kernel Phase in Fizeau Interferometry," ApJ **724**, 464–469 (Nov. 2010).

[2] Baldwin, J. E., Haniff, C. A., Mackay, C. D., and Warner, P. J., "Closure phase in high-resolution optical imaging," Nature **320**, 595–597 (Apr. 1986).

[3] Martinache, F., "The Asymmetric Pupil Fourier Wavefront Sensor," PASP **125**, 422–430 (Apr. 2013).

[4] Pope, B., Cvetojevic, N., Cheetham, A., Martinache, F., Norris, B., and Tuthill, P., "A demonstration of wavefront sensing and mirror phasing from the image domain," MNRAS **440**, 125–133 (May 2014).

[5] Koechlin, L., Lawson, P. R., Mourard, D., Blazit, A., Bonneau, D., Morand, F., Stee, P., Tallon-Bosc, I., and Vakili, F., "Dispersed fringe tracking with the multi-r_o apertures of the Grand Interferometre a 2 Telescopes," Appl. Opt. **35**, 3002–3009 (June 1996).

[6] Martinache, F., "Global wavefront sensing for interferometers and mosaic telescopes: the dispersed speckles principle," *Journal of Optics A: Pure and Applied Optics* **6**, 216–220 (Feb. 2004).

[7] Fienup, J. R., "Reconstruction of an object from the modulus of its Fourier transform," *Optics Letters* **3**, 27–29 (July 1978).

[8] Perrin, G., Lacour, S., Woillez, J., and Thiébaut, É., "High dynamic range imaging by pupil single-mode filtering and remapping," MNRAS **373**, 747–751 (Dec. 2006).

[9] Huby, E., Perrin, G., Marchis, F., Lacour, S., Kotani, T., Duchêne, G., Choquet, E., Gates, E. L., Woillez, J. M., Lai, O., Fédou, P., Collin, C., Chapron, F., Arslanyan, V., and Burns, K. J., "FIRST, a fibered aperture masking instrument. I. First on-sky test results," A&A **541**, A55 (May 2012).

[10] Jovanovic, N., Tuthill, P. G., Norris, B., Gross, S., Stewart, P., Charles, N., Lacour, S., Ams, M., Lawrence, J. S., Lehmann, A., Niel, C., Robertson, J. G., Marshall, G. D., Ireland, M., Fuerbach, A., and Withford, M. J., "Starlight demonstration of the Dragonfly instrument: an integrated photonic pupil-remapping interferometer for high-contrast imaging," MNRAS **427**, 806–815 (Nov. 2012).

[11] Lopez, B., Lagarde, S., Jaffe, W., Petrov, R., Schöller, M., Antonelli, P., Beckmann, U., Berio, P., Bettonvil, F., Glindemann, A., Gonzalez, J.-C., Graser, U., Hofmann, K.-H., Millour, F., Robbe-Dubois, S., Venema, L., Wolf, S., Henning, T., Lanz, T., Weigelt, G., Agocs, T., Bailet, C., Bresson, Y., Bristow, P., Dugué, M., Heininger, M., Kroes, G., Laun, W., Lehmitz, M., Neumann, U., Augereau, J.-C., Avila, G., Behrend, J., van Belle, G., Berger, J.-P., van Boekel, R., Bonhomme, S., Bourget, P., Brast, R., Clausse, J.-M., Connot, C., Conzelmann, R., Cruzalèbes, P., Csepany, G., Danchi, W., Delbo, M., Delplancke, F., Dominik, C., van Duin, A., Elswijk, E., Fantei, Y., Finger, G., Gabasch, A., Gay, J., Girard, P., Girault, V., Gitton, P., Glazenborg, A., Gonté, F., Guitton, F., Guniat, S., De Haan, M., Haguenauer, P., Hanenburg, H., Hogerheijde, M., ter Horst, R., Hron, J., Hugues, Y., Hummel, C., Idserda, J., Ives, D., Jakob, G., Jasko, A., Jolley, P., Kiraly, S., Köhler, R., Kragt, J., Kroener, T., Kuindersma, S., Labadie, L., Leinert, C., Le Poole, R., Lizon, J.-L., Lucuix, C., Marcotto, A., Martinache, F., Martinot-Lagarde, G., Mathar, R., Matter, A., Mauclert, N., Mehrgan, L., Meilland, A., Meisenheimer, K., Meisner, J., Mellein, M., Menardi, S., Menut, J.-L., Merand, A., Morel, S., Mosoni, L., Navarro, R., Nussbaum, E., Ottogalli, S., Palsa, R., Panduro, J., Pantin, E., Parra, T., Percheron, I., Duc, T. P., Pott, J.-U., Pozna, E., Przygodda, F., Rabbia, Y., Richichi, A., Rigal, F., Roelfsema, R., Rupprecht, G., Schertl, D., Schmidt, C., Schuhler, N., Schuil, M., Spang, A., Stegmeier, J., Thiam, L., Tromp, N., Vakili, F., Vannier, M., Wagner, K., and Woillez, J., "An Overview of the MATISSE Instrument - Science, Concept and Current Status," *The Messenger* **157**, 5–12 (Sept. 2014).

[12] Blind, N., Eisenhauer, F., Gillessen, S., Kok, Y., Lippa, M., Perrin, G., Dembet, R., Fedou, P., Lacour, S., Perraut, K., Jocou, L., Burtscher, L., Hans, O., Haug, M., Haussmann, F., Huber, S., Janssen, A., Kellner, S., Ott, T., Pfuhl, O., Sturm, E., Weber, J., Wieprecht, E., Amorim, A., Brandner, W., and Straubmeier, C., "GRAVITY: the VLTI 4-beam combiner for narrow-angle astrometry and interferometric imaging," *ArXiv e-prints* (Mar. 2015).

[13] Monnier, J. D., Berger, J.-P., Millan-Gabet, R., and ten Brummelaar, T. A., "The Michigan Infrared Combiner (MIRC): IR imaging with the CHARA Array," in [*New Frontiers in Stellar Interferometry*], Traub, W. A., ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **5491**, 1370 (Oct. 2004).

[14] Cheetham, A. C., Tuthill, P. G., Sivaramakrishnan, A., and Lloyd, J. P., "Fizeau interferometric cophasing of segmented mirrors," *Optics Express* **20**, 29457 (Dec. 2012).

[15] Cheetham, A., Cvetojevic, N., Norris, B., Sivaramakrishnan, A., and Tuthill, P., "Fizeau interferometric cophasing of segmented mirrors: experimental validation," *Optics Express* **22**, 12924 (June 2014).

[16] Sivaramakrishnan, A., Lafrenière, D., Ford, K. E. S., McKernan, B., Cheetham, A., Greenbaum, A. Z., Tuthill, P. G., Lloyd, J. P., Ireland, M. J., Doyon, R., Beaulieu, M., Martel, A., Koekemoer, A., Martinache, F., and Teuben, P., "Non-redundant Aperture Masking Interferometry (AMI) and segment phasing with JWST-NIRISS," in [*Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*], Proc. SPIE **8442**, 84422S (Sept. 2012).

Retour au fringe tracking ?

A ce moment de la lecture du document, j'ai je l'espère réussi à rendre plus apparent le lien formel existant entre les images à la limite de la diffraction et les données interférométriques, qu'elles soit produites par un masque non-redondant dans la pupille d'un unique télescope ou par un instrument au foyer d'un interféromètre longue base.

Il fera sans doute sourire mes collègues spécialistes de l'interférométrie longue base, de m'entendre évoquer ce derniers cas de figure comme un scénario simplifié en comparaison de celui des images au foyer d'un télescope unique. Il ne s'agit pas pour moi de dénigrer la vraie difficulté opérationnelle de l'interférométrie longue base qui dans l'optique demande une infrastructure complexe, mais de mettre en évidence le fait que le nombre de degrés de liberté du problème à résoudre est beaucoup plus petit dans le cas d'un recombineur à quatre faisceaux destiné au VLTI que les milliers de modes de front d'onde à prendre en compte pour un système XAO comme SCExAO.

Du point de vue de l'analyse, le cas d'un *fringe tracker* pour les quatre faisceaux du VLTI est un cas particulier, restreint, d'un analyseur de front d'onde plan focal : tout ce qui peut être fait avec SCExAO peut être transposé ici. A cette fin, dans le cadre du projet KERNEL, un banc optique implémentant un miroir déformable segmenté comme moyen d'action sur le front d'onde et une caméra haute cadence et haute sensibilité placée en plan focal offre un moyen de démonstration complémentaire du travail que je continue de mener avec SCExAO.

La difficulté pour un *fringe tracker* de cette approche plan focal reste le domaine de capture limité. Couvrir de façon continue un domaine spectral comme proposé dans la publication précédente demanderait un spectrographe intégral de champ, dont la réalisation reste délicate, avec un grand détecteur s'accompagnant de contraintes de coût et de vitesse de lecture rendant cette solution peu viable. Un compromis implémentable en laboratoire dans l'année qui vient, consiste en un système capable de fournir quatre interférogrammes dans quatre filtres larges, couvrant toute les bandes J et H. Ces images fournissent suffisamment de contraintes pour contraindre la solution du front d'onde présent en entrée du système, tout en faisant une utilisation efficace de la lumière disponible à la métrologie. La Figure 85 présente un mode de recombinaison 2D non-redondant (donc asymétrique !) compact qui permet de contraindre assez facilement pour toutes les inconnues du problème.

Les valeurs des pistons et tip-tilt différentiels peuvent en effet directement lus dans la transformée de Fourier de l'interférogramme, au premier ordre en phase et au deuxième ordre, en incluant l'intégramité de la visibilité complexe.

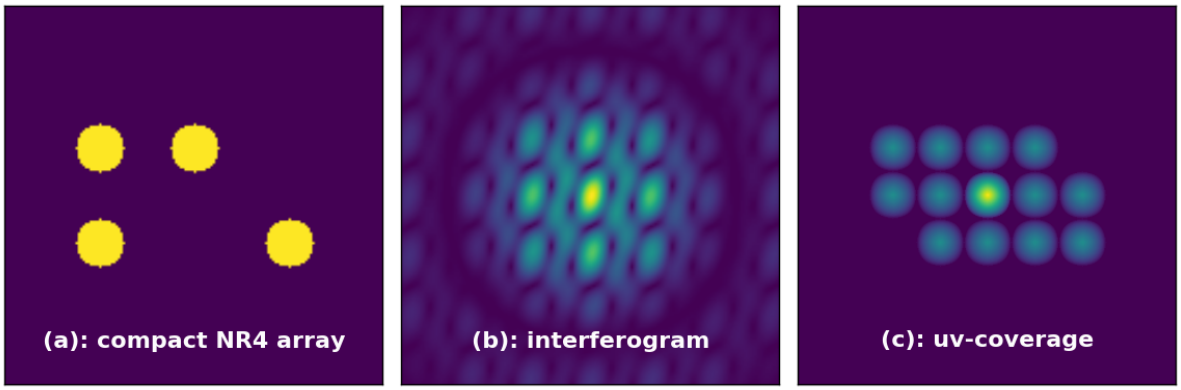


FIGURE 85: Exemple de configuration non-redondante compacte à deux dimensions pour un recombineur 4T semblant adaptée pour un *fringe tracker* offrant en plus du piston, accès au tip-tilt différentiel entre les sous-ouvertures.

ARRAKIS TEACHES THE ATTITUDE OF THE KNIFE — CHOPPING OFF WHAT'S INCOMPLETE AND SAYING : "NOW IT'S COMPLETE BECAUSE IT'S ENDED HERE".

COLLECTED SAYINGS OF MUAD'DIB, FRANK HERBERT, *DUNE*

Conclusion : vers un haut contraste robuste ?

L'application ambitieuse des techniques et méthodes présentées dans ce mémoire reste encore la détection directe de planètes extrasolaires en orbite autour des étoiles du voisinage Solaire. Le paysage scientifique de la thématique des exoplanètes est aujourd'hui beaucoup plus riche qu'à la conclusion de ma thèse de doctorat : on peut maintenant affirmer avec assurance que les planètes sont plus nombreuses que les étoiles dans notre Galaxie, que les planètes rocheuses sont particulièrement fréquentes et la recherche se concentre désormais sur les planètes dites habitables, où les conditions d'apparition de la vie semblent réunies.

Plusieurs instruments d'imagerie haut contraste dédiés à cette recherche ont également été construits et utilisés sur les plus grands télescopes optiques en opération à ce jour. Les campagnes d'imagerie directe n'ont pour le moment découvert qu'une dizaine de planètes et la dernière génération d'instruments n'a que très marginalement contribué à augmenter ce total. Avoir eu cet objectif ambitieux a malgré tout conduit à des systèmes capables de produire des images de qualité exceptionnelle, permettant entre autres d'étudier la structure dynamique des disques protoplanétaires dans l'infrarouge proche et de permettre des observations avec une résolution angulaire encore augmentée dans le visible.

En parlant du travail que j'ai mené sur SCExAO, le deuxième chapitre de ce mémoire a tenté d'expliquer pourquoi les performances des coronographes modernes, théoriquement capables de délivrer des contrastes très élevés, restent fortement limitées par les résidus de correction adaptative et les très faibles biais instrumentaux encore non-corrigés. La technologie progressant, notamment celle des détecteurs opérant dans l'infrarouge accompagnée par une puissance de calcul en temps réel accrue, les systèmes d'optique adaptative disposent au jour d'aujourd'hui d'un fort potentiel de progression, en se mettant à utiliser l'information présente dans le plan focal à des fins de métrologie. J'espère bien pouvoir continuer à contribuer à cette amélioration avec les méthodes de métrologie plan focal mises au point dans le cadre du projet KERNEL, généralisant le travail accompli jusque là en explorant les aspects multi-λ et multi-source (tomographie), avec un projet de thèse qui n'a pas encore démarré.

Je continue cependant de penser que bâtir une méthodologie observationnelle, la coronographie, en pariant sur la perfection de la correction du front d'onde reste une stratégie extrêmement risquée.

Pour arriver à un potentiel de détection et de caractérisation rivalisant avec ce que permettent déjà les méthodes indirectes, plusieurs ordres de grandeur sont encore à gagner sur la performance en contraste. Les résidus de contrôle de front d'onde, si petits deviennent ils au fur et à mesure des améliorations des systèmes d'optique adaptative, resteront encore longtemps le terme dominant les limites de détection directe des systèmes d'imagerie haut contraste.

CONTRASTANT AVEC LA CORONOGRAPHIE, les méthodes d'interprétation interférométrique des données exploitant des quantités observables auto-étalonnées comme la clôture de phase et leur généralisation sous formes de kernels offrent une alternative intéressante, capable de s'affranchir de ces erreurs résiduelles. Cette robustesse intrinsèque des mesures auto-étalonnées est particulièrement attirante. Malheureusement, la dynamique de cette approche reste limitée et empêche leur utilisation dans le cas des détections très haut contraste. On aimerait pouvoir intégrer l'extraction de quantités observables auto-étalonnées d'images haut-contraste. Pour conclure cette thèse de HDR, j'aimerais par conséquent évoquer deux pistes de travail prometteuses qui réussissent à combiner un système d'atténuation du bruit de photon d'une source brillante à une méthode de prise et d'analyse de données conduisant à la formation de kernels. La première application concerne les observations haut contraste avec un seul télescope. La deuxième exploite le potentiel de l'interférométrie longue base et mène à un concept d'instrument dont la réalisation pourrait être une suite logique au projet KERNEL.

Les kernels d'une pupille apodisée ?

La coronographie, combinant un masque plan focal à un masque de Lyot, n'est pas la seule manière de faire du haut contraste. L'apodisation de pupille peut également être utilisée seule pour créer dans le plan image, des régions où la PSF résultante est intrinsèquement plus contrastée. Si la prise d'image respecte les critères d'échantillonnage, de non saturation et se place dans le régime des faibles aberrations, les données peuvent tout à fait faire l'objet d'une analyse kernel. Il faut simplement construire un modèle discret de la pupille qui soit adapté à la pupille apodisée.

Sans trop trahir les subtiles nuances qui font du design de ce genre d'apodisation un art à part entière [92], on peut dire que le choix d'une apodisation de pupille est le résultat d'un compromis entre transmission, contraste maximal et séparation angulaire opérationnelle minimale (ou IWA). A moins d'utiliser une apodisation de type PIAA, une apodisation résultant en des images fortement contrastées se traduit par une faible transmission et un grand IWA. Si enfin on impose au masque d'être à transmission binaire, sa géométrie tend à se complexifier. La construction d'un modèle discret pour une telle pupille demanderait un modèle d'une très grande finesse, dont l'efficacité reste encore à démontrer.

92. M. N'Diaye, R. Soummer, L. Pueyo, A. Carlotti, C. C. Stark, and M. D. Perrin. Apodized Pupil Lyot Coronagraphs for Arbitrary Apertures. V. Hybrid Shaped Pupil Designs for Imaging Earth-like planets with Future Space Observatories. **ApJ**, 818:163, Feb. 2016

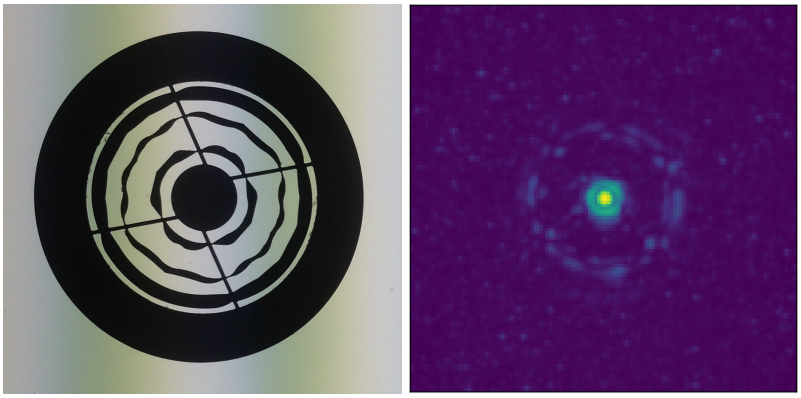


FIGURE 86: Gauche : photographie du masque apodiseur conçu pour une analyse kernel. Il est découpé au laser sur un substrat carbone, normalement utilisé pour fabriquer les masques de l'instrument MOIRCS de Subaru. Le diamètre externe du disque visible est de 25.4 mm pour montage dans une roue standard. Droite : exemple d'image focale obtenue sur le ciel affichée avec une échelle verticale logarithmique.

Ainsi, plutôt que de chercher tout de suite une solution optimale, il me paraît adapté de rapidement mettre à l'épreuve une apodisation aux ambitions modestes, présentant une transmission raisonnable élevée, et pour laquelle un modèle discret peut facilement être construit. En collaboration avec Mamadou N'Diaye (ancien postdoctorant du projet KERNEL et maintenant CR CNRS au laboratoire Lagrange), Romain Laugier (doctorant KERNEL) a donc fait le design d'une apodisation répondant à ce cahier des charges. Lors d'une visite récente (juin 2018) à Subaru pour des observations avec SCExAO, nous avons utilisé une machine à découpe laser de Subaru pour fabriquer cet apodiseur, dont une photographie est visible dans le panneau de gauche de la figure 86. A sa droite, une image de l'étoile Véga (juin 2018) prise en présence de ce masque (avec la caméra CRED2 de SCExAO) révèle la région annulaire haut contraste autour du coeur de la PSF.

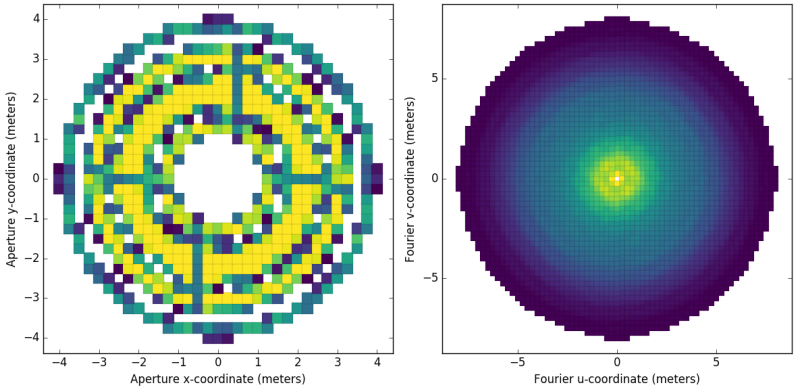


FIGURE 87: Représentation par XARA des propriétés du modèle discret utilisé pour l'analyse des données du masque apodiseur compatible avec l'analyse kernel, avec sans le panneau de gauche, une représentation grise de la pupille et à droite, la couverture du plan UV (module de la transformée de Fourier) offerte par ce modèle.

Dans la zone annulaire de plus haut contraste, il est donc possible de simultanément bénéficier de l'effet de réduction du bruit de photon de l'étoile centrale et du filtrage du bruit de phase résiduelle apporté par l'analyse kernel, ce qui devrait se traduire par une sensibilité accrue. La figure 87 présente à cette fin un exemple de modèle discret à transmission non-binaire, de façon à mieux décrire avec un nombre pas trop important de sous-ouvertures virtuelles, les détails de la structure du masque.

Etant donné la contrainte de non-saturation, cette démonstration va devoir bénéficier de la haute sensibilité de la caméra CRED1 du projet KERNEL, temporairement installée sur le port visiteur IR de SCExAO (voir la figure 88). Nos premiers tests de juin étant très encourageants, des observations complémentaires sont prévues dans la deuxième moitié d'octobre 2018.

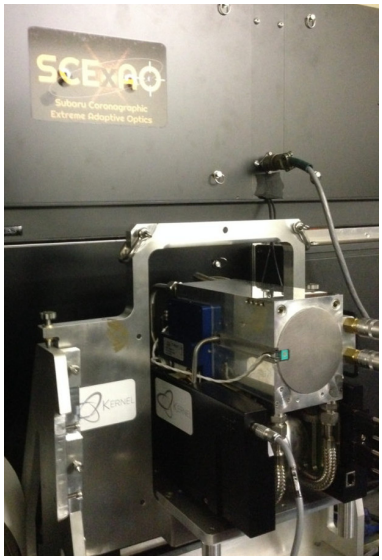


FIGURE 88: Photographie de la caméra CRED1 du projet KERNEL (K-CAM) après installation sur le port IR visiteur de l'instrument SCExAO (juin 2018).

Pour pousser encore plus cette méthode, nous étudions également dans le cadre de cette thèse, la possibilité de délibérément laisser saturer le coeur de la PSF qui peut être reconstruit par interpolation de façon à ce que le signal des structures faibles de l'image ressorte mieux, ce qui devrait se traduire par une dynamique plus élevée. Cette approche a été appliquée avec succès sur des données en provenance de HST au sujet desquelles nous sommes en train de finaliser un article. Il y a ici mise à profit combinée de plusieurs des développements faits dans le cadre du projet KERNEL : une meilleure prise en compte des bruits de mesure et de leur covariance, la mise en oeuvre d'un modèle discret avec transmission grise et récupération des effets de saturation. Toutes ces idées font l'objet de travaux en cours, menés par Romain Laugier et Alban Ceau (tous deux doctorants KERNEL) et feront l'objet de publications dans les mois à venir. Ce seront les premiers résultats importants du projet KERNEL.

Les kernels d'un interféromètre annulant ?

Lors d'une visite récente (novembre et décembre 2018) à l'Observatoire du Mont Stromlo, sponsorisée par un programme de l'Université Nationale Australienne (ANU), des discussions répétées avec mon collègue M. Ireland nous ont mené à imaginer un instrument interférométrique annulant d'un genre nouveau [93], produisant à partir des faisceaux en provenance de quatre stations d'observation, des signaux interférométriques pouvant être combinés de façon à former des kernels. L'article, en cours de révision après une première lecture, devrait être accepté pour publication au moment de la soutenance de cette thèse de HDR : je vais tenter de décrire les principes généraux sur lesquels ce concept repose. J'évoquerai finalement la possible généralisation des idées ayant mené à ce concept et leur impact sur le design d'instruments haut-contraste à venir.

93. F. Martinache and M. J. Ireland. Kernel-nulling for a robust direct interferometric detection of extrasolar planets. *ArXiv e-prints*, Feb. 2018

On a vu dans la première partie de ce mémoire que pour un interféromètre à deux télescopes, l'effet d'une erreur de cophasage, induite par l'atmosphère ou l'infrastructure, est indistinguable d'un déplacement de la source sur le ciel, et se manifeste par un changement de la phase de Fourier (voir l'équation 4). C'est en tirant avantage des effets de couplage entre des observations simultanées le long de lignes de base multiples qu'on a pu isoler une quantité observable invariante : la clôture de phase. L'idée est de trouver une façon de construire une telle quantité observable, au foyer d'un interféromètre annulant.

L'interférométrie annulante (j'emploierai également le terme de *nulling*) est une technique dont la description au premier abord est

déconcertante, surtout à la lumière de ce que je viens de rappeler. A la distinction d'un interféromètre classique, un interféromètre annulant (ou *nuller*) à deux télescopes requiert l'introduction sur un de ses deux bras, d'un déphasage achromatique, de façon à ce que lorsque l'interféromètre est cophasé, le système soit aligné sur une frange sombre. Comme l'acquisition de franges d'interférences dans un instrument classique (ie. pas annulant) requiert l'échantillonnage d'au moins une alternance de la fonction d'intensité décrivant les franges, soit par modulation dans le cadre d'un mode de recombinaison co-axial (un scan), soit en utilisant un système multi-axial, on peut se demander l'intérêt de la manoeuvre.

Pour que son utilisation ait du sens, un *nuller* requiert en réalité un système stabilisé en amont par un compensateur actif, équivalent de l'optique adaptative et un simple détecteur mono-pixel mesurant le flux produit en sortie du *nuller*. Un tel système est en quelque sorte l'équivalent d'un coronographe stabilisé par optique adaptative, pour lequel on ne mesurerait que le flux intégré en sortie du Lyot stop.

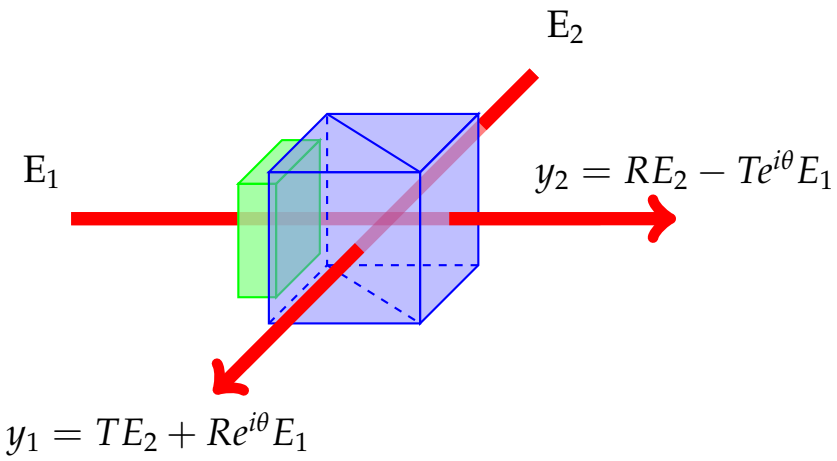


FIGURE 89: Implémentation possible d'un recombineur interférométrique annulant à deux faisceaux. Les deux champs électriques E_1 et E_2 collectés par les deux ouvertures (supposées de même pouvoir collecteur) sont recombinés à l'aide d'un cube séparateur (représenté en bleu) caractérisé par une transmission T et une réflectivité R (telles que $R^2 + T^2 = 1$). Sur le chemin du champ E_1, une lame déphasante achromatique (représentée en vert) introduit un déphasage $e^{i\theta}$. Pour ce cas simple, $T = R = 1/\sqrt{2}$ et $\theta = \pi$ conduisent à un *nuller*.

La figure 89 présente une implémentation possible pour un *nuller* à deux faisceaux[94]. Lorsque l'interféromètre alimentant ce recombineur observe une source ponctuelle, le déphasage de π affectant le champ E_1 avant son entrée dans le cube séparateur ($R = T = 1/\sqrt{2}$) résulte en une interférence parfaitement destructive avec le champ E_2. Ce recombineur permet de monitorer l'intensité $\mathbf{I} = \mathbf{y}^2$ collectée aux deux sorties :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}. \tag{68}$$

Dans des conditions d'observations idéales, si l'intensité I_1 diffère de zéro, le *nuller* enregistre la présence d'une structure située hors-axe. En combinant ensemble des observations pour différentes orientations de l'interféromètre[95], on peut reconstruire une carte de l'objet.

94. O. Guyon, B. Mennesson, E. Serabyn, and S. Martin. Optimal Beam Combiner Design for Nulling Interferometers. **PASP**, 125:951, Aug. 2013

95. R. N. Bracewell. Detecting nonsolar planets by spinning infrared interferometer. **Nature**, 274:780, Aug. 1978

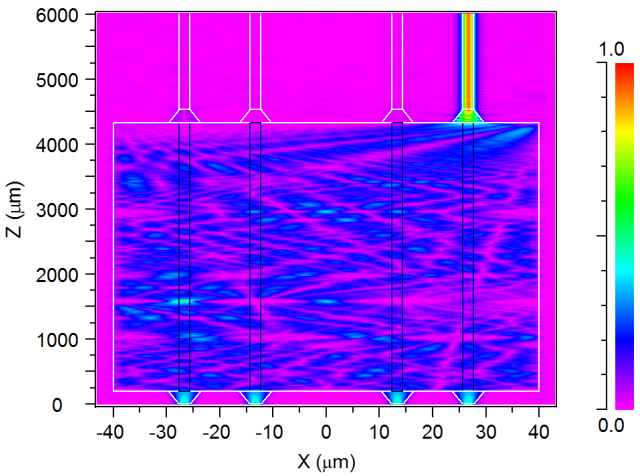


FIGURE 90: Implémentation possible de recombineur pour quatre faisceaux décrit par l'équation 69 utilisant la technologie MMI : les quatre entrées sont visibles sur la partie base de l'image. Ces entrées sont couplées dans un bloc de verre chalcogénide à l'intérieur duquel une fonction d'indice a été imprimée après optimisation pour permettre dans une bande la plus large possible, de rejeter le flux associé à une source située sur l'axe en la redirigeant intégralement sur une des quatre sorties située en haut de la figure. La lumière associée à des structures situées hors-axe va filtrer dans les trois autres sorties "sombres" (image par Harry Kenchington Goldsmith).

A l'instar d'un coronographe, les sorties d'un *nuller* sont sensibles à la présence de perturbations de phase instrumentale. Dans l'espoir de construire une ou plusieurs relations de clôture, je vais tout de suite considérer un *nuller* impliquant quatre sous-ouvertures. Un tel système peut être décrit par un modèle semblable à l'équation 68 avec un opérateur de *nulling* **N** agissant sur les quatre champs électriques collectés par les quatre sous-ouvertures :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \frac{1}{\sqrt{4}} \times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{bmatrix}. \tag{69}$$

Le facteur global $1/\sqrt{4}$ est nécessaire pour la contrainte de conservation du flux global : $||y||^2 = ||\mathbf{N} \cdot \mathbf{E}||^2 = ||\mathbf{E}||^2$. Bien qu'il soit possible de construire un tel nuller à partir d'un assemblage de composants discrets semblables à celui représenté dans la figure 89, la technologie de l'optique intégrée, par exemple à base d'interféromètres multi-modes (MMI, voir la figure 90)[96] qui permet d'achromatiser les propriétés de la fonction recherchée, semble pertinente.

96. H.-D. Kenchington Goldsmith, N. Cvetojevic, M. Ireland, and S. Madden. Fabrication tolerant chalcogenide mid-infrared multimode interference coupler design with applications for Bracewell nulling interferometry. *Optics Express*, 25:3038, Feb. 2017

Contrairement au cas d'un recombineur interférométrique classique pour lequel la relation entre la phase instrumentale φ et la phase de Fourier Φ est linéaire (voir l'équation 40), l'intensité **I** mesurée en sortie d'un *nuller* est une fonction **quadratique** de cette même perturbation φ. On se place dans l'hypothèse où les faisceaux entrant dans *nuller* sont stabilisés par un *fringe-tracker*. On peut alors supposer que les excursions de phase sont de faible amplitude et utiliser un développement linéaire au premier ordre pour chaque terme de champ électrique E_i affecté par une perturbation φ_i :

$$E_i = E_0 \times e^{-i\varphi} \approx E_0 \times (1 - i\varphi).$$

En l'absence de perturbation ($\varphi = 0$), le *nuller* est dans sa position de référence et redirige intégralement la lumière de l'étoile pointée (supposée non-résolue) vers la sortie y_1 de l'équation 69, laissant les

trois sorties sombres. Une perturbation de phase va redistribuer l'intensité et de la lumière de l'étoile cible va être couplées dans les sorties y_2, y_3 et y_4. On choisit la dernière des sous-ouvertures comme référence de phase instrumentale ce qui laisse trois termes de phase $(\varphi_i)_{i=1}^3$. La perturbation des sorties $\Delta\mathbf{I}$ s'écrit comme le résultat de [97] :

$$\Delta\mathbf{I} = E_0^2 \times ||\mathbf{N} \cdot [\varphi_1, \varphi_2, \varphi_3, 0]^T||^2. \tag{70}$$

Avec trois phases $(\varphi_i)_{i=1}^3$, on peut construire six termes de perturbation du deuxième ordre. On peut donc ré-écrire cette dernière équation comme le produit d'une nouvelle matrice $\mathbf{A}$ (qui va jouer un rôle analogue à la matrice de transfert du modèle des noyaux de phase) et d'un vecteur contenant ces six perturbations du deuxième ordre [98] :

$$\Delta\mathbf{I} = \mathbf{A} \cdot \left[\varphi_1^2, \varphi_2^2, \varphi_3^2, \varphi_1 \times \varphi_2, \varphi_1 \times \varphi_3, \varphi_2 \times \varphi_3\right]^T \tag{71}$$

La structure de $\mathbf{A}$ va décider de la possibilité de construire ou pas des kernels. Si on intègre la perturbation de la sortie brillante parmi les termes de $\Delta\mathbf{I}$, cette matrice $\mathbf{A}$ s'écrit [99] :

$$\mathbf{A} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 \end{bmatrix} \tag{72}$$

On peut rapidement identifier un kernel $\mathbf{K} = [1\,1\,1\,1]$ qui vérifie $\mathbf{K} \cdot \mathbf{A} = 0$: ce kernel correspond tout simplement à la conservation du flux global [100]. En plus de ce kernel trivial, aucun autre kernel ne peut être construit pour ce système. Pour sortir de cette situation, il faut modifier l'architecture du *nuller* et identifier un design qui fasse que les intensités des différentes sorties répondent de façon asymmétrique aux différentes perturbations.

La solution que nous avons proposé consiste à rajouter un étage de mélange des signaux des différentes sorties sombres du *nuller* $\mathbf{N}$. Chaque sortie $(y_i)_{i=2}^4$ est séparée en quatre faisceaux de même intensité, et un déphasage $e^{i\theta}$ est appliqué sur deux de ces quatre faisceaux. Les faisceaux en provenance de sorties différentes sont ensuite mélangés les uns avec les autres. En retenant un déphasage de $\pi/2$, l'effet de cette étape de mélange se résume à une nouvelle opération de *scrambling* $\mathbf{S}$:

$$\mathbf{S} = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & i & 0 \\ i & 1 & 0 \\ 1 & 0 & i \\ i & 0 & 1 \\ 0 & 1 & i \\ 0 & i & 1 \end{bmatrix}. \tag{73}$$

$\mathbf{S}$ est une matrice 6×3, complexe, prenant les trois sorties $[y_2\, y_3\, y_4]^T$ du *nuller* $\mathbf{N}$ pour les mélanger et produire six nouvelles sorties, dont on va pouvoir cette fois, tirer des kernels non-triviaux. La figure 91

97. Pour rester homogènes, les équations devraient faire apparaître le facteur global E_0^2 comme dans l'équation en cours de discussion. Dans la suite (on ne discute pas de fluctuations photométriques), les équations seront normalisées et ce facteur global éliminé pour augmenter leur lisibilité.

98. Tous ces termes correspondent aux dérivées secondes de $\mathbf{I}$. Y a t'il un intérêt à réécrire ces équations en faisant intervenir le Hessien de $\mathbf{I}$?

99. Il est intéressant de noter que les trois premiers termes de chaque ligne, correspondant aux termes non-croisés, sont tous égaux. C'est parmi les signes associés termes croisés qu'on observe de la diversité.

100. La lumière qui filtre vers les sorties sombres provient de la sortie brillante.

présente la façon dont on propose d'interfacer l'étage du *nuller* **N** à l'étage du *scrambler* **S**.

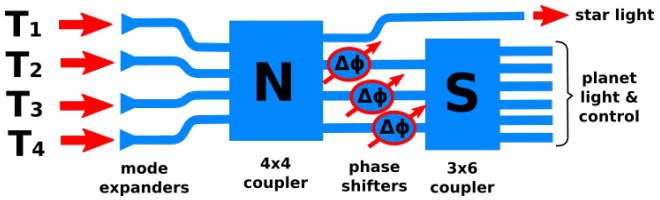


FIGURE 91: Architecture proposée pour le concept de *nuller* modifié pour permettre la formation de kernels. Les trois sorties sombres du *nuller* servent d'entrée à un étage de *scrambling* qui les mélange et conduit à la formation de six sorties.

On peut comme précedemment déterminer la dépendance de ces six nouvelles sorties mélangées à une perturbation de phase φ, et déterminer les entrées d'une nouvelle matrice **A** décrivant la dépendance du système aux six perturbations de deuxième ordre. On va ici délaisser la sortie brillante et ne s'intéresser qu'aux six nouvelles sorties sombres. La matrice **A** s'écrit alors :

$$\mathbf{A} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & -1 & 0 \end{bmatrix}. \tag{74}$$

Avec cette architecture il est de nouveau facile d'identifier trois kernels, correspondant chacun à la soustraction de paires consécutives des six sorties [101] :

$$\mathbf{K} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}. \tag{75}$$

101. Comme on va le voir, ces trois nouveaux kernels sont non-triviaux et apportent tous les trois une information utile sur la structure de la source observée.

POUR RÉSUMER, la proposition de recombineur interférométrique intègre deux fonctions distinctes (voir figure 91) de *nulling* et de *scrambling* produisant une sortie brillante vers laquelle le signal de la source (non-résolue) pointée est intégralement redirigé et six sorties sombres dans lesquelles la signature de structures situées dans le voisinage direct de l'étoile visée peut filtrer. Cet assemblage de fonctions est toujours sensible aux perturbations de phase cependant les six sorties qui en résultent peuvent être combinées pour former des kernels. Ces nouveaux kernels, de nature très différente de celle des noyaux de phase, sont des quantités observables insensibles aux perturbation du second ordre qui affectent naturellement la sortie d'un *nuller*. Un recombineur classique (c'est à dire qui ne soit pas annulant) au foyer d'un réseau à n télescopes offre accès instantané à jusqu'à $n_b = n \times (n-1)/2$ mesures de visibilité complexe qui peuvent être combinées pour former $n_K = (n-1) \times (n-2)/2$ clôtures (ou kernels) de phase. Il est plutôt satisfaisant de voir qu'avec ses trois kernels, l'architecture proposée permet de former dans le cas d'un *nuller* à quatre ouvertures, un nombre d'observables en accord avec le nombre optimal d'un recombineur classique.

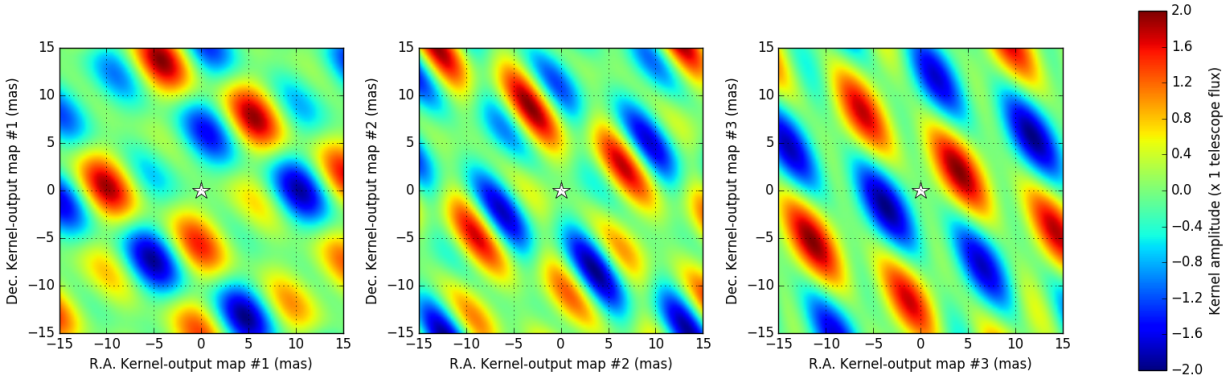


FIGURE 92: Trois cartes montrant comment les trois valeurs des kernels construits à partir des six sorties du *scrambler* évoluent en fonction de la position dans le champ. Les amplitudes des kernels sont données en multiples du flux collecté par un télescope.

La réponse sur le ciel d'un *nuller* est en grande partie conditionnée par la géométrie au sol du réseau interférométrique et la couverture effective du plan (u,v) qui dépend en plus des coordonnées de l'objet au moment de l'observation. Pour alimenter cette première description du concept, nous avons choisi d'utiliser la géométrie des quatre UTs du VLTI, dans le cas d'une observation se passant au zénith. L'étude utilise l'architecture du *nuller* qui vient d'être présentée pour former des kernels. La figure 92 présente comment l'amplitude des trois kernels de l'instrument proposé évolue en fonction de la position d'une source ponctuelle dans un champ de ± 15 mas centré sur la position de référence de l'instrument.

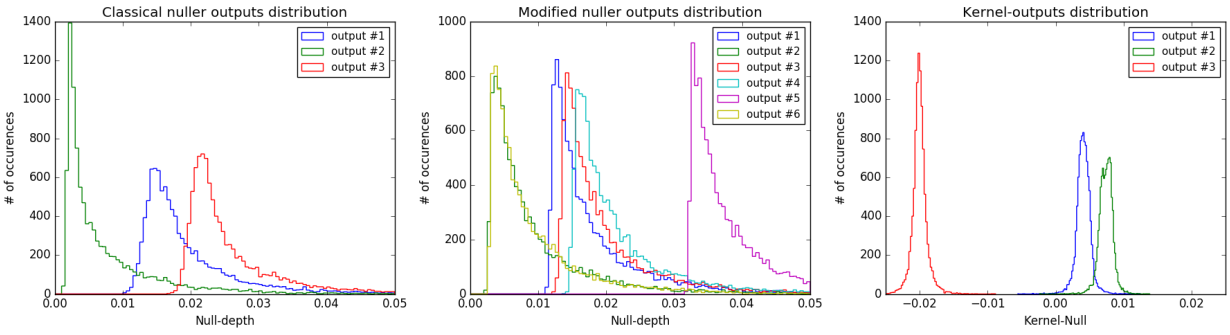


FIGURE 93: Distribution des sorties du *nuller* lors de l'observation d'un système binaire avec un compagnon de contraste $c = 10^{-2}$ aux coordonnées (+4.8,+1.8) dans les cartes précédentes en présence d'un résidu d'excursions de piston caractérisé par un RMS de 50 nm. A gauche : les histogrammes des trois sorties sombres qui seraient enregistrées en n'utilisant que l'étage du *nuller*. Au centre : les histogrammes des six sorties sombres enregistrées en sortie du *scrambler*. A droite : les histogrammes des trois kernels construits à partir des six sorties. Les lignes pointillées marquent les valeurs attendues des différentes mesures en l'absence de piston résiduel.

L'effet du *nulling* n'est pas uniforme : les trois kernels présentent ainsi de multiples minima et maxima pour différentes positions dans le champ couvert. Sur l'axe, l'effet du *nuller* est par contre clair et les trois kernels sont bien nuls. Les cartes de kernels présentent par contre toutes les trois une propriété intéressante : elles sont en effet anti-symétriques. Une seule observation apporte une contrainte forte sur la position d'un éventuel compagnon à l'objet observé.

Le véritable intérêt de ce concept est révélé lorsqu'on inclut des erreurs de phase. La figure 93 permet de visualiser ce que l'étage de mélange et la construction des kernels apporte. La simulation inclut le signal produit par un système binaire, avec un compagnon

de contraste $c = 10^{-2}$ situé au point de coordonnées (+4.8,+1.8) mas sur les cartes de la figure 92 en présence de turbulence. Les histogrammes des valeurs mesurées par le *nuller* seul sont fortement asymétriques révèlent des excursions importantes du *null*. Les histogrammes des valeurs brutes mesurées en sortie du *scrambler* sont très semblablement distribuées et la réduction apparente du domaine d'excursion des valeurs mesurées s'explique simplement par la réduction par un facteur deux du flux par sortie (en passant de trois à six sorties). Chaque série de mesures instantanées des six sorties peut cependant être combinée pour former trois kernels : la suppression des erreurs du second ordre affectant les sorties brutes du *nuller* se matérialise par des histogrammes beaucoup plus piqués et centrés sur la valeur attendue : la mesure qui est offerte par les kernels est effectivement rendue robuste aux excursions de piston.

C'EST LA PREMIÈRE FOIS qu'un concept d'instrument haut-contraste combine à la fois les effets positifs d'un *nuller* qui atténue autant que possible la contribution du bruit de photon de la source centrale tout en étant robuste aux excursions de contrôle de front d'onde. Ce concept de *kernel-nulling* nous a naturellement mené à évoquer dans notre article la possibilité d'un instrument haut-contraste pour le VLTI, que nous avons appelé VIKiNG (acronyme de VLTI Infrared Kernel NullinG) et identifie une douzaine de planètes détectées par vitesse radiale qui pourraient observées par cet instrument. Ce concept d'instrument me semble constituer un progrès important : il est conçu pour fonctionner dans des conditions plutôt favorables, c'est à dire en bénéficiant de l'effet stabilisateur d'un *fringe tracker* mais qui ne parie pas sur la perfection du contrôle de front d'onde. J'espère que ma collaboration avec M. Ireland, qui dure depuis mon premier contrat de post-doctorant, conduira à la matérialisation d'un tel instrument qui présente un fort potentiel pour véritablement contribuer à des détections directes de planètes extrasolaires.

Vers une kernel-coronographie ?

Peut on extrapoler les idées qui ont mené au concept de *kernel-nulling* et les appliquer au scénario de la coronographie ? Le problème est plus compliqué à formaliser car il implique un très grand nombre de degrés de liberté. Le nombre fini de faisceaux recombinés dans le cas du *nulling* rend en effet le problème relativement facile à écrire. Avec n sous-ouvertures, on a $n-1$ sources de perturbation de phase et $(n-1) \times n/2$ termes de perturbation du second ordre pour lesquels on va chercher à assembler des kernels. Une première étape vers cette généralisation serait de voir si une architecture de *nuller* + *scrambler* peut toujours être trouvée pour un nombre arbitraire de sous-ouvertures ou s'il existe des combinaisons non-permises. Une architecture pour un réseau à six télescopes pourrait conduire à un instrument *kernel-nuller* adapté à l'interféromètre CHARA.

Une autre étape serait d'explorer ce qui est possible de réaliser avec un masque coronographique adapté, au foyer d'un masque à ouverture non-redondante, qui reste après tout un scénario assez simple dont on peut encore apprendre des choses. Complété par un Lyot-stop lui même non-redondant, ce montage est comparable à un *nuller*. Un équivalent de l'étage de *scrambling* pourrait être un masque de phase placé dans le plan de Lyot. Je n'ai cependant pas encore d'idée claire sur la façon dont il faudrait aborder l'analyse des interférogrammes produits par un tel système.

Une autre approche, plus proche de l'imagerie coronographique telle qu'elle est implémentée aujourd'hui est d'inclure à l'image focale, les informations nécessaires à la réalisation d'une bonne métrologie qui permet de séparer la contribution de la phase instrumentale de celle des structures du voisinage de la source observée. L'étude d'une utilisation de l'analyse de front d'onde plan focal coronographique avec un Lyot-stop asymétrique figure ainsi au programme du projet KERNEL. Cette démarche nous ramène également à l'idée de *speckle-nulling* dont j'ai parlé dans le deuxième chapitre de cette thèse. Une alternative au *speckle-nulling* qui requiert une modulation temporelle de la perturbation, est, d'employer un codage spatial. La *self-coherent camera* (SCC) [102] est une proposition existante qui répond en partie au besoin mais à la mise au point délicate qui n'offre à mon sens pas encore suffisamment de robustesse. Pour le moment, utilisée à des fins de contrôle de front d'onde, il serait intéressant de voir si à partir de mesures d'amplitude complexe dans le plan focal, un assemblage de kernels est possible qui permettrait de s'affranchir des limites résultant des effets subtils de diffraction de Fresnel introduits par les optiques elles mêmes.

102. R. Galicher, P. Baudoz, and G. Rousset. Wavefront error correction and Earth-like planet detection by a self-coherent camera in space. **A&A**, 488:L9–L12, Sept. 2008

Pour conclure

SE DOTER DE TECHNIQUES ET DE PRINCIPES D'OBSERVATION ROBUSTES AUX PERTURBATIONS DE SURFACE D'ONDE me semble être une étape cruciale pour que l'imagerie haut contraste reste une proposition pertinente, susceptible de faire une contribution significative dans le domaine des exoplanètes depuis le sol. Utilisant des concepts instrumentaux considérablement moins sophistiqués que les imageurs haut-contraste, les techniques d'observation de la mesure de vitesse radiale, l'observation de lentilles gravitationnelles, le suivi de transit photométrique et la spectroscopie des transits ont permis des progrès considérables et couvert une large partie de l'espace des paramètres. La performance de ces techniques est également dégradée par le *seeing* mais contrairement au haut-contraste, cette dégradation est progressive. L'imagerie haut-contraste reste par conséquent encore une contribution très mineure qui n'a pas encore permis la révolution qu'elle a le potentiel d'accomplir.

Avec les télescopes de la classe des 20-30 mètres, l'astronomie est en train de se doter de moyens d'observation pour lesquels l'imagerie directe des planètes extrasolaires était le cas scientifique moteur et pourtant aucun des trois grands télescopes de cette classe ne prévoit d'instrument véritablement haut contraste en première génération. On peut faire un constat semblable pour les missions spatiales : au moment de ma thèse de doctorat, le haut contraste semblait être une évidence, que ce soit pour la coronographie avec TPF-C et pour l'interférométrie avec SIM, SIM-lite, TPF-I et Darwin. Aucune de ces missions spatiales n'a malheureusement vu le jour. Les missions dédiées au transits comme Kepler (lancé en 2009), TESS (lancé en 2018) et PLATO (dont le lancement prévu en 2026) dominent le paysage "exoplanète" de l'astrophysique spatiale. Cette attitude des agences vis à vis du haut contraste est révélatrice d'une forme de crise de confiance. Une solution à cette crise viendra peut être du mariage des techniques, par exemple celui de l'imagerie haut-contraste combinée à la mesure de vitesse radiale grâce à l'injection dans une ou plusieurs fibres optiques. Des observations partiellement résolues de transits photométriques avec des systèmes multi-fibres bénéficiant de correction XAO pourraient peut être également apporter aux transits des contraintes sur la configuration spatiale des systèmes observés.

L'attitude saine de l'optimisme réaliste est de toujours espérer le meilleur en se ne cessant jamis de se préparer au pire. Pour faire de la coronographie depuis le sol, il faut être résolument optimiste, car beaucoup d'éléments conspirent contre nos efforts dans l'objectif du haut-contraste mais l'optimisme seul ne suffit pas. Je doute que les dernières idées proposées dans ce chapitre de conclusion soient finales : elles sont cependant au jour d'aujourd'hui ma tentative sincère d'introduire du réalisme à la grande ambition de l'imagerie haut-contraste. Il n'y a qu'en acceptant pleinement la nature impermanente de nos observations, qu'on peut espérer mettre en évidence des observables qui la transcendent. L'optique adaptative nous offre les moyens de contrôler cette impermanence mais ce contrôle reste partiel : avec des niveaux d'exigence croissants, les résidus de contrôle contribueront toujours significativement à un plancher de performance.

Cette thèse de HDR est nécessairement un bilan des différentes activités scientifiques accomplies depuis la soutenance de ma thèse de doctorat. L'imagerie directe des planètes extrasolaires habitables, bénéficie encore d'une aura positive et a permis de justifier beaucoup des développements auxquels j'ai eu la chance d'être associé. Je me suis permis d'aller au delà du simple bilan et ai profité de cette conclusion pour partager une réflexion et proposer une démarche visant à rendre service à cette cause. J'espère que les quelques pistes proposées dépasseront le cadre du projet KERNEL dans lequel je suis pour le moment engagé. Je vous remercie de votre attention et espère que la lecture de ce document aura réussi à stimuler votre réflexion.

Bibliographie

[1] K. Adler. *The Measure of All Things*. Free Press, 2002.

[2] D. Bernat, F. Martinache, M. Ireland, P. Tuthill, and J. Lloyd. The Use of Spatial Filtering with Aperture Masking Interferometry and Adaptive Optics. **ApJ**, 756 :8, Sept. 2012.

[3] J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, J. Antichi, A. Baruffolo, P. Baudoz, A. Berton, A. Boccaletti, M. Carbillet, J. Charton, R. Claudi, M. Downing, P. Feautrier, E. Fedrigo, T. Fusco, R. Gratton, N. Hubin, M. Kasper, M. Langlois, C. Moutou, L. Mugnier, J. Pragt, P. Rabou, M. Saisse, H. M. Schmid, E. Stadler, M. Turrato, S. Udry, R. Waters, and F. Wildi. SPHERE : A 'Planet Finder' Instrument for the VLT. *The Messenger*, 125, Sept. 2006.

[4] P. J. Bordé and W. A. Traub. High-Contrast Imaging from Space : Speckle Nulling in a Low-Aberration Regime. **ApJ**, 638 :488–498, Feb. 2006.

[5] B. P. Bowler. Imaging Extrasolar Giant Planets. **PASP**, 128(10) :102001, Oct. 2016.

[6] R. N. Bracewell. Detecting nonsolar planets by spinning infrared interferometer. **Nature**, 274 :780, Aug. 1978.

[7] A. P. Digby, S. Hinkley, B. R. Oppenheimer, A. Sivaramakrishnan, J. P. Lloyd, M. D. Perrin, L. C. Roberts, Jr., R. Soummer, D. Brenner, R. B. Makidon, M. Shara, J. Kuhn, J. Graham, P. Kalas, and L. Newburgh. The Challenges of Coronagraphic Astrometry. **ApJ**, 650 :484–496, Oct. 2006.

[8] T. M. Evans, M. J. Ireland, A. L. Kraus, F. Martinache, P. Stewart, P. G. Tuthill, S. Lacour, J. M. Carpenter, and L. A. Hillenbrand. Mapping the Shores of the Brown Dwarf Desert. III. Young Moving Groups. **ApJ**, 744 :120, Jan. 2012.

[9] R. Galicher, P. Baudoz, and G. Rousset. Wavefront error correction and Earth-like planet detection by a self-coherent camera in space. **A&A**, 488 :L9–L12, Sept. 2008.

[10] M. Golay. Point arrays having compact, nonredundant autocorrelations. *JOSA*, 61 :272–273, 1971.

[11] R. A. Gonsalves. Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21 :829–832, Oct. 1982.

[12] O. Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging. **A&A**, 404 :379–387, June 2003.

[13] O. Guyon, T. Matsuo, and R. Angel. Coronagraphic Low-Order Wave-Front Sensor : Principle and Application to a Phase-Induced Amplitude Coronagraph. **ApJ**, 693 :75–84, Mar. 2009.

[14] O. Guyon, B. Mennesson, E. Serabyn, and S. Martin. Optimal Beam Combiner Design for Nulling Interferometers. **PASP**, 125 :951, Aug. 2013.

[15] O. Guyon, E. Pluzhnik, F. Martinache, J. Totems, S. Tanaka, T. Matsuo, C. Blain, and R. Belikov. High-Contrast Imaging and Wavefront Control with a PIAA Coronagraph : Laboratory System Validation. **PASP**, 122 :71–84, Jan. 2010.

[16] O. Guyon, E. A. Pluzhnik, M. J. Kuchner, B. Collins, and S. T. Ridgway. Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs. **ApJS**, 167 :81–99, Nov. 2006.

[17] K. W. Hodapp, R. Suzuki, M. Tamura, L. Abe, H. Suto, R. Kandori, J. Morino, T. Nishimura, H. Takami, O. Guyon, S. Jacobson, V. Stahlberger, H. Yamada, R. Shelton, J. Hashimoto, A. Tavrov, J. Nishikawa, N. Ukita, H. Izumiura, M. Hayashi, T. Nakajima, T. Yamada, and T. Usuda. HiCIAO : the Subaru Telescope's new high-contrast coronographic imager for adaptive optics. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7014 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Aug. 2008.

[18] M. J. Ireland. Phase errors in diffraction-limited imaging : contrast limits for sparse aperture masking. **MNRAS**, 433 :1718–1728, Aug. 2013.

[19] R. C. Jennison. A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent. **MNRAS**, 118 :276–+, 1958.

[20] N. Jovanovic, F. Martinache, O. Guyon, C. Clergeon, G. Singh, T. Kudo, V. Garrel, K. Newman, D. Doughty, J. Lozi, J. Males, Y. Minowa, Y. Hayano, N. Takato, J. Morino, J. Kuhn, E. Serabyn, B. Norris, P. Tuthill, G. Schworer, P. Stewart, L. Close, E. Huby, G. Perrin, S. Lacour, L. Gauchet, S. Vievard, N. Murakami, F. Oshiyama, N. Baba, T. Matsuo, J. Nishikawa, M. Tamura, O. Lai, F. Marchis, G. Duchene, T. Kotani, and J. Woillez. The Subaru Coronagraphic Extreme Adaptive Optics System : Enabling High-Contrast Imaging on Solar-System Scales. **PASP**, 127 :890–910, Oct. 2015.

[21] N. Jovanovic, P. G. Tuthill, B. Norris, S. Gross, P. Stewart, N. Charles, S. Lacour, M. Ams, J. S. Lawrence, A. Lehmann, C. Niel, J. G. Robertson, G. D. Marshall, M. Ireland, A. Fuerbach, and M. J. Withford. Starlight demonstration of the Dragonfly instrument : an integrated photonic pupil-remapping interferometer for high-contrast imaging. **MNRAS**, 427 :806–815, Nov. 2012.

[22] H.-D. Kenchington Goldsmith, N. Cvetojevic, M. Ireland, and S. Madden. Fabrication tolerant chalcogenide mid-infrared multimode interference coupler design with applications for Bracewell nulling interferometry. *Optics Express*, 25 :3038, Feb. 2017.

[23] L. Koechlin, P. R. Lawson, D. Mourard, A. Blazit, D. Bonneau, F. Morand, P. Stee, I. Tallon-Bosc, and F. Vakili. Dispersed fringe tracking with the multi-r_o apertures of the Grand Interferometre a 2 Telescopes. **Appl. Opt.**, 35 :3002–3009, June 1996.

[24] A. L. Kraus and M. J. Ireland. LkCa 15 : A Young Exoplanet Caught at Formation ? **ApJ**, 745 :5, Jan. 2012.

[25] A. L. Kraus, M. J. Ireland, F. Martinache, and L. A. Hillenbrand. Mapping the Shores of the Brown Dwarf Desert. II. Multiple Star Formation in Taurus-Auriga. **ApJ**, 731 :8, Apr. 2011.

[26] A. L. Kraus, M. J. Ireland, F. Martinache, and J. P. Lloyd. Mapping the Shores of the Brown Dwarf Desert. I. Upper Scorpius. **ApJ**, 679 :762–782, May 2008.

[27] A. Labeyrie. Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier Analysing Speckle Patterns in Star Images. **A&A**, 6 :85, May 1970.

[28] A.-M. Lagrange, M. Bonnefoy, G. Chauvin, D. Apai, D. Ehrenreich, A. Boccaletti, D. Gratadour, D. Rouan, D. Mouillet, S. Lacour, and M. Kasper. A Giant Planet Imaged in the Disk of the Young Star β Pictoris. *Science*, 329 :57–, July 2010.

[29] B. Lopez, S. Lagarde, W. Jaffe, and collaborators. An Overview of the MATISSE Instrument - Science, Concept and Current Status. *The Messenger*, 157 :5–12, Sept. 2014.

[30] J. Lozi, O. Guyon, N. Jovanovic, G. Singh, S. Goebel, B. Norris, and H. Okita. Characterizing and mitigating vibrations for SCExAO. In *Adaptive Optics Systems V*, volume 9909 of **Proc. SPIE**, page 99090J, July 2016.

[31] J. Lozi, F. Martinache, and O. Guyon. Phase-Induced Amplitude Apodization on Centrally Obscured Pupils : Design and First Laboratory Demonstration for the Subaru Telescope Pupil. **PASP**, 121 :1232–1244, Nov. 2009.

[32] B. Macintosh, J. R. Graham, P. Ingraham, Q. Konopacky, C. Marois, M. Perrin, L. Poyneer, B. Bauman, T. Barman, A. S.

Burrows, A. Cardwell, J. Chilcote, R. J. De Rosa, D. Dillon, R. Doyon, J. Dunn, D. Erikson, M. P. Fitzgerald, D. Gavel, S. Goodsell, M. Hartung, P. Hibon, P. Kalas, J. Larkin, J. Maire, F. Marchis, M. S. Marley, J. McBride, M. Millar-Blanchaer, K. Morzinski, A. Norton, B. R. Oppenheimer, D. Palmer, J. Patience, L. Pueyo, F. Rantakyro, N. Sadakuni, L. Saddlemyer, D. Savransky, A. Serio, R. Soummer, A. Sivaramakrishnan, I. Song, S. Thomas, J. K. Wallace, S. Wiktorowicz, and S. Wolff. First light of the Gemini Planet Imager. *Proceedings of the National Academy of Science*, 111 :12661–12666, Sept. 2014.

[33] F. Malbet, J. W. Yu, and M. Shao. High-Dynamic-Range Imaging Using a Deformable Mirror for Space Coronography. **PASP**, 107 :386, Apr. 1995.

[34] F. Martinache. Global wavefront sensing for interferometers and mosaic telescopes : the dispersed speckles principle. *Journal of Optics A : Pure and Applied Optics*, 6 :216–220, Feb. 2004.

[35] F. Martinache. Kernel Phase in Fizeau Interferometry. **ApJ**, 724 :464–469, Nov. 2010.

[36] F. Martinache. The Asymmetric Pupil Fourier Wavefront Sensor. **PASP**, 125 :422–430, Apr. 2013.

[37] F. Martinache. Spectrally dispersed Fourier-phase analysis for redundant apertures. In *Optical and Infrared Interferometry and Imaging V*, volume 9907 of **Proc. SPIE**, page 990712, Aug. 2016.

[38] F. Martinache, O. Guyon, C. Clergeon, and C. Blain. Speckle Control with a Remapped-Pupil PIAA Coronagraph. **PASP**, 124 :1288–1294, Dec. 2012.

[39] F. Martinache, O. Guyon, N. Jovanovic, C. Clergeon, G. Singh, T. Kudo, T. Currie, C. Thalmann, M. McElwain, and M. Tamura. On-Sky Speckle Nulling Demonstration at Small Angular Separation with SCExAO. **PASP**, 126 :565–572, June 2014.

[40] F. Martinache, O. Guyon, E. A. Pluzhnik, R. Galicher, and S. T. Ridgway. Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronograph. II. Performance. **ApJ**, 639 :1129–1137, Mar. 2006.

[41] F. Martinache and M. J. Ireland. Kernel-nulling for a robust direct interferometric detection of extrasolar planets. *ArXiv e-prints*, Feb. 2018.

[42] F. Martinache, J. P. Lloyd, M. J. Ireland, R. S. Yamada, and P. G. Tuthill. Precision Masses of the Low-Mass Binary System GJ 623. **ApJ**, 661 :496–501, May 2007.

[43] F. Martinache, B. Rojas-Ayala, M. J. Ireland, J. P. Lloyd, and P. G. Tuthill. Visual Orbit of the Low-Mass Binary GJ 164 AB. **ApJ**, 695 :1183–1190, Apr. 2009.

[44] D. Mawet, E. Serabyn, K. Liewer, R. Burruss, J. Hickey, and D. Shemo. The Vector Vortex Coronagraph : Laboratory Results and First Light at Palomar Observatory. **ApJ**, 709 :53–57, Jan. 2010.

[45] A. A. Michelson and F. G. Pease. Measurement of the diameter of alpha Orionis with the interferometer. **ApJ**, 53 :249–259, May 1921.

[46] Y. Minowa, Y. Hayano, S. Oya, M. Watanabe, M. Hattori, O. Guyon, S. Egner, Y. Saito, M. Ito, H. Takami, V. Garrel, S. Colley, T. Golota, and M. Iye. Performance of Subaru adaptive optics system AO188. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2010.

[47] J. D. Monnier. An Introduction to Closure Phases. In P. R. Lawson, editor, *Principles of Long Baseline Stellar Interferometry*, pages 203–+, 2000.

[48] T. Nakajima, S. R. Kulkarni, P. W. Gorham, A. M. Ghez, G. Neugebauer, J. B. Oke, T. A. Prince, and A. C. S. Readhead. Diffraction-limited imaging. II - Optical aperture-synthesis imaging of two binary stars. **AJ**, 97 :1510–1521, May 1989.

[49] M. N'Diaye, F. Martinache, N. Jovanovic, J. Lozi, O. Guyon, B. Norris, A. Ceau, and D. Mary. Calibration of the island effect : Experimental validation of closed-loop focal plane wavefront control on Subaru/SCExAO. *ArXiv e-prints*, Dec. 2017.

[50] M. N'Diaye, R. Soummer, L. Pueyo, A. Carlotti, C. C. Stark, and M. D. Perrin. Apodized Pupil Lyot Coronagraphs for Arbitrary Apertures. V. Hybrid Shaped Pupil Designs for Imaging Earth-like planets with Future Space Observatories. **ApJ**, 818 :163, Feb. 2016.

[51] B. R. Oppenheimer, C. Beichman, D. Brenner, R. Burruss, E. Cady, J. Crepp, L. Hillenbrand, S. Hinkley, E. R. Ligon, T. Lockhart, I. Parry, L. Pueyo, E. Rice, L. C. Roberts, J. Roberts, M. Shao, A. Sivaramakrishnan, R. Soummer, G. Vasisht, F. Vescelus, J. K. Wallace, C. Zhai, and N. Zimmerman. Project 1640 : the world's first ExAO coronagraphic hyperspectral imager for comparative planetary science. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8447 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2012.

[52] P. Pathak, O. Guyon, N. Jovanovic, J. Lozi, F. Martinache, Y. Minowa, T. Kudo, H. Takami, Y. Hayano, and N. Narita. A High-precision Technique to Correct for Residual Atmospheric Dispersion in High-contrast Imaging Systems. **PASP**, 128(12) :124404, Dec. 2016.

[53] G. Perrin, S. Lacour, J. Woillez, and É. Thiébaut. High dynamic range imaging by pupil single-mode filtering and remapping. **MNRAS**, 373 :747–751, Dec. 2006.

[54] B. J. S. Pope. Kernel phase and kernel amplitude in Fizeau imaging. **MNRAS**, 463 :3573–3581, Dec. 2016.

[55] L. Pueyo, J. Kay, N. J. Kasdin, T. Groff, M. McElwain, A. Give'on, and R. Belikov. Optimal dark hole generation via two deformable mirrors with stroke minimization. **Appl. Opt.**, 48 :6296, Nov. 2009.

[56] A. C. S. Readhead, T. S. Nakajima, T. J. Pearson, G. Neugebauer, J. B. Oke, and W. L. W. Sargent. Diffraction-limited imaging with ground-based optical telescopes. **AJ**, 95 :1278–1296, Apr. 1988.

[57] J.-F. Sauvage, T. Fusco, M. Lamb, J. Girard, M. Brinkmann, A. Guesalaga, P. Wizinowich, J. O'Neal, M. N'Diaye, A. Vigan, D. Mouillet, J.-L. Beuzit, M. Kasper, M. Le Louarn, J. Milli, K. Dohlen, B. Neichel, P. Bourget, P. Haguenauer, and D. Mawet. Tackling down the low wind effect on SPHERE instrument. In *Adaptive Optics Systems V*, volume 9909 of **Proc. SPIE**, page 990916, July 2016.

[58] J.-F. Sauvage, T. Fusco, C. Petit, A. Costille, D. Mouillet, J.-L. Beuzit, K. Dohlen, M. Kasper, M. Suarez, C. Soenke, A. Baruffolo, B. Salasnich, S. Rochat, E. Fedrigo, P. Baudoz, E. Hugot, A. Sevin, D. Perret, F. Wildi, M. Downing, P. Feautrier, P. Puget, A. Vigan, J. O'Neal, J. Girard, D. Mawet, H. M. Schmid, and R. Roelfsema. SAXO : the extreme adaptive optics system of SPHERE (I) system overview and global laboratory performance. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(2) :025003, Apr. 2016.

[59] G. Singh, F. Martinache, P. Baudoz, O. Guyon, T. Matsuo, N. Jovanovic, and C. Clergeon. Lyot-based Low Order Wavefront Sensor for Phase-mask Coronagraphs : Principle, Simulations and Laboratory Experiments. **PASP**, 126 :586, June 2014.

[60] A. Sivaramakrishnan, D. Lafrenière, K. E. S. Ford, B. McKernan, A. Cheetham, A. Z. Greenbaum, P. G. Tuthill, J. P. Lloyd, M. J. Ireland, R. Doyon, M. Beaulieu, A. Martel, A. Koekemoer, F. Martinache, and P. Teuben. Non-redundant Aperture Masking Interferometry (AMI) and segment phasing with JWST-NIRISS. In *Space Telescopes and Instrumentation 2012 : Optical, Infrared, and Millimeter Wave*, volume 8442 of **Proc. SPIE**, page 84422S, Sept. 2012.

[61] M. Tamura. Subaru Strategic Exploration of Exoplanets and Disks with HiCIAO/AO188 (SEEDS). In T. Usuda, M. Tamura,

and M. Ishii, editors, *American Institute of Physics Conference Series*, volume 1158 of *American Institute of Physics Conference Series*, pages 11–16, Aug. 2009.

[62] G. Torres. Substellar Companion Masses from Minimal Radial Velocity or Astrometric Information : a Monte Carlo Approach. **PASP**, 111 :169–176, Feb. 1999.

[63] P. Tuthill, J. Lloyd, M. Ireland, F. Martinache, J. Monnier, H. Woodruff, T. ten Brummelaar, N. Turner, and C. Townes. Sparse-aperture adaptive optics. In *Advances in Adaptive Optics II. Edited by Ellerbroek, Brent L. ; Bonaccini Calia, Domenico. Proceedings of the SPIE, Volume 6272, pp. (2006).*, July 2006.

[64] F. P. A. Vogt, F. Martinache, O. Guyon, T. Yoshikawa, K. Yokochi, V. Garrel, and T. Matsuo. Coronagraphic Low-Order Wavefront Sensor : Postprocessing Sensitivity Enhancer for High-Performance Coronagraphs. **PASP**, 123 :1434–1441, Dec. 2011.

[65] J. K. Wallace, R. S. Burruss, R. D. Bartos, T. Q. Trinh, L. A. Pueyo, S. F. Fregoso, J. R. Angione, and J. C. Shelton. The Gemini Planet Imager calibration wavefront sensor instrument. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7736 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2010.