Problèmes inverses en Haute Résolution Angulaire

Laurent M. Mugnier

ONERA

THE FRENCH AEROSPACE LAB

Dossier d'Habilitation à Diriger des Recherches
Université Paris Diderot – Paris 7

Discipline : PHYSIQUE

Problèmes inverses en Haute Résolution Angulaire

par Laurent MUGNIER
Office National d'Études et de Recherches Aérospatiales

Habilitation soutenue le 18 octobre 2011 devant le jury composé de :

Cécile Ferrari	rapporteur
Jean-François Giovannelli	rapporteur
David Mouillet	examinateur
Sylvie Roques	rapporteur
Gérard Rousset	examinateur

Table des matières

Table des figures

Première partie

Éléments de curriculum vitae et activités d'encadrement

Curriculum vitae

Laurent MUGNIER
ONERA
Dép^{t} Optique Théorique et Appliquée
BP 72, 92322 Châtillon cedex.

Français, né le 8 Octobre 1964 à Asnières.
Tél. : 01 46 73 47 47.
Fax : 01 46 73 41 71.
Mél : Laurent.Mugnier@onera.fr

EXPÉRIENCE

Depuis 1994 : chercheur puis Maître de Recherche (2000), Office National d'Études et de Recherches Aérospatiales.
En charge de la coordination des activités *problèmes inverses / traitement de données* de l'équipe Haute Résolution Angulaire.
Développement et application de méthodes de restauration et de reconstruction d'images pour l'imagerie à haute résolution angulaire : optique adaptative (OA) pour l'astronomie et l'observation de satellites, imagerie coronographique corrigée par OA, imagerie rétinienne, imagerie par analyse de front d'onde, imagerie interférométrique (depuis le sol et depuis l'espace), interférométrie à frange noire (*nulling*), synthèse d'ouverture active, observation satellitaire.
Étude de méthodes d'étalonnage d'un système optique : analyse de front d'onde et cophasage par diversité de phase, estimation de fonction de transfert myope sur image contrastée.
Modélisation et optimisation de systèmes imageurs à synthèse d'ouverture optique passive et active.
Chef de projet sur deux études de faisabilité d'un système d'observation de la Terre par interférométrie optique.
Encadrant principal de cinq doctorants et participation à l'encadrement de dix autres doctorants.
Enseignements réguliers à l'Institut d'Optique Graduate School (anciennement École Supérieure d'Optique) en 3^{e} année, au Master 2 Recherche de l'École Doctorale "Astronomie et Astrophysique d'Île-de-France", et depuis 2009 au M2 Pro de la même École Doctorale ainsi qu'à Supélec (campus de Metz). Enseignements à des écoles thématiques (*Imagerie Très Haute Dynamique et détection d'exoplanètes 2005*, et *Physics and Ophtalmology 2005*).

1993 : chercheur post-doctoral, Northwestern University, Département Electrical Engineering and Computer Science. Restauration d'images par utilisation du bispectre.

1989 à 1992 : chercheur doctorant, Télécom Paris. Sujet de thèse : *Vers une inversion des hologrammes conoscopiques*.
Développement de méthodes d'inversion de données interférométriques en vue de la reconstruction d'une scène tridimensionnelle. Conception, réalisation et exploitation d'un dispositif automatisé d'acquisition d'hologrammes conoscopiques. Délégué des doctorants auprès de l'administration de Télécom Paris.

FORMATION

Novembre 1992 :		**Doctorat**, spécialité Traitement du Signal et des Images, Télécom Paris (E.N.S.T.). Mention très honorable avec les félicitations du jury.
1988/89	:	Mastère en Traitement d'Images à Télécom Paris.
1985 à 1988	:	**École Polytechnique**.
Juin 1982	:	Baccalauréat C, mention Bien.

LANGUES

Anglais	:	bilingue.
Allemand	:	scolaire.

PUBLICATIONS

- Sept contributions à des ouvrages de synthèse :
  - « Blind Image Deconvolution », L. Blanc-Féraud, L. Mugnier & A. Jalobeanu, chap. 3 de *Inverse Problems in Vision and 3D Tomography*, pp. 97-121, sous la direction de Ali Mohammad-Djafari, ISTE / John Wiley, 2010.
  - « Déconvolution aveugle d'image, L. Blanc-Féraud, L. Mugnier & A. Jalobeanu », chap. 3 de *Problèmes inverses en imagerie et en vision*, pp. 107-132, sous la direction de Ali Mohammad-Djafari, Hermes, 2009.
  - « Des données à la connaissance de l'objet : le problème inverse », L. Mugnier, chap. 9, pp. 606–629 de *L'observation en astrophysique*, sous la direction de P. Léna, EDP Sciences, 2008.
  - « Inversion in optical imaging through atmospheric turbulence », L. M. Mugnier, G. Le Besnerais et S. Meimon, chapitre 10 de *Bayesian approach to inverse problems*, sous la direction de J. Idier, ISTE / John Wiley, 2008.
  - « Phase Diversity: a technique for Wave-Front Sensing and for Diffraction-Limited Imaging », L. Mugnier, A. Blanc et J. Idier, dans *Advances in Imaging & Electron Physics*, sous la direction de P. Hawkes, volume 141, chap. 1, pp. 1–76, Academic Press, 2006.
  - « Data processing in nulling interferometry: case of the DARWIN mission », L. M. Mugnier, E. Thiébaut et A. Belu, dans *Astronomy with High Contrast Imaging III*, EAS Publications Series, EDP Sciences, 2006.
  - « Problèmes inverses en imagerie optique à travers la turbulence », L. Mugnier et G. Le Besnerais, chapitre 10 de *Approche bayésienne pour les problèmes inverses*, pp. 241–270, sous la direction de J. Idier, Hermès, 2001 ;

- Trois brevets :
  - S. Demoustier, A. Brignon, J.-P. Huignard, L. Mugnier et J. Primot, *Source laser à recombinaison cohérente de faisceaux*, Brevet Thales déposé le 12 août 2005,
  - F. Cassaing et L. Mugnier, *Procédé et dispositif de mesure d'au moins une déformation d'une surface d'onde*, Brevet Onera déposé le 11 mai 2007,
  - F. Cassaing, I. Mocœur et L. Mugnier, *Procédé d'estimation d'au moins une déformation du front d'onde d'un système optique ou d'un objet observé par le système optique et dispositif associé*, Brevet Onera déposé le 19 juillet 2007 ;

- plus de 40 publications dans des revues internationales à comité de lecture ;

- plus de 90 communications dans des conférences avec actes.

Activités d'encadrement de doctorants

- Je suis l'encadrant principal de Leonardo Blanco *(90%)* : *restauration myope d'images 3D de la rétine*, Paris 7, 2009–2012, qui fait suite à celle de Guillaume Chenegros. Cette thèse s'effectue dans le cadre d'une collaboration avec le LESIA de l'Observatoire de Paris et est codirigée par François Lacombe, aujourd'hui chez Mauna Kea Technologies. À ce jour un article à paraître [A2, Annexe C page 215] et plusieurs communications dont l'une avec actes [C2].
- Je coencadre Marie Ygouf : *Calibration et traitement d'images 3D pour les systèmes d'imagerie à haut contraste*, Univ. Joseph Fourier, Grenoble, 2009–2012. Thèse dirigée par Jean-Luc Beuzit, coencadrée par David Mouillet et Thierry Fusco. À ce jour plusieurs communications avec actes [C1, C8].
- Je participe occasionnellement à l'encadrement de Sarah Dandy : *Analyse de front d'onde pour l'optique adaptative extrême et l'imagerie à haut contraste*, Paris 7, 2008–1011. Thèse encadrée par Jean-François Sauvage et dirigée par Thierry Fusco. À ce jour une communication avec actes [C12].
- J'ai été l'encadrant principal d'Alberto Cornia *(90%)* : *Traitement d'images différentielles pour la détection de planètes extra-solaires*, Paris 7, 2010 [T1]. Cette thèse a été codirigée par Gérard Rousset (Obs. Paris / LESIA). Publications à ce jour : [A10, Annexe E page 233] et plusieurs communications. Alberto Cornia est aujourd'hui Ingénieur R&D chez Evolution Energie, à Arcueil.
- J'ai été l'encadrant principal de Guillaume Chenegros *(90%)* : *Restauration d'images de la rétine corrigées par optique adaptative*, Paris 7, 2008 [T2]. Thèse codirigée par François Lacombe. Publications : [A18, Annexe H page 269] et plusieurs communications. Guillaume Chenegros est aujourd'hui Maître de Conférences à l'Université Paris 6.
- J'ai participé à l'encadrement *(30%)* d'Isabelle Mocœur : *Analyse de front d'onde en plan focal: développement d'algorithmes temps-réel et application au cophasage de télescopes multipupilles imageurs*, Paris 11, 2008 [T3]. Thèse encadrée par Frédéric Cassaing et dirigée par Denis Mourard (OCA). Publications : un brevet [B1], deux articles [A8], [A9, Annexe D page 229] et plusieurs communications. Isabelle Mocœur est aujourd'hui Ingénieur Expert Technique en Optronique à la DGA, Arcueil.
- J'ai participé à l'encadrement de Jean-François Sauvage *(30%)* : *Calibration et méthodes d'inversion en imagerie haute dynamique pour la détection directe d'exoplanètes*, Paris 7,

2007 [T4]. Thèse codirigée par Anne-Marie Lagrange et Gérard Rousset. Publications : [A5], [A15, Annexe G page 257], et deux communications. Jean-François Sauvage est aujourd'hui un collègue de l'équipe DOTA/HRA.

– J'ai participé à l'encadrement de Sébastien Demoustier *(10%)* : *Recombinaison cohérente de fibres laser*, Paris 11, 2006 [T5]. Thèse dirigée par Jean-Louis Meyzonnette. Publications communes : un brevet [B3] et une communication. Sébastien Demoustier est aujourd'hui Manager des Propositions et Programmes chez Thales/TRT, Palaiseau.

– J'ai été l'encadrant principal de Serge Meimon *(70%)* : *Reconstruction d'images astronomiques en interférométrie optique*, Paris 11, 2005 [T6]. Thèse coencadrée par Guy Le Besnerais (ONERA/DTIM), et dirigée par Guy Demoment (Paris 11). Publications : [A22, A24, A16, A14], [A7, Annexe F page 243] et cinq communications. Serge Meimon est aujourd'hui un collègue de l'équipe DOTA/HRA.

– J'ai été le responsable ONERA de l'encadrement de Damien Gratadour *(40%)* : *Restauration d'images astronomiques en optique adaptative, appliquée à l'étude des noyaux actifs de galaxie avec NAOS*, Paris 7, 2005 [T7]. Cette thèse s'est effectuée dans le cadre d'une collaboration avec l'Observatoire de Paris-Meudon et a été dirigée par Daniel Rouan. Publications : [A26, A20] et [A23, Annexe I page 279]. Damien Gratadour est aujourd'hui Maître de Conférences à l'Université Paris 7.

– J'ai participé à l'encadrement de la thèse de Fabien Baron *(25%)* : *Définition et test d'un capteur de cophasage sur télescope multipupilles : application à la détection d'exoplanètes et à l'observation de la Terre*, Paris 6, 2005 [T8]. Thèse encadrée par Frédéric Cassaing et dirigée par Jean Gay (OCA). Publications : [A17] et plusieurs communications. Fabien Baron est actuellement en CDD à l'Université du Michigan.

– J'ai participé à l'encadrement de Brice Le Roux *(20%)* : *Analyse de front d'onde et commande en optique adaptative multiconjuguée*, Univ. Nice, 2003 [T9]. Thèse encadrée par Jean-Marc Conan (ONERA) et dirigée par Julien Borgnino (Univ. Nice). Publications : [A28] et plusieurs communications. Brice Leroux est actuellement Maître de Conférences à l'Université Aix-Marseille 1.

– J'ai été l'encadrant principal d'Amandine Blanc *(60%)* : *identification de réponse impulsionnelle et restauration d'images : apports de la diversité de phase*, Paris 11, 2002 [T10]. Cette thèse a été codirigée par Jérôme Idier (CNRS). Outre des publications dans des revues à comité de lecture [A31, A30, A29], [A21, Annexe J page 289], cette thèse a donné lieu à une contribution dans un ouvrage de synthèse [L6, Annexe B page 133]. Amandine Blanc est actuellement Professeur dans l'enseignement secondaire.

– J'ai participé régulièrement à l'encadrement de Thierry Fusco *(30%)* : *Imagerie à haute résolution en dehors du domaine isoplanétique*, Univ. Nice, 2000 [T11]. Cette thèse a été encadrée par Jean-Marc Conan et codirigée par Julien Borgnino et Gérard Rousset. Publications communes : [A39, A37, A36, A33], [A27, Annexe K page 301]. Thierry Fusco est aujourd'hui un collègue de l'équipe DOTA/HRA.

– J'ai également participé occasionnellement à l'encadrement de Ludovic Meynadier *(≤10%)* : *Analyse de surface d'onde pour le contrôle actif d'un télescope spatial*, Univ. Nice, 1997 [T12]. Thèse encadrée par Vincent Michau et codirigée par Claude Aime (Univ. Nice) et Gérard Rousset. Publication : [A38]. Ludovic Meynadier est actuellement responsable internet des Éditions Lavoisier.

Participation à des jurys de Thèse de Doctorat et d'Habilitation à Diriger des Recherches

La liste des jurys auxquels j'ai participé, sans avoir participé à l'encadrement de la thèse, est donnée ci-dessous, avec mention du titre de la thèse lorsqu'elle ne figure pas dans la bibliographie et, entre parenthèses, de mon rôle dans le jury (invité, examinateur ou rapporteur) :

– mars 2011 : thèse de Daniel Burke intitulée *optimal post processing of AO corrected astronomical images: application to faint companion detection and characterisation*, dirigée par Nicholas Devaney, Université de Galway, Irlande (*extern examiner* c'est-à-dire seul membre du jury de trois personnes externe au laboratoire).

– novembre 2010 : thèse de Stéphanie Renard intitulée *Imager les zones de formation des planètes atour des étoiles jeunes dans le cadre de la reconstruction d'images pour le VLTI*, dirigée par F. Malbet et É. Thiébaut, Université Joseph Fourier, Grenoble (rapporteur).

– septembre 2009 : thèse de Raphaël Galicher intitulée *Étude de techniques d'imagerie à haut contraste basées sur la cohérence*, dirigée par G. Rousset, Université Paris 7 (examinateur).

– novembre 2007 : thèse de Xavier Rondeau intitulée *Imagerie à travers la turbulence : mesure inverse du front d'onde et centrage optimal*, dirigée par R. Foy, Université Claude Bernard/Lyon I (rapporteur).

– décembre 2005 : Habilitation à Diriger des Recherches de Mireille Guillaume intitulée *Imagerie multicomposantes : de l'estimation de paramètres à la reconstruction*, Université Aix-Marseille III (examinateur).

– novembre 2000 : thèse de Stéphanie Cabanillas intitulée *Approche bayésienne pour la reconstruction d'images astronomiques à partir de séquences d'images à faible niveau de photons*, dirigée par Ph. Réfrégier, Université Aix-Marseille III (examinateur).

Pour mémoire, voici la liste des jurys auxquels j'ai participé comme encadrant :

- décembre 2010 : thèse d'Alberto Cornia, codirigée par G. Rousset et moi-même, École Doctorale d'Astronomie et d'Astrophysique d'île-de-France (codirecteur) [T1] ;
- novembre 2008 : thèse de Guillaume Chenegros, codirigée par F. Lacombe et moi-même, Université Paris 7 (codirecteur) [T2].
- juillet 2008 : thèse d'Isabelle Mocœur, dirigée par D. Mourard, Université Paris 11 (invité) [T3].
- décembre 2007 : thèse de Jean-François Sauvage, dirigée par A.-M. Lagrange, Université Paris 7 (invité) [T4].
- novembre 2005 : thèse de Serge Meimon, codirigée par G. Demoment et moi-même, Université Paris 11 (codirecteur) [T6].
- novembre 2005 : thèse de Damien Gratadour, codirigée par D. Rouan et moi-même, École Doctorale d'Astronomie et d'Astrophysique d'île-de-France (codirecteur) [T7].
- juillet 2002 : thèse d'Amandine Blanc, codirigée par J. Idier et moi-même, Université Paris 11 (codirecteur) [T10].

Actions d'enseignement

Mon expérience dans l'enseignement a commencé il y a une douzaine d'années et se poursuit sans interruption depuis lors, avec un volume actuel d'une vingtaine d'heures par an. Depuis la rentrée 2010-2011, je suis également responsable d'une mineure en troisième année de Supélec, campus de Metz, ce qui représente 24 h de cours consacrés à la HRA.

Ces activités d'enseignement, conjuguées avec mes participations à des ouvrages pédagogiques [L3, L4, L5, L6, L7], me permettent de contribuer à la formation des jeunes scientifiques et d'exprimer mon goût croissant pour la pédagogie.

- Institut d'Optique Graduate School : je donne tous les ans depuis 1999, avec mon collègue Jean-Marc Conan, un cours en 3e année sur la restauration d'images dégradées par la turbulence (12 h de cours et travaux dirigés) ;
- École thématique Imagerie à Très Haute Dynamique et détection d'exoplanètes (ITHD '05) : j'ai donné un cours à cette école organisée en 2005 par l'Université de Nice, intitulé « Traitement des données en interférométrie coronographique : cas de la mission DARWIN » ;
- École d'été *Physics and Ophtalmology: opening a window on the Living* : j'ai donné deux cours à cette école organisée en 2005 par le réseau européen *Sharp Eye*, l'un sur l'optique adaptative et l'autre sur les techniques de déconvolution, intitulés respectivement « High Resolution Imaging: Adaptive Optics, wavefront sensing and MultiConjugate AO » et « Beyond adaptive optics: deconvolution techniques » ;
- Master 2 Recherche *Astrophysique et Méthodes Associées* de l'E. D. d'Astronomie et Astrophysique d'Île-de-France : je donne depuis 2006 un cours, assortis de travaux dirigés, sur les problèmes inverses en imagerie astronomique à haute résolution angulaire (2×4 h) ;
- Supélec Metz (3e année) : je donne depuis 2009 un cours sur les techniques et traitements en haute résolution angulaire (2×3 h) ;
- Master 2 Pro de l'E. D. d'Astronomie et Astrophysique d'Île-de-France : je donne depuis 2009, avec mes collègues Jean-Marc Conan et Cyril Petit, un cours sur l'estimation et la commande en optique adaptative (3×3 h) ;
- École thématique *High Angular Resolution* 2010 : j'ai donné un cours sur les problèmes inverses et la déconvolution dans cette école organisée par Guy Perrin, de l'Observatoire de Paris-Meudon.

Partenariats scientifiques, groupes de travail

– Je suis membre du conseil scientifique du Groupement d'Intérêt Scientifique Partenariat Haute résolution Angulaire Sol-Espace (GIS PHASE) liant l'ONERA et l'Observatoire de Paris, qui vient d'être étendu à l'IPAG et au LAM.

– Je suis membre du conseil scientifique du Jean-Marie Mariotti Center (JMMC), initiative française pour la coordination des activités de traitement des données issues d"interféromètres optiques.

– J'ai coordonné pour l'ONERA la mise en place d'une convention de recherche liant l'ONERA, l'Observatoire de Paris, l'Hôpital des XV-XX et l'INSERM, dans le cadre de l'Institut de la Vision récemment créé. Je suis membre du conseil scientifique de ce partenariat, intitulé ŒIL-HRS, qui porte sur l'imagerie à haute résolution spatiale de la rétine.

– J'ai initié et participé à la mise en place d'un partenariat avec le Laboratoire des Signaux et Systèmes (laboratoire CNRS/Univ. Orsay/Supélec) sur l'identification de réponse impulsionnelle à l'aide de la diversité de phase, concrétisée par la codirection ONERA/CNRS de la thèse d'A. Blanc.

– Je suis représentant de l'ONERA dans le Centre de Compétence Technique du CNES consacré au Traitement du Signal et des Images (CCT TSI).

– Je fais partie du groupe de travail *Mathématiques appliquées* du DOTA, et j'ai organisé dans ce cadre en 2008 une journée consacrée aux problèmes inverses combinant formation et ateliers.

– J'ai fait partie du SOC de l'École thématique *High Angular Resolution* 2010 organisée par Guy Perrin, de l'Observatoire de Paris-Meudon.

– J'ai fait partie du SOC de la deuxième conférence internationale *the spirit of Bernard Lyot*, organisée en 2010 par l'Observatoire de Paris-Meudon et l'ONERA.

Deuxième partie

Exposé synthétique des recherches et perspectives

Acronymes

ADI	*Angular Differential Imaging*
ANDROMEDA	*ANgular DiffeRential OptiMal Exoplanet Detection Algorithm*
ANR	Agence Nationale de la Recherche
ASO	Analyseur de Surface d'Onde
BOA	Banc d'Optique Adaptative de l'ONERA
BRISE	Banc Reconfigurable d'Interférométrie sur Sources Étendues de l'ONERA
CCD	*Charged Coupled Device*
CIC	Centre d'Investigations Cliniques (de l'Hôpital des Quinze-Vingts)
CNES	Centre National d'Études Spatiales
CONICA	*COudé Near-Infrared CAmera*
DECASO	DEConvolution par Analyse de Surface d'Onde
DGA	Direction Générale de l'Armement
DMLA	Dégénérescence Maculaire Liée à l'Âge
DOTA	Département Optique Théorique et Appliquée
DSP	Densité Spectrale de Puissance
DWARF	*DarWin AstRonomical Fringe sensor*
ELT	*Extremely Large Telescope*
E-ELT	*European ELT*
EII	*European Interferometry Initiative*
ELI	*Extreme Light Infrastructure*
EPICS	*Exo-Planets Imaging Camera and Spectrograph (of the E-ELT)*
EQM	Écart Quadratique Moyen
ESA	*European Space Agency*
EUCLID	*EUropean Cooperation for the Long term In Defense*
ESO	*European Southern Observatory*
FEP	Fonction d'Etalement de Point
FFT	*Fast Fourier Transform*

FITTEST	*Filter for Interferometric Test data and Terrestrial planet Exploration Software Tools*
FP7	*Seventh Framework Programme*
FTA	Fonction de Transfert Atmosphérique
FTM	Fonction de Transfert de Modulation
FTO	Fonction de Transfert Optique
FTS	Fonction de Transfert de Speckle
FUSCHIA	*Fast Unambiguous Sensors for CopHasing Interferometric Arrays*
GEO	*Geostationnary Earth Orbit*
GIS	Groupement d'Intérêt Scientifique
GPI	*Gemini Planet Imager*
HRA	Haute Résolution Angulaire
HS	Hartmann-Shack
IDL	*Interactive Data Langage*
IFS	*Integral Field Spectrograph*
INCA	Identification Nuit-jour de Cibles Aériennes
INOVEO	INstrumentation à Optique adaptatiVE pour l'Ophtalmologie
IPAG	Institut de Planétologie et d'Astrophysique de Grenoble
IR	Infra-Rouge
JMMC	*Jean-Marie Mariotti Center*
JWST	*James Webb Space Telescope*
LAM	Laboratoire d'Astrophysique de Marseille
LAOG	Laboratoire d'AstrOphysique de Grenoble (désormais fusionné dans l'IPAG)
LESIA	Laboratoire d'Études Spatiales et d'Instrumentation en Astronomie
MAP	Maximum *a posteriori*
MIRA	*Multi-aperture Image Reconstruction Algorithm*
MISTRAL	*Myopic Iterative STep-preserving Restoration ALgorithm*
MV	Maximum de Vraisemblance
NAOS	*Nasmyth Adaptive Optics System* , première optique adaptative du VLT
NASA	*National Aeronautics and Space Administration*
NCPA	*Non-Common Path Aberrations*
OA	Optique Adaptative
OAMC	Optique Adaptative Multi-Conjuguée
OCT	*Optical Coherence Tomography*
ONERA	Office National d'Études et de Recherches Aérospatiales

PCRD	Programme-Cadre de Recherche et de Développement
PHASE	Partenariat Haute-résolution Angulaire Sol-Espace
RESSP	*Reconstruction of Exo-Solar System Properties*
RI	Réponse Impulsionnelle
RNTS	Réseau National des Technologies pour la Santé
RSB	Rapport Signal à bruit
SAXO	*Sphere Ao for eXoplanet Observation*
SC	Senseur de Cophasage
SED	*Spectral Energy Distribution*
SICLOPE	*Single Image with Calibrated Local Offset for Phase Estimation*
SOC	*Scientific Organizing Committee*
SOO	Synthèse d'Ouverture Optique
SOTISE	Satellite d'Observation de la Terre par Interférométrie sur Scènes Étendues
SPHERE	*Spectro-Polarimetry High-contrast Exoplanet REsearch*
TAS	Thales Alenia Space
TPF-I	*Terrestrial Planet Finder Interferometer*
VLT	*Very Large Telescope*
VLTI	*Very Large Telescope Interferometer*
WISARD	*Weak-phase Interferometric Sample Alternating Reconstruction Device*

Chapitre 1

Résumé des recherches et plan

Mes activités de recherche portent sur les techniques d'imagerie optique à haute résolution, et plus particulièrement sur les méthodes, dites d'*inversion*, de traitement de données issues de ces techniques afin d'estimer des paramètres physiques d'intérêt. Mes travaux se situent donc à la croisée des chemins entre l'imagerie optique et le traitement du signal et des images.

Mes travaux de thèse ont été réalisés à l'École Nationale Supérieure des Télécommunications (ENST), à Paris, de 1989 à 1992. Ils ont permis d'obtenir la première reconstruction tridimensionnelle (3D) d'un objet opaque à partir d'un hologramme en lumière spatialement incohérente enregistré numériquement, dit hologramme conoscopique.

J'ai été embauché à l'ONERA en 1994 dans l'équipe alors dénommée *Imagerie Optique à Haute Résolution* et dirigée par Marc Séchaud, aujourd'hui unité de recherche *Haute Résolution Angulaire* (HRA) animée par Vincent Michau et faisant partie du Département Optique Théorique et Appliquée (DOTA). J'y ai bénéficié de la grande diversité de compétences de l'équipe et du goût de ses membres pour les échanges scientifiques ; j'en ai profité pour développer, en collaboration avec mes collègues, des techniques d'inversion de données pour une large palette d'instruments et d'applications.

Ces travaux sont dans la suite classés en deux familles, l'étalonnage d'instrument d'une part, l'imagerie d'autre part – même si ce classement est quelque peu arbitraire du fait que les deux familles ont une intersection non vide et d'ailleurs particulièrement intéressante, à savoir l'imagerie à l'aide d'instruments imparfaitement étalonnés.

Les domaines à l'origine de ces travaux sont l'astronomie depuis le sol ou l'espace, l'observation de la Terre, et plus récemment l'imagerie de la rétine.

1.1 Étalonnage d'instrument

Mes apports sur l'étalonnage d'instrument concernent avant tout la technique de diversité de phase pour l'estimation d'aberrations, instrumentales ou turbulentes. Ils incluent :

- des travaux méthodologiques : développement d'un nouvel estimateur, dit estimateur marginal, pour la technique de diversité de phase, et étude du comportement asymptotique de l'estimateur habituellement utilisé [L6, Annexe B page 133], [A29], [A21, Annexe J page 289] ;
- une contribution à des évaluations expérimentales de la diversité de phase [A38, A31, A30], [C51, C39] ;
- des travaux algorithmiques : développement d'un algorithme potentiellement temps-réel pour la diversité de phase [A9, Annexe D page 229] [A8] [B1] et pour le *phase retrieval* [A4] ;
- une extension de cette technique pour l'imagerie à haute dynamique permettant de la faire fonctionner en ligne c'est-à-dire pendant une observation astronomique et de n'estimer que les aberrations statiques de l'instrument [A15, Annexe G page 257] ;
- une adaptation de cette technique pour la mise en phase de lasers fibrés [B3].

Mes apports concernent également des variations autour de la technique de diversité de phase n'utilisant qu'une seule image au voisinage du plan focal, dans des cas particuliers présentant un intérêt pratique avéré : analyse de front d'onde sur objet ponctuel observé par un télescope monopupille [B2] et algorithme temps-réel associé [A4], analyse de front d'onde sur objet ponctuel observé par un interféromètre à pupille de type non redondant [A17], objet de type scène terrestre présentant des ruptures franches entre des régions uniformes (ou « bords de plages ») [C72, C68].

Ils concernent enfin, en collaboration avec des collègues de l'ONERA, quelques sujets que je n'aborderai guère dans ce document : l'analyse de front d'onde anisoplanétique [A37, A33], l'estimation de réponse impulsionnelle longue pose isoplanétique ou non [A36] ainsi que la commande [A28], pour un analyseur de front d'onde linéaire tel qu'un Hartmann-Shack.

1.2 Imagerie

Depuis mon entrée à l'ONERA, j'ai développé, avec des étudiants ou en collaboration avec des collègues, des méthodes de traitement pour un éventail de modalités d'imagerie, listées brièvement ci-dessous et détaillées dans le document :

- imagerie 2D mono-télescope :
  - déconvolution myope d'images courtes poses turbulentes non corrigées par OA mais assorties de mesures de front d'onde [A34, Annexe L page 317] ; ici le terme myope signifie que l'on utilise les mesures de front d'onde tout en modélisant explicitement le fait qu'elles ne donnent pas une estimée idéale de la réponse instantanée de l'instrument ;
  - recalage de piles d'images corrigées par OA en vue de leur déconvolution [A23, Annexe I page 279] ;
  - déconvolution myope d'images longues poses corrigées par OA [A40, A39] [A27, Annexe K page 301] ; ici le terme myope signifie que l'on modélise explicitement le fait que la mesure de réponse impulsionnelle longue pose, supposée obtenue par ailleurs, n'est

pas exactement celle qui a donné lieu à l'image. Cette méthode, appelée MISTRAL, a été appliquée à de nombreuses données expérimentales, dont [A32, A25, A20, A13] ;
  - détection d'exoplanètes par imagerie différentielle angulaire à partir d'une série d'images coronographiques corrigées par OA – méthode ANDROMEDA pour le projet SPHERE [A10, Annexe E page 233] ;
- imagerie 3D : déconvolution myope d'images 3D en vue de l'imagerie rétinienne corrigée par OA : ces travaux incluent une extension à trois dimensions de la technique de diversité de phase [A18, Annexe H page 269] ;
- imagerie interférométrique :
  - reconstruction d'images à partir des mesures d'un interféromètre optique au sol type VLTI – méthode WISARD [A7, Annexe F page 243] [A16, A14] ;
  - reconstruction d'images à partir des mesures d'un interféromètre coronographique type DARWIN – méthode FITTEST [L5, Annexe A page 115] [C36].

La grande majorité de ces méthodes a été validée sur des données expérimentales, qui constituent l'ultime juge de leur intérêt. Les deux seules exceptions – pour l'instant – sont la détection d'exoplanètes par imagerie différentielle angulaire ou ADI développée pour SPHERE et la reconstruction d'images développée pour DARWIN. Dans les deux cas nous ne disposons pas encore de données expérimentales issues de ces instruments puisqu'ils ne sont pas encore opérationnels... On peut d'ailleurs noter que notre technique d'ADI a été tout récemment validée sur des images expérimentales provenant de l'instrument opérationnel NAOS-CONICA [C14], dans l'attente de données issues de SPHERE. Il ne faut pas croire pour autant que ces méthodes de traitement aient été développés prématurément : au contraire, un intérêt de ces méthodes lorsqu'elles sont développées suffisamment tôt, comme ici, est de pouvoir « reboucler » sur la spécification de l'instrument ou de son étalonnage en regard des performances finales visées et ainsi d'optimiser la conception globale de la chaîne d'acquisition et de traitement.

C'est selon cette conception que j'ai développé une méthode d'optimisation de l'agencement relatif des télescopes ou *configuration pupillaire* d'un interféromètre imageur [A41, Annexe M page 329], méthode qui prend en compte la qualité finale des images après restauration. J'ai appliqué cette méthode pour la conception et la simulation de bout en bout d'interféromètres imageurs pour l'observation de la Terre lors de deux projets que j'ai par ailleurs dirigés, l'un durant sa dernière année (projet EUCLID), l'autre pendant toute sa durée (projet financé par la DGA).

1.3 Plan

Le plan du reste du document est le suivant : le chapitre 2 présente l'intérêt et quelques caractéristiques importantes de l'inversion de données, puis rappelle quelques éléments de formation d'image et d'analyse de front d'onde. Ensuite, le chapitre 3 propose une synthèse de mes contributions à l'étalonnage d'instruments en HRA et le chapitre 4 résume mes travaux en imagerie, plus précisément en restauration d'images, en détection et en reconstruction d'images.

Enfin, le chapitre 5 dégage des perspectives de recherches. La bibliographie suit immédiatement le chapitre de perspectives et commence page 89. Elle contient des références dont je ne suis pas coauteur ainsi que quelques rapports ONERA.

La liste de mes publications et des thèses que j'ai encadrées ou coencadrées constitue la partie III, page 93 et suivantes. Afin de faciliter la lecture du texte, les contributions à des livres sont préfixées par la lettre L, les brevets par B, les articles par A, les thèses par T et les communications par C.

Les annexes, constituées de quelques contributions à des ouvrages de synthèse et d'articles particulièrement significatifs, sont rassemblées dans la partie IV.

Chapitre 2

Introduction

Dans ce chapitre sont introduits d'abord l'intérêt et quelques caractéristiques importantes, à mon avis, de l'inversion de données. Puis, dans les sections 2.2 et 2.3, quelques rappels sont effectués sur la formation d'image et l'analyse de front d'onde, qui constituent mon terrain de jeu pour l'inversion.

2.1 Inversion de données

En imagerie par optique adaptative les données sont des images (légèrement) floues et bruitées de l'objet d'intérêt, et la restauration de ces images permet d'en améliorer sensiblement l'exploitation et l'interprétation.

Souvent, contrairement au cas de la restauration d'image évoqué ci-dessus, les données fournies par l'observation ne « ressemblent » pas à la grandeur physique d'intérêt, mais lui sont liées par des lois physiques connues. Un exemple représentatif est la reconstruction d'image en interférométrie optique, coronographique ou non : les données sont tout simplement inexploitables sans un traitement approprié.

Dans le cas d'école d'un analyseur de front d'onde de Hartmann-Shack mesurant des aberrations (propres à un télescope ou dues à la turbulence), la grandeur physique d'intérêt est la phase dans la pupille du télescope, qui représente les aberrations. Les données sont les pentes moyennes de la phase sur chaque sous-pupille du Hartmann-Shack selon deux directions perpendiculaires, que l'on peut concaténer dans un vecteur de pentes mesurées. Le calcul des pentes connaissant la phase est un *problème direct*, classique en Physique, qui nécessite le choix d'un modèle de formation des données[1] ou *modèle direct*.

Estimer ou « remonter à » la phase à partir des pentes est le *problème inverse* correspondant, qui nécessite d'inverser le modèle de formation des données. Le choix de ce modèle direct n'est pas toujours uniquement dicté par la Physique, il peut résulter d'un compromis entre le souhait d'une modélisation fine des données et le coût lié à l'inversion. Dans l'exemple élémentaire

[1] l'analogue pour des mesures quelconques du modèle de formation des images pour un instrument imageur.

du Hartmann-Shack, on peut envisager d'utiliser un modèle de propagation diffractif lors de la conception de l'analyseur mais on se restreindra à un modèle d'optique géométrique pour l'estimation des aberrations dans le calculateur temps-réel d'une optique adaptative, car celui-ci conduit à un modèle de données linéaire (en fonction des aberrations à mesurer), d'inversion aisée.

Plus radicalement, le choix même de ce que *l'on considère comme* les données résulte de ce type de compromis et requiert à la fois une compréhension de la physique de la formation des données et des compétences en inversion de données pour évaluer les impacts du choix du modèle direct. Dans ce même exemple du Hartmann-Shack, on pourrait – et cela a été fait – considérer que les données sont les imagettes au foyer de chaque micro-lentille, mais ce choix a pour prix un traitement non linéaire bien plus lourd. Trois cas réels traités dans la suite illustrent bien cet aspect important du *choix de modèle direct* :

- en imagerie coronographique pour la détection d'exoplanètes j'ai choisi de considérer comme mesures des différences d'images brutes dans lesquelles la planète éventuelle s'est déplacée du fait de la rotation de champ [A10, Annexe E page 233]. Ceci a permis d'éviter l'estimation délicate de l'image coronographique de l'étoile, qui n'est pas constante au cours de la nuit et dont on ne possédait pas de modèle analytique exploitable au début de ces travaux ;
- en reconstruction d'image pour l'interférométrie optique depuis le sol nous avons choisi de considérer comme mesures des visibilités complexes compatibles avec les visibilités carrées et les clôtures de phase constituant les données brutes [A7, Annexe F page 243]. Ceci a permis d'aboutir à un modèle de données dont l'inversion est – à turbulence donnée – associée à la minimisation d'un critère convexe, donc plus aisée ;
- en reconstruction d'image pour l'interférométrie coronographique depuis l'espace nous avons choisi de considérer comme mesures des combinaisons linéaires des données brutes qui éliminent toute composante paire de la scène observée (nuages zodiacal et exo-zodiacal, fuites stellaires, émission instrumentale) [L5, Annexe A page 115]. Ceci a permis de ne rechercher, lors de la reconstruction, que les éventuelles exoplanètes.

On le verra dans la suite, l'inversion peut souvent tirer parti des connaissances statistiques sur les incertitudes de mesures, que l'on modélise généralement comme des bruits. Un modèle direct complet inclut donc la modélisation de la formation des données jusqu'à la détection, voire jusqu'au support de stockage dans le cas de transmission des données avec compression. Il prend ainsi en compte le bruit photonique, le bruit du détecteur, le bruit de quantification du numériseur s'il n'est pas négligeable, le bruit éventuel de compression, etc.

Le traitement de données expérimentales, en Physique et en particulier en Astronomie, consiste donc essentiellement à résoudre un problème inverse, en pratique après une étape de *réduction* ou pré-traitement des données, dont le but est de corriger les défauts instrumentaux de façon à ce que les données puissent être correctement décrites par le modèle direct adopté.

Les méthodes d'inversion naïves ont souvent la caractéristique d'être « instables » au sens où l'inévitable bruit de mesure est amplifié de manière non contrôlée lors de l'inversion, et conduit

à une solution inacceptable. Dans ce cas où les données seules ne suffisent pas à obtenir une solution acceptable, il est nécessaire de développer des méthodes d'inversion plus sophistiquées dites *régularisées* qui incorporent des contraintes supplémentaires pour imposer à la solution une certaine régularité compatible avec nos connaissances *a priori* sur celle-ci.

Concevoir explicitement le traitement de données comme l'inversion d'un problème direct est généralement très fructueux. Cela oblige à modéliser l'ensemble du processus de formation de données pour le prendre en compte dans l'inversion. Cela permet aussi d'analyser telle ou telle méthode existante et en particulier d'en expliciter les hypothèses sous-jacentes. Cela permet alors de concevoir des méthodes tirant parti à la fois des connaissances sur le processus de formation des données et de celles que l'on a *a priori*, c'est-à-dire avant de faire les mesures, sur la grandeur physique d'intérêt. Le lecteur plus familier de la HRA que des problèmes inverses pourra trouver une introduction que j'espère pédagogique aux techniques modernes de résolution de ces problèmes dans ma contribution [L3] à l'ouvrage [1] dirigé par P. Léna.

Enfin, la conception de l'inversion comme une partie intégrante la chaîne complète acquisition & traitement permet d'optimiser la conception même de l'instrument. C'est ce que l'on nomme aujourd'hui la co-conception et que j'ai mis en œuvre dans les années 90 en interférométrie optique [A41, Annexe M page 329] pour l'optimisation du positionnement des télescopes. De la même manière en imagerie coronographique depuis le sol, le développement suffisamment précoce de la méthode de détection d'exoplanètes pour l'instrument IRDIS de SPHERE a permis récemment de quantifier l'impact d'erreurs d'étalonnage de l'instrument et donc de mieux évaluer la précision requise pour l'étalonnage [C14].

2.2 Éléments de formation des images

Le pouvoir de résolution théorique d'un télescope est limité par son diamètre. Pour un instrument réel, la présence d'aberrations optiques empêche souvent d'atteindre cette limite théorique, appelée (résolution de la) limite de diffraction. Ces aberrations peuvent provenir du télescope lui-même ainsi que du milieu de propagation des ondes lumineuses. Dans le cas de l'observation astronomique depuis le sol, ces aberrations sont avant tout dues à la turbulence atmosphérique. Plusieurs techniques ont été développées pour améliorer la résolution des instruments d'observation et s'affranchir des dégradations apportées par la turbulence.

Cette section rappelle, pour des lecteurs spécialistes de problèmes inverses peu familiers de la HRA, quelques notions élémentaires sur l'imagerie optique et en particulier sur les effets optiques de la turbulence. Puis les diverses techniques d'imagerie haute résolution à travers la turbulence sont passées en revue. Une introduction plus complète peut être trouvée dans ma contribution [L4] à l'ouvrage [2] dirigé par J. Idier.

2.2.1 Diffraction et fonction de transfert optique

La formation des images est bien décrite par la théorie scalaire de la diffraction, exposée en détail dans des ouvrages de référence comme [3, 4, 5] ; on y trouvera une introduction synthétique et moderne dans [6]. Elle peut être modélisée par une convolution, en tout cas à l'intérieur du domaine dit isoplanétique de l'instrument. Aux longueurs d'onde visibles, ce domaine est typiquement de l'ordre du degré quand on considère uniquement les aberrations propres d'un télescope et de quelques secondes d'arc ($1\,\text{arcsec} = 1/3\,600°$) pour un télescope observant l'espace à travers la turbulence.

La réponse impulsionnelle (RI) instantanée d'un télescope ou du système « télescope + atmosphère » est égale au module carré de la transformée de Fourier de l'amplitude complexe du champ $\psi = P\exp(j\varphi)$ présent à l'instant t dans la pupille P de l'instrument lorsque l'objet observé est ponctuel :

$$h_t(\boldsymbol{\xi}) = \left|\mathrm{TF}^{-1}\left(P(\lambda\boldsymbol{u})\,e^{j\varphi(\lambda\boldsymbol{u})}\right)\right|^2(\boldsymbol{\xi}) \tag{2.1}$$

où λ est la longueur d'onde de l'imagerie supposée quasi monochromatique. Cette réponse est de plus normalisée, par convention, à une intégrale unité. Dans l'expression (2.1), la transformée de Fourier est la transformation du champ effectuée par le télescope entre le plan pupille et le plan focal, et le module carré est dû à la détection quadratique, *i.e.*, en intensité. Le vecteur $\boldsymbol{\xi} = [\xi, \zeta]^T$ est constitué d'angles sur le ciel, en radians. Pour un télescope parfait et en l'absence de turbulence, P est constant dans la pupille et φ est nul[2]. Pour un télescope réel, les variations du champ $P\exp(j\varphi)$ sont dues à la fois aux aberrations propres du télescope et à celles introduites par la turbulence.

Dans ce qui suit, on supposera que P est simplement l'indicatrice de la pupille, c'est-à-dire que les variations d'intensité dans la pupille d'entrée sont négligeables ; cette hypothèse est généralement valide en imagerie astronomique et est appelée approximation de champ proche.

L'équation (2.1) indique que la fonction de transfert optique ou FTO est l'autocorrélation de $\psi = P\,e^{j\varphi}$ dilatée de l'inverse de la longueur d'onde, ce que l'on écrit :

$$\tilde{h}_t(\boldsymbol{u}) = P\,e^{j\varphi} \otimes P\,e^{j\varphi}(\lambda\boldsymbol{u}), \tag{2.2}$$

où $\boldsymbol{u}$ est la fréquence angulaire, en radians^{-1}. En l'absence d'aberrations c'est-à-dire lorsque la phase φ est nulle, la FTO est donc l'autocorrélation de la pupille P; elle a une fréquence spatiale de coupure égale à $D/\lambda\,\mathrm{rd}^{-1}$, où D est le diamètre de la pupille, et est strictement nulle au-delà. La résolution ultime d'un télescope (appelé parfois télescope monolithique par opposition aux interféromètres décrits plus bas) est donc limitée par son diamètre D. Celui-ci est limité par la technologie actuelle à une dizaine de mètres pour des télescopes au sol et à quelques mètres pour des télescopes embarqués sur satellite, par suite de contraintes d'encombrement et de poids. L'interférométrie optique ou synthèse d'ouverture optique (SOO) est une technique permettant de dépasser la limitation en résolution qui en résulte.

[2] La réponse impulsionnelle correspondante est appelée tache d'Airy.

2.2.2 Principe de l'interférométrie optique et fonction de transfert

Cette technique consiste à faire interférer les champs électromagnétiques reçus en chaque pupille d'un réseau de pupilles ; l'instrument qui en résulte est un interféromètre ou un télescope multi-pupilles. À ce jour, cette technique n'a été utilisée que sur des instruments au sol. Les pupilles peuvent soit être elles-mêmes chacune un télescope, comme en astronomie (VLTI, NPOI par ex.) soit être des segments d'un miroir primaire commun. Si les segments sont adjacents, comme pour le télescope Keck, on parle alors d'un télescope segmenté plutôt que d'un interféromètre, même si cet instrument en est conceptuellement un. En ce qui concerne les missions spatiales, des interféromètres sont prévus en astronomie, avec un télescope segmenté pour le JWST et un interféromètre à pupilles très espacées pour les missions TPF-I (Terrestrial Planet Finder Interferometer) de la NASA et DARWIN de l'ESA. Ils sont également envisagés pour des télescopes d'observation de la Terre à haute résolution [C57, C51].

Pour chaque paire (k, l) de pupilles, les données contiennent de l'information haute résolution à (ou autour de) la fréquence spatiale angulaire $\boldsymbol{B}_{k,l}/\lambda$, où $\boldsymbol{B}_{k,l}$ est le vecteur séparant les pupilles ou *base*. Cette fréquence spatiale peut être bien plus grande que la fréquence de coupure D/λ des pupilles individuelles.

Selon le type d'interféromètre et de recombinaison des faisceaux, on peut soit former et mesurer directement une image de l'objet d'intérêt (l'interféromètre est dit imageur), soit mesurer un ensemble discret de fréquences spatiales dudit objet (l'interféromètre peut être appelé « corrélateur » car il mesure la corrélation des champs électromagnétiques entre pupilles élémentaires [7]). Le lecteur intéressé par une typologie plus précise des instruments à SOO est invité à consulter [A35].

Pour un télescope monolithique comme pour un interféromètre, la fonction de transfert est l'autocorrélation de la pupille d'entrée (*cf.* Eq. (2.2), pourvu que, si l'interféromètre est de type corrélateur, on assimile les pupilles à des points. Pour un interféromètre à grande base, *i.e.*, lorsque les bases sont grandes devant le diamètre des pupilles individuelles – ce qui est généralement le cas des interféromètres corrélateurs – la différence entre imageur et corrélateur s'estompe du point de vue de l'information enregistrée dans les données. Les fonctions de transfert d'un télescope monolithique, d'un interféromètre imageur et d'un interféromètre corrélateur sont illustrées figure 2.1. Pour un interféromètre imageur, le traitement à réaliser est en bonne approximation une déconvolution, avec une réponse impulsionnelle toujours donnée par l'équation (2.1) mais plus irrégulière qu'avec un télescope monolithique du fait de la forme de la pupille.

Pour un interféromètre corrélateur le problème de traitement de données change de nature, puisqu'il s'agit désormais de reconstruire un objet à partir de coefficients de Fourier, problème appelé synthèse de Fourier (SF). C'est ce problème qui sera abordé dans la section 4.7.

Une façon intuitive de se représenter la formation des données dans un interféromètre à grande base est l'expérience des trous d'Young, dans laquelle la pupille de chaque télescope est un (petit) trou laissant passer la lumière qui vient de l'objet d'intérêt situé à grande distance. Chaque paire (k, l) de télescopes donne alors dans un plan focal un réseau de franges à la fréquence spatiale $\boldsymbol{B}_{k,l}/\lambda$. Le contraste et la position de ces franges, s'ils sont tous deux mesurés,

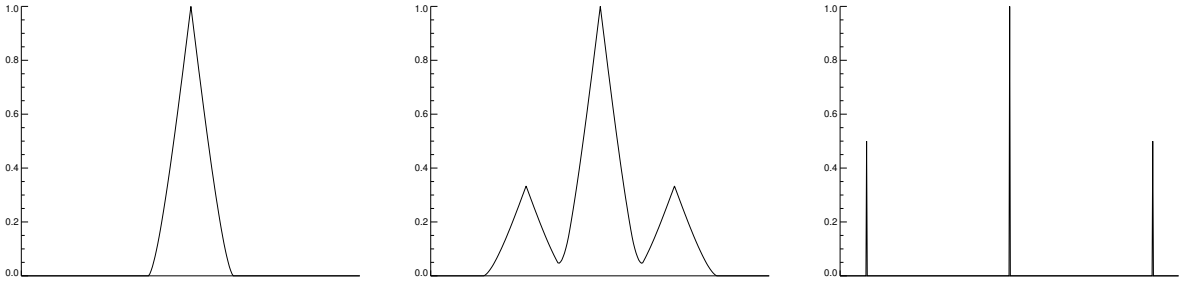


FIG. 2.1 – Coupes des fonctions de transfert d'un télescope monolithique (gauche), d'un interféromètre imageur à trois télescopes (milieu) et d'un interféromètre corrélateur à deux télescopes (droite).

peuvent être naturellement regroupés en un nombre appelé « visibilité complexe » qui donne, dans un cadre idéal et en l'absence de turbulence, la valeur normalisée à la fréquence nulle de la TF de l'objet $\boldsymbol{o}$ observé, soit $\tilde{\boldsymbol{o}}(\boldsymbol{B}_{k,l}/\lambda)/\tilde{\boldsymbol{o}}(\boldsymbol{0})$ (théorème de Van Cittert-Zernike [8, 6]).

2.2.3 Effets de la turbulence sur la formation des images

Turbulence et phase

Les inhomogénéités de température de l'air dans l'atmosphère engendrent des inhomogénéités de l'indice de réfraction de l'air, qui perturbent la propagation des ondes lumineuses à travers l'atmosphère. Ces perturbations se traduisent par des variations spatiales et temporelles de la phase pupillaire φ, qui doit alors être modélisée par un processus aléatoire. Dans cette section sont rappelés quelques résultats permettant de modéliser au second ordre la phase pupillaire turbulente. On fera l'hypothèse généralement bien vérifiée, au moins pour des échelles spatiales inférieures à la dizaine de mètres, que les variations aléatoires d'indice de l'air obéissent à la loi de Kolmogorov : elles suivent une loi de probabilité gaussienne, de moyenne nulle et de densité spectrale de puissance (DSP) proportionnelle à $|\boldsymbol{\nu}|^{-11/3}$, où $\boldsymbol{\nu}$ est la fréquence spatiale 3D [9].

Par intégration le long du chemin optique et dans le cadre de l'approximation de champ proche, on peut en déduire la statistique spatiale de la phase dans la pupille du télescope, pour une onde plane en entrée de l'atmosphère. La phase dans la pupille est gaussienne, car résultante de la somme de toutes les perturbations d'indice depuis la haute atmosphère jusqu'au sol [9]. La DSP de cette phase ne dépend que du paramètre r_0 et vaut [10] :

$$S_\varphi(\boldsymbol{\rho}) = 0{,}023\, r_0^{-5/3} \rho^{-11/3} \tag{2.3}$$

où $\boldsymbol{\rho}$ est la fréquence spatiale 2D dans la pupille, ρ son module, et r_0 le paramètre-clé quantifiant la force de la turbulence, appelé diamètre de Fried [11]. r_0 est d'autant plus petit que la turbulence est forte ; il dépend de la longueur d'onde et vaut typiquement $10\,\text{cm}$ dans le visible dans un relativement bon site.

Un temps typique d'évolution de la phase turbulente dans la pupille est le temps de bouillonnement τ du front d'onde, donné par le rapport entre l'échelle caractéristique r_0 de cette phase et l'écart-type Δv de la distribution des modules des vitesses de vent [12] :

$$\tau = r_0/\Delta v. \tag{2.4}$$

Pour $r_0 \approx 10\,\text{cm}$ et $\Delta v \approx 10\,\text{m.s}^{-1}$, on obtient $\tau \approx 10^{-2}\,\text{sec}$. On appellera donc longue pose une image correspondant à une intégration nettement plus longue que ce temps, et courte pose une image de temps d'intégration plus court. Pour un traitement complet de la statistique temporelle de la phase turbulente, on pourra consulter [13].

Imagerie longue pose

La FTO longue pose turbulente est le produit de la FTO dite statique $\tilde{h}^s$ du télescope sans atmosphère et d'une fonction de transfert de l'atmosphère $\tilde{h}^a$, de fréquence de coupure r_0/λ [9] :

$$\tilde{h}^{\text{lp}}(\boldsymbol{u}) \triangleq \langle \tilde{h}_t(\boldsymbol{u}) \rangle = \tilde{h}^s(\boldsymbol{u})\, \tilde{h}^a(\boldsymbol{u}) \;\text{ où } \tilde{h}^a(\boldsymbol{u}) = \exp\left\{-3{,}44\,(\lambda u/r_0)^{5/3}\right\}, \tag{2.5}$$

où la notation $\langle \cdot \rangle$ désigne une moyenne temporelle sur une durée arbitrairement longue. On voit donc que pour un télescope de grand diamètre $D \gg r_0$, la résolution en imagerie longue pose est limitée par la turbulence et n'est pas meilleure que celle d'un télescope de diamètre r_0.

Imagerie courte pose

Comme l'a remarqué Antoine Labeyrie [14], lorsque le temps de pose est assez court pour figer la turbulence (typiquement inférieur à 10 ms, *cf.* (2.4)) les images conservent des informations haute fréquence sous la forme de tavelures ou *speckles*, de taille typique λ/D et de position aléatoire. Ceci est illustré sur la figure 2.2, qui montre l'image simulée d'une étoile à travers la turbulence ($D/r_0 = 10$) en courte pose (à gauche) et en longue pose (à droite).

Il est possible de quantifier cette information haute fréquence présente dans les images courtes poses en évaluant la fonction de transfert de speckle (FTS), définie comme le moment d'ordre deux de la fonction de transfert optique instantanée, $\langle |\tilde{h}_t(\boldsymbol{u})|^2 \rangle$. Pour un télescope de grand diamètre ($D \gg r_0$), et moyennant une approximation sur la statistique de la turbulence, on peut trouver une expression approchée de la FTS [9] :

$$\langle |\tilde{h}_t(\boldsymbol{u})|^2 \rangle \approx \langle \tilde{h}_t(\boldsymbol{u}) \rangle^2 + 0{,}435\,(r_0/D)^2\, \tilde{h}_0^s(\boldsymbol{u}) \tag{2.6}$$

où $\tilde{h}_0^s$ est la fonction de transfert d'un télescope parfait (*i.e.*, dépourvu d'aberrations) de diamètre D.

Cette expression permet de décrire la FTS comme la somme du carré de la fonction de transfert longue pose, qui est basse fréquence (BF), et d'une composante haute fréquence (HF) qui s'étend jusqu'à la fréquence de coupure du télescope avec une atténuation proportionnelle à $(D/r_0)^2$. Il est donc possible, par un traitement d'un ensemble d'images courtes poses plus judicieux qu'une simple moyenne, de restituer une image haute résolution de l'objet observé.

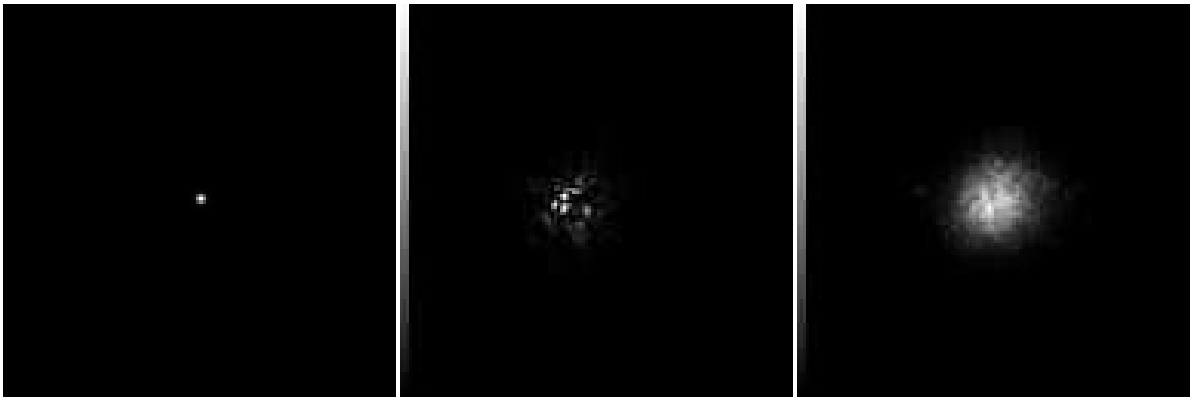


FIG. 2.2 – Images d'une étoile simulées sans turbulence atmosphérique (gauche) et à travers la turbulence (courte pose au centre et longue pose à droite). Le rapport D/r_0 vaut 10 ; l'échantillonnage des images respecte la condition de Shannon.

Cas d'un interféromètre à grande base

L'équation (2.5) s'applique quelle que soit la forme de la pupille de l'instrument, donc en particulier à un interféromètre. En longue pose le contraste des franges mesurées pour la base $\boldsymbol{B}_{k,l}/\lambda$ est donc multiplié par $\tilde{h}^a(\boldsymbol{B}_{k,l}/\lambda)$ et atténué si fortement qu'il rend la mesure de $\tilde{\boldsymbol{o}}(\boldsymbol{B}_{k,l}/\lambda)$ inutilisable.

En courte pose, pour un interféromètre dont chaque pupille est de diamètre inférieur au diamètre de Fried r_0 ou est corrigée de la turbulence par optique adaptative (voir 2.2.4), l'impact de la turbulence sur les mesures de l'interféromètre est aisé à modéliser : dans l'analogie des trous d'Young évoquée plus haut, chaque trou k ajoute un déphasage (ou piston) $\varphi_k(t)$ à l'onde le traversant, du fait des aberrations introduites par la turbulence devant cette pupille. Les interférences entre deux pupilles k et l sont donc déphasées du « piston différentiel » $\varphi_l(t)-\varphi_k(t)$ qui se traduit, en courte pose, par un déplacement aléatoire des franges sans atténuation du contraste. L'atténuation du contraste en longue pose résulte du moyennage de ces déplacements aléatoires. La section 4.7 présentera des techniques de moyennage permettant de s'affranchir des pistons différentiels.

La fonction de transfert en courte pose, à la fréquence $\boldsymbol{B}_{k,l}/\lambda$, s'écrit :

$$\tilde{h}_t(\boldsymbol{B}_{k,l}/\lambda) = \eta_{k,l}(t)\, e^{j(\varphi_l(t)-\varphi_k(t))} \tag{2.7}$$

où $\eta_{k,l}(t)$ est un nombre souvent appelé « visibilité instrumentale ». En l'absence des nombreuses sources potentielles de perte de visibilité (perturbations de front d'onde résiduelles sur chaque télescope, tilts différentiels entre télescopes, effets de polarisation différentielle, largeur spectrale non nulle, etc.), $\eta_{k,l}(t)$ vaut l'inverse du nombre de pupilles interférant simultanément (équ. 2.2 en considérant que P est constitué de Diracs). En pratique cette visibilité instrumentale est étalonnée sur une étoile réputée non résolue par l'interféromètre avant d'observer l'objet d'intérêt,

et est compensée lors des prétraitements des données brutes. En prenant en compte cet étalonnage on peut donc remplacer $\eta_{k,l}(t)$ par 1 dans l'équation (2.7).

Notons que la base $\boldsymbol{B}_{k,l}$ de la mesure entre les pupilles k et l dépend du temps : en effet, la configuration pupillaire vue depuis l'objet évolue avec la rotation terrestre. Ceci est utilisé en « super-synthèse », technique qui consiste, lorsque l'émission de la source est stationnaire, à répéter les mesures au cours de la nuit d'observation pour augmenter la couverture fréquentielle de l'interféromètre.

2.2.4 Techniques d'imagerie à travers la turbulence

Pour un instrument imageur tel qu'un un télescope monolithique de diamètre D, l'imagerie HRA depuis le sol a pour objectif de restaurer les HF au-delà de la fréquence de coupure r_0/λ de l'imagerie longue pose jusqu'à la fréquence de coupure de l'instrument D/r_0. Ceci est rendu possible par diverses techniques expérimentales qui évitent l'intégration temporelle des défauts de phase introduits par la turbulence. La qualité de la technique peut alors se mesurer par le rapport signal à bruit (RSB) résultant dans les hautes fréquences spatiales.

Techniques speckle

Les premières techniques haute résolution reposent sur l'acquisition d'une série d'images courtes poses et le calcul de moments empiriques. L'interférométrie[3] des tavelures [14], ou *speckle interferometry*, utilise la moyenne quadratique des transformées de Fourier des images, qui permet d'estimer l'autocorrélation de l'objet observé. Knox et Thomson [15], puis Weigelt [16] ont proposé des méthodes de traitement, utilisant respectivement l'interspectre (ou *cross-spectrum*) et le bispectre des images courtes poses, qui permettent d'estimer l'objet et pas seulement son autocorrélation. Ces méthodes nécessitent d'effectuer des moyennes sur un très grand nombre d'images, même pour des objets simples, à la fois pour que l'estimation des quantités statistiques estimées soit valide et pour améliorer le RSB.

Déconvolution par analyse de front d'onde

Une amélioration notable des résultats d'imagerie courte pose à travers la turbulence est alors survenue, non d'une amélioration des traitements, mais d'un changement de la technique expérimentale elle-même. Jean-Claude Fontanella a proposé en 1985 [17] une nouvelle technique d'imagerie : la déconvolution par analyse de front d'onde, aussi appelée déconvolution par analyse de surface d'onde (DECASO). Cette technique, fondée sur l'utilisation d'un dispositif appelé analyseur de front d'onde ou analyseur de surface d'onde (ASO), a été validée expérimentalement peu après [18].

[3] Le terme d'interférométrie pourrait laisser croire, à tort, que l'instrument utilisé ici est un interféromètre. Il n'en est rien et les interférences dont il s'agit proviennent d'une pupille monolithique.

Le but des ASOs, qui jusqu'alors n'étaient utilisés que pour vérifier la qualité de surface des miroirs de télescopes, est de mesurer les aberrations de systèmes optiques (la phase φ de l'Équ. (2.1)). Certains d'entre eux, comme l'analyseur de HS utilisé en déconvolution par analyse de front d'onde, fonctionnent même si l'on observe un objet d'intérêt étendu et non un point source. Si l'objet d'intérêt est dans l'espace et non dans le laboratoire, les aberrations mesurées sont alors celles du système instrument plus atmosphère. Aujourd'hui, les ASOs sont des composants essentiels des instruments d'imagerie à haute résolution. La section 2.3 donnera un aperçu rapide de ces dispositifs.

La technique de déconvolution par analyse de front d'onde consiste à enregistrer simultanément une série d'images courtes poses et des mesures de front d'onde par HS. En pratique, il faut utiliser typiquement au moins une dizaine d'images courtes poses pour couvrir correctement l'ensemble des fréquences spatiales jusqu'à la coupure du télescope (Équ. (2.6)). Ce nombre d'images doit être plus grand si l'objet observé est peu brillant.

La déconvolution par analyse de front d'onde représente une amélioration importante par rapport aux autres techniques courtes poses évoquées plus haut. En effet, tout d'abord, comme Knox-Thomson ou le bispectre, elle permet d'estimer non pas l'autocorrélation de l'objet, mais l'objet lui-même. Ensuite, contrairement aux précédentes techniques courtes poses, cette technique évite l'enregistrement d'images d'une étoile de référence ; c'est d'ailleurs pour quoi elle est appelée *self-referenced speckle interferometry*. Enfin, elle permet une mesure efficace en termes de photons collectés ; en effet, comme les images courtes poses doivent être quasi monochromatiques pour ne pas brouiller les tavelures, tous les photons restants peuvent être détournés sur l'ASO sans aucune perte de signal sur la voie image. Cette technique permet donc d'enregistrer plus d'information que les précédentes techniques courtes poses et contrairement à elles, possède un RSB qui n'est pas limité par le bruit de speckle à fort flux [19], du fait de son caractère autoréférencé. Ceci explique que les techniques d'interférométrie des tavelures soient aujourd'hui délaissées.

Optique adaptative

La technique d'imagerie la plus performante en termes de RSB est l'optique adaptative (OA), qui consiste en une compensation en temps réel des aberrations introduites par la turbulence atmosphérique, généralement par réflexion sur un miroir dont la surface est déformée à chaque instant *via* une boucle d'asservissement, en fonction des mesures d'un ASO.

L'ASO le plus couramment utilisé est l'analyseur de HS (*cf.* 2.3.2). Le miroir déformable associé comporte des actionneurs constitués par exemple d'empilements de plaques piézoélectriques, lesquels sont commandés par des hautes tensions. La technique de l'OA a été proposée par Babcock dès 1953, puis développée à partir des années 1970 pour les besoins de la défense, d'abord aux Etats-Unis puis en France, mais il a fallu attendre la fin des années 1980 pour que le premier système d'OA pour l'astronomie voie le jour [20]. Le lecteur intéressé par une présentation détaillée de l'OA pourra consulter un ouvrage de référence tel que [21].

Cette technique permet donc d'enregistrer des images longue pose (typiquement de quelques secondes à plusieurs dizaines de minutes) en conservant les HF de l'objet observé jusqu'à la fréquence de coupure propre du télescope. Ces HF sont néanmoins atténuées car la correction est partielle [22] et une déconvolution est nécessaire.

L'OA « classique » est aujourd'hui une technique mature ; de même la restauration d'images corrigées par OA et correspondant à un modèle d'imagerie convolutif est désormais bien maîtrisée. Les systèmes d'observation actuellement en développement sont plus complexes et les traitements auront vraisemblablement un grand rôle à jouer dans ceux-ci. Des exemples représentatifs sont les systèmes à OA grand champ dite multi-conjuguée [23], pour laquelle la réponse impulsionnelle ne pourra pas être considérée invariante spatialement et les systèmes comme SPHERE [24] ou GPI combinant une OA haute performance dite extrême et un coronographe en vue de détecter des exo-planètes. Pour de tels systèmes l'imagerie est foncièrement non convolutive et des traitements spécifiques doivent être développés. L'OA a également trouvé depuis quelques années une application à l'imagerie de la rétine et plusieurs équipes développent des systèmes opérationnels – voir par exemple [25, 26] et les références incluses dans ce dernier article. Dans ce contexte, l'image mesurée et l'objet à restaurer sont tri-dimensionnels.

2.2.5 Modèle d'image discret

Ce paragraphe vise à modéliser de manière correcte les images enregistrées au foyer d'un instrument, en distinguant les quantités continues comme les fonctions de transfert des quantités discrètes comme les images.

Les images sont enregistrées par un détecteur tel qu'une caméra CCD, qui intègre le flux incident sur une grille de pixels. Ceci peut être modélisé par une convolution du signal 2D incident par une réponse impulsionnelle (RI) h^{det} suivie d'un échantillonnage. La RI globale de l'instrument est donc la convolution de la RI du détecteur et de la RI optique, longue pose ou courte pose selon le mode d'acquisition :

$$h = h^{\text{det}} \bigstar h^{\text{opt}}. \tag{2.8}$$

Du fait du bruit inévitable lors de l'enregistrement des données (bruit de photons et bruits de détecteur en particulier), l'image enregistrée s'écrit :

$$\boldsymbol{i} = [h \bigstar o]_{\text{Ш}} + \boldsymbol{b} \tag{2.9}$$

où $[\,\cdot\,]_{\text{Ш}}$ désigne l'opération d'échantillonnage et $\boldsymbol{b}$ représente le bruit. Si celui-ci n'est pas additif et indépendant de l'image non bruitée, par exemple s'il est dominé par le bruit de photons, l'équation (2.9) devrait s'écrire $\boldsymbol{i} = [h \bigstar o]_{\text{Ш}} \diamond \boldsymbol{b}$, où le symbole $\diamond$ représente une opération pixel à pixel. Dans un souci de lisibilité je conserverai néanmoins la notation additive.

Ce modèle est généralement approximé par une convolution discrète entre la version échantillonnée $\boldsymbol{h}$ de h et la version échantillonnée $\boldsymbol{o}$ de l'objet continu observé o. Elle s'écrit alors sous la forme matricielle suivante :

$$\boldsymbol{i} = \boldsymbol{h} \star \boldsymbol{o} + \boldsymbol{b} = \mathbf{H}\,\boldsymbol{o} + \boldsymbol{b}, \tag{2.10}$$

où $\mathbf{H}$ est la matrice représentant la convolution discrète par $\boldsymbol{h}$, $\boldsymbol{i}$ est le vecteur obtenu en concaténant les colonnes de l'image correspondante, et $\boldsymbol{o}$ le vecteur obtenu en concaténant les colonnes de l'objet échantillonné.

2.3 Éléments d'analyse de front d'onde : plan pupille *vs* plan focal

2.3.1 Introduction

L'ASO est aujourd'hui un élément-clé d'un instrument d'imagerie à haute résolution, car il permet de mesurer les aberrations de celui-ci et de la turbulence atmosphérique afin de les compenser, soit en temps réel (OA) soit *a posteriori*, par traitement.

Il existe aujourd'hui un grand nombre d'ASOs, qui sont passés en revue de manière détaillée dans [27] et peuvent être classés en deux familles : les analyseurs plan focal et les analyseurs plan pupille.

Les systèmes d'OA actuellement opérationnels sur les grands télescopes utilisent généralement soit un analyseur de HS [28], bien décrit dans [17], soit un analyseur à courbure [29][4]. Tous deux appartiennent à la famille des ASOs plan pupille et utilisent une partie de la lumière incidente détournée au moyen d'une lame séparatrice (dichroïque). Pour l'OA ils ont tous deux les propriétés agréables qu'ils fonctionnent avec une large bande spectrale (parce qu'ils sont bien décrits par l'optique géométrique) et que la relation entre les aberrations inconnues et les données est linéaire, de sorte qu'elle peut être inversée en temps réel. La sous-section suivante présente le principe de fonctionnement de l'analyseur de HS, qui est utilisé dans la suite pour la technique de DECASO et est le plus utilisé en OA.

La famille des analyseurs de front d'onde en plan focal est née de l'idée naturelle que l'image d'un objet donné contient de l'information non seulement sur cet objet mais également sur le front d'onde. Un analyseur de front d'onde en plan focal requiert par conséquent peu ou pas d'autre optique que le capteur d'imagerie. Il est aussi le seul moyen pour être sensible à toutes les aberrations jusqu'au plan focal.

La sous-section 2.3.3 présente brièvement la technique d'analyse de front d'onde plan focal appelée diversité de phase [32]. Cette technique est à la fois très simple du point de vue matériel et, comme le Hartmann-Shack, fonctionne sur des objets très étendus.

Notons qu'il existe des ASOs particuliers appelés senseurs de cophasage qui permettent de mesurer les pistons différentiels entre pupilles, lesquels sont les aberrations spécifiques des interféromètres. La diversité de phase peut d'ailleurs à la fois être utilisée comme ASO et comme senseur de cophasage. À l'heure actuelle, ces pistons différentiels ne sont pas encore corrigés sur les interféromètres en fonctionnement.

[4] L'analyseur dit à pyramide [30] est depuis peu également installé dans un système d'OA d'un grand télescope [31].

L'essentiel de mes contributions à l'analyse de front d'onde concerne les analyseurs plan focal. Dans la suite de ce manuscrit, je ne détaillerai donc que mes travaux relatifs à cette dernière famille, même si j'ai effectué des contributions relatives à l'analyseur HS, en particulier en analyse de front d'onde pour l'OA multi-conjuguée [A37, A33] et en commande [A28].

2.3.2 L'analyseur de Hartmann-Shack

Le principe de cet analyseur est illustré Fig. 2.3 : une matrice de $N_{ml} \times N_{ml}$ microlentilles est placée dans un plan pupille (image de la pupille d'entrée du télescope) ; elle échantillonne, c'est-à-dire « découpe » le front d'onde incident. Au foyer de cette matrice, un ensemble de détecteurs (caméra CCD par exemple) enregistre les N_{ml}^2 imagettes, qui sont chacune l'image de l'objet observé à travers la portion de pupille découpée par la microlentille correspondante. Lorsque le front d'onde est perturbé par des aberrations, chaque microlentille voit approximativement un front d'onde plan incliné et l'imagette correspondante est donc décalée par rapport à sa position de référence d'une grandeur proportionnelle à la pente moyenne du front d'onde. Dans le cas d'aberrations dues à la turbulence atmosphérique, N_{ml} doit être choisi pour que la dimension de chaque microlentille, ramenée dans la pupille d'entrée de l'instrument, soit de l'ordre du diamètre de Fried r_0. On mesure la position du centre de gravité de chaque imagette, ce qui fournit une carte des pentes moyennes du front d'onde sur une grille[5] $N_{ml} \times N_{ml}$.

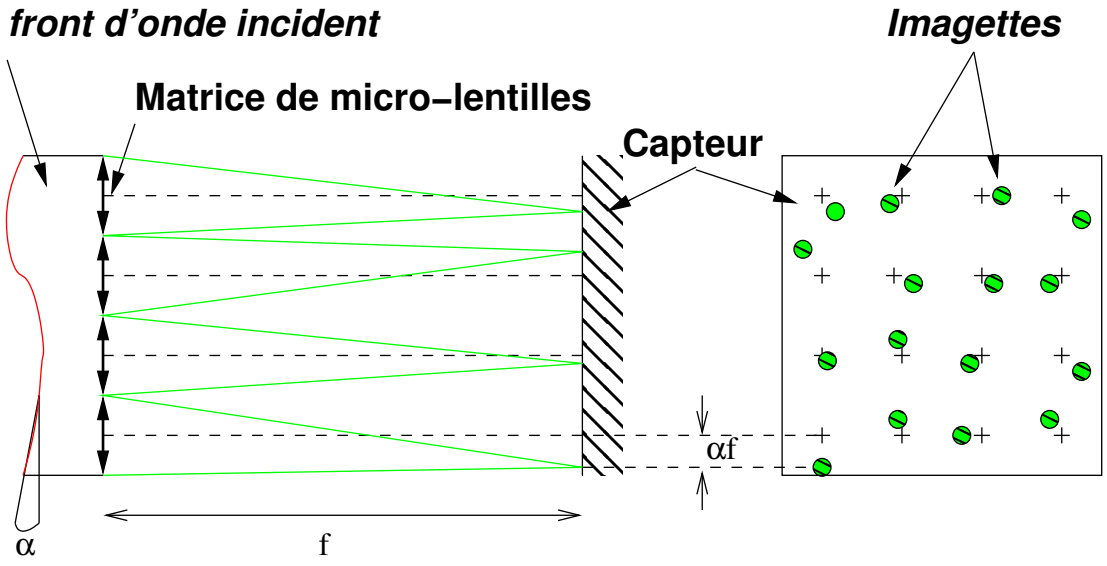


FIG. 2.3 – Principe de l'analyseur de HS

La phase inconnue à l'instant t, notée φ_t, est décomposée sur une base telle que celle des

[5]On peut envisager de considérer comme mesures non pas cette carte des pentes locales, mais directement l'ensemble des imagettes brutes ; en pratique, ces imagettes engendrent un flot de données important et ne sont donc généralement pas stockées sur disque. En effet, pour une utilisation en imagerie à travers la turbulence, il faut échantillonner le front d'onde à plusieurs dizaines, voire centaines de Hertz.

polynômes de Zernike [10] et on note ϕ_t^q les coefficients de cette décomposition :

$$\varphi_t(\boldsymbol{r}) = \sum_q \phi_t^q Z_q(\boldsymbol{r}) \tag{2.11}$$

où $\boldsymbol{r}$ est le point courant dans la pupille. Le problème direct peut donc se mettre sous la forme linéaire suivante :

$$\boldsymbol{s}_t = \mathbf{D}\boldsymbol{\phi}_t + \boldsymbol{b}'_t$$

où $\boldsymbol{s}_t$ est le vecteur concaténant les $2N_{ml}^2$ mesures de pente (x et y), $\boldsymbol{\phi}_t$ le vecteur des coordonnées de la phase inconnue et $\mathbf{D}$ essentiellement un opérateur de dérivation échantillonné appelé « matrice ASO ».

Le bruit est généralement supposé blanc gaussien et homogène. L'indépendance entre les mesures des différentes sous-pupilles est naturelle et le caractère gaussien justifie parce que résultant d'une estimation de centre de gravité sur un « grand » nombre de pixels (typiquement quelques dizaines)

La solution classiquement utilisée pour estimer la phase, en particulier sous des contraintes de temps réel (OA), est l'estimation aux moindres carrés (MC). Les problèmes rencontrés dans l'approche MC et les solutions pratiques généralement adoptées sont discutées dans [L4].

Comme l'on possède de bonnes connaissances statistiques sur la turbulence (*cf.* références du 2.2.3), une approche bayésienne est plus appropriée et donne de meilleurs résultats. Néanmoins, l'estimation aux moindres carrés est encore aujourd'hui très largement utilisée en OA.

Le problème de reconstruction de la phase étant linéaire et gaussien, il conduit à un estimateur MMSE/MAP analytique, sous forme dite *covariance* dans [33] et sous forme dite *information* dans [34, 35]. L'estimation MAP de chacune des phases correspond à minimiser le critère mixte $J_{\text{MAP}}^{\phi} = J_s + J_\phi$, avec :

$$J_s = \frac{1}{2}(\boldsymbol{s}_t - \mathbf{D}\boldsymbol{\phi}_t)^T \mathbf{C}_{b'}^{-1}(\boldsymbol{s}_t - \mathbf{D}\boldsymbol{\phi}_t) \tag{2.12}$$

$$\text{et} \quad J_\phi = \frac{1}{2}\boldsymbol{\phi}_t^T \mathbf{C}_\phi^{-1}\boldsymbol{\phi}_t \tag{2.13}$$

où $\mathbf{C}_{b'}$ est la matrice de covariance du bruit de mesure des pentes (diagonale, de diagonale à peu près constante) et $\mathbf{C}_\phi$ la matrice de covariance de la phase turbulente dans la base de Zernike, qui est déduite de l'Equ. (2.3) et fonction uniquement de r_0. La solution bien connue est :

$$\widehat{\boldsymbol{\phi}}_t = (\mathbf{D}^T\mathbf{C}_{b'}^{-1}\mathbf{D} + \mathbf{C}_\phi^{-1})^{-1}\mathbf{D}^T\mathbf{C}_{b'}^{-1}\boldsymbol{s}_t. \tag{2.14}$$

Cette solution tire parti des connaissances sur la statistique spatiale de la turbulence. Pour une utilisation en OA, où la fréquence d'échantillonnage est généralement bien supérieure à $1/\tau_0$, il est judicieux d'opter pour une extension naturelle de cet estimateur MMSE qui utilise également les connaissances *a priori* sur la statistique temporelle de la turbulence ; cette extension est l'estimateur optimal du filtrage de Kalman [A28], [36].

2.3.3 *Phase retrieval* et diversité de phase

Le « phase retrieval » consiste à estimer les aberrations vues par un instrument à partir de l'image d'un point-source. Ceci revient à inverser l'équation (2.1), c'est-à-dire à en estimer la phase φ à partir d'une mesure de h. Cette technique, née en microscopie électronique [37] puis redécouverte en optique [38], a deux limitations majeures : d'une part elle ne fonctionne qu'avec un objet ponctuel, d'autre part la solution obtenue souffre d'une ambiguïté de signe et n'est généralement pas unique (voir par exemple [L6, sect. 1.C]).

Gonsalves [32] a montré qu'en utilisant une deuxième image comportant une variation connue des aberrations par rapport à la première, par exemple une défocalisation, il était possible d'estimer les aberrations même si l'objet était étendu et inconnu. De plus cette deuxième image lève l'indétermination mentionnée plus haut et les aberrations estimées sont uniques, en pratique, pour de faibles aberrations. Cette technique est appelée diversité de phase par analogie avec une technique utilisée en télécommunications.

Soient ϕ la phase inconnue et ϕ_d cette phase dite de diversité, les images focale $\boldsymbol{i}_f$ et défocalisée $\boldsymbol{i}_d$ enregistrées s'écrivent, d'après le modèle des équations (2.8) à (2.10) :

$$\boldsymbol{i}_f = \boldsymbol{h}^{\text{det}} \star \boldsymbol{h}^{\text{opt}}(\phi) \star \boldsymbol{o} + \boldsymbol{b_f} \tag{2.15}$$

$$\boldsymbol{i}_d = \boldsymbol{h}^{\text{det}} \star \boldsymbol{h}^{\text{opt}}(\phi + \phi_d) \star \boldsymbol{o} + \boldsymbol{b_d}. \tag{2.16}$$

La diversité de phase est utilisée dans deux contextes différents : on peut vouloir imager un objet à distance, par exemple en astronomie solaire, ou l'on peut vouloir mesurer les aberrations vues par un instrument, pour les corriger en temps réel ou *a posteriori*. Ces deux problématiques sont liées mais néanmoins distinctes. Dans les deux cas, la base de l'inversion consiste à estimer les aberrations et l'objet qui sont les plus compatibles avec les images mesurées. L'approche historique, qui est également la plus simple, est une estimation conjointe de l'objet et de la phase [32] fondée sur la minimisation, en l'objet et les aberrations inconnus, d'un critère MC.

Le lecteur intéressé pourra trouver dans [L6, Annexe B page 133] un historique plus complet et une revue des applications de la diversité de phase ainsi qu'une étude détaillée des propriétés des différentes méthodes d'estimation associées à cet ASO.

Chapitre 3

Travaux en étalonnage d'instrument : analyse de front d'onde au plan focal et estimation de réponse instrumentale

3.1 Travaux méthodologiques en diversité de phase

Comme évoqué en introduction à la diversité de phase, section 2.3.3, le traitement usuel des données en diversité de phase consiste en une estimation conjointe des aberrations et de l'objet observé [39]. L'interprétation bayésienne de cette approche est qu'elle consiste à calculer l'estimateur MAP *conjoint* (MAPJ) suivant :

$$\begin{aligned}(\hat{\boldsymbol{o}}, \hat{\boldsymbol{\phi}})_{\text{MAP}} &= \arg\max_{\boldsymbol{o},\boldsymbol{\phi}} p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}) \\ &= \arg\max_{\boldsymbol{o},\boldsymbol{\phi}} p(\boldsymbol{i}_1|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}_b)\, p(\boldsymbol{i}_2|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}_b)\, p(\boldsymbol{o}; \boldsymbol{\theta}_o)\, p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi). \end{aligned} \tag{3.1}$$

où $p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta})$ est la densité de probabilité conjointe des données (images $\boldsymbol{i}_1, \boldsymbol{i}_2$), de l'objet $\boldsymbol{o}$ et des aberrations $\boldsymbol{\phi}$. Cette densité dépend généralement d'un ensemble d'hyper-paramètres liés au bruit, à l'objet recherché et aux aberrations, que l'on regroupe dans un vecteur $\boldsymbol{\theta} = (\boldsymbol{\theta}_b, \boldsymbol{\theta}_o, \boldsymbol{\theta}_\phi)$. La vraisemblance des données $\boldsymbol{i}_k$ est notée $p(\boldsymbol{i}_k|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta})$; $p(\boldsymbol{o}; \boldsymbol{\theta}_o)$ et $p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi)$ sont les densités de probabilité *a priori* de $\boldsymbol{o}$ et $\boldsymbol{\phi}$. Sous les hypothèses gaussiennes souvent adoptées pour le bruit et la loi *a priori* pour l'objet, l'objet optimal s'écrit analytiquement à phase donnée, et le critère MAPJ peut donc se réécrire en y injectant cet objet optimal. Ceci donne un critère que nous appellerons *critère conjoint réinjecté*, qui ne dépend explicitement que de la phase, ce qui rend l'optimisation plus efficace sans changer la solution. Nous appellerons *critère conjoint réinjecté* ce critère qui ne dépend explicitement que de la phase.

La très grande majorité des structures d'estimation utilisées en diversité de phase peut se réécrire sous la forme de l'équation (3.1), même si celles-ci n'ont pas été introduites dans un cadre bayésien à leur origine – voir [L6, sect. 3] pour un historique complet. Cette structure est pourtant connue pour ne *pas* avoir les bonnes propriétés asymptotiques (consistance, efficacité

asymptotique) du MAP ou du MV du fait de son caractère conjoint : si l'on augmente le nombre de données, par exemple en augmentant la taille des images, le nombre d'inconnues liées à l'objet augmente dans les mêmes proportions, et donc le contraste statistique – que l'on peut définir simplement comme le rapport du nombre de mesures au nombre d'inconnues – ne s'améliore pas.

Lors de la thèse d'Amandine Blanc [T10], nous avons proposé un nouvel estimateur, appelé estimateur marginal, qui reconstruit uniquement les aberrations inconnues $\boldsymbol{\phi}$ (et les hyper-paramètres $(\boldsymbol{\theta}_b, \boldsymbol{\theta}_o)$) en intégrant l'objet hors du problème[1]. C'est le « vrai » estimateur MAP des aberrations, qui s'obtient en intégrant la densité de probabilité conjointe des données :

$$\begin{aligned}\hat{\boldsymbol{\phi}}_{\text{MAP}} &= \arg\max_{\boldsymbol{\phi}} p(\boldsymbol{i}_f, \boldsymbol{i}_d, \boldsymbol{\phi}; \boldsymbol{\theta}) = \arg\max_{\boldsymbol{\phi}} \int p(\boldsymbol{i}_f, \boldsymbol{i}_d, \boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{o} \\ &= \arg\max_{\boldsymbol{\phi}} \int p(\boldsymbol{i}_f|\boldsymbol{\phi}, \boldsymbol{o}; \boldsymbol{\theta}) \, p(\boldsymbol{i}_d|\boldsymbol{\phi}, \boldsymbol{o}; \boldsymbol{\theta}) \, p(\boldsymbol{\phi}; \boldsymbol{\theta}) \, p(\boldsymbol{o}; \boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{o}.\end{aligned} \tag{3.2}$$

Cet estimateur, proposé dans [A29], est présenté dans [L6, Annexe B page 133]. Sous les hypothèses gaussiennes adoptées pour le bruit et la loi *a priori* pour l'objet, le critère MAP s'écrit analytiquement, et une relation simple lie le critère MAP au critère joint classique.

Les simulations détaillées de [L6, sect. 4] illustrent les bonnes propriétés que cet estimateur possède :

- le critère MAP converge effectivement, en pratique, vers la vraie valeur des aberrations lorsqu'on augmente la taille des images, en accord avec la propriété de consistance de l'estimateur MAP ;
- les minima locaux parfois présents dans l'estimateur joint régularisé sont absents avec l'estimateur marginal pour des images assez grandes, ce qui est en accord avec le caractère asymptotiquement convexe du critère MAP.

L'estimateur marginal nécessite un bon réglage des hyper-paramètres relatifs à l'objet et au bruit – les DSPs de ces deux quantités. Fort heureusement, il est possible d'estimer ces hyper-paramètres en même temps que les aberrations, avec un impact négligeable sur la qualité d'estimation de celles-ci dès que la taille des images devient suffisante – en pratique 128×128 dans les simulations effectuées.

Nous avons également étudié par simulations l'estimateur conjoint, très utilisé dans la littérature bien que l'estimation conjointe ait en général de mauvaises propriétés statistiques. Nous avons pu constater, sur de nombreuses simulations, qu'il était possible d'obtenir une bonne estimée des aberrations lorsque l'on sous-régularisait l'objet, pour des images présentant un assez bon RSB. L'objet obtenu est alors inexploitable, mais les aberrations estimées sont bien meilleures qu'avec la régularisation objet déduite de la DSP vraie de celui-ci. Nous avons pu, de plus, comprendre théoriquement ce comportement : nous avons montré que l'estimateur conjoint

[1] En probabilités, intégrer hors du problème c'est-à-dire marginaliser une quantité signifie calculer une loi marginale, par une intégrale sur toutes les valeurs possibles de la quantité en question.

réinjecté est un estimateur des aberrations consistant[2] si la régularisation objet tend vers zéro quand la taille des images devient arbitrairement grande, *cf.* [A21, Annexe J page 289]. Dans cet article on montre également que le fait de pré-filtrer les images, comme le font certains auteurs pour enlever le bruit HF, ne fait pas perdre la consistance de l'estimateur. Ces études permettent d'éclairer le fait que l'estimation des aberrations sans régularisation objet, avec ou sans filtrage des images, ait pu être utilisée avec succès dans la littérature. Elles donnent également des indications pour une meilleure utilisation de l'estimateur conjoint. Si la quantité d'intérêt ultime est l'objet observé, une stratégie efficace est d'effectuer, après l'estimation des aberrations et des hyper-paramètres, une restauration d'images utilisant ceux-ci. Cette restauration d'images est alors entièrement non-supervisée, comme pour l'estimateur marginal.

Deux limitations souvent rencontrées en diversité de phase, qui ont empêché ce capteur d'être utilisé comme ASO temps-réel en imagerie à travers la turbulence, sont d'une part le fait que cette technique fonctionne souvent mal pour des aberrations de grande amplitude, typiquement supérieures à 2π, et d'autre part le grand temps de calcul requis du fait du caractère très non-linéaire du modèle de données.

La première limitation est liée au fait que les données ne contiennent de l'information sur la phase que modulo 2π. Nous avons montré que l'utilisation, classique, d'une base de Zernike tronquée pour décrire la phase inconnue rendait le critère MAPJ et le critère MAP multi-modaux c'est-à-dire introduisait des minima locaux. Nous avons suggéré d'utiliser une base de pixels pour décrire la phase, assortie d'une régularisation originale qui impose une régularité à la phase modulo 2π. Ce choix de base et cette régularisation permettent d'éviter l'introduction de minima locaux dans le critère. Cette étude est décrite dans [L6, sect. 8].

Nos travaux visant à une réduction significative du temps de calcul sont décrits dans la section suivante.

3.2 Travaux algorithmiques en diversité de phase

Durant les quinze dernières années, un certain nombre d'auteurs ont apporté des améliorations à la rapidité d'estimation des aberrations en diversité de phase, améliorations fondées pour certaines sur de meilleurs algorithmes d'optimisation [40, 41], pour d'autres sur des modifications du critère à optimiser [42, 43, 44]. Néanmoins tous ces algorithmes nécessitent plus d'une itération à notre connaissance.

Lors de la thèse d'Isabelle Mocœur [T3], j'ai proposé une approche originale fondée sur une approximation quadratique du critère à optimiser, laquelle permet d'obtenir une solution de manière non-itérative dans l'approximation des faibles phases. Cette approche a fait l'objet d'un brevet [B1] et de deux articles [A8],[A9, Annexe D page 229]. Les paragraphes ci-dessous en décrivent les grandes lignes. Sous les hypothèses gaussiennes pour le bruit et la loi *a priori* pour

[2] Un estimateur est consistant s'il tend vers la vraie valeur des paramètres inconnus lorsque la taille des données tend vers l'infini.

l'objet, le critère conjoint réinjecté s'écrit, dans le domaine de Fourier discret et à régularisation quasi-nulle :

$$J'_{\text{MAPJ}}(\boldsymbol{\phi}) = \frac{1}{2}\sum_{\boldsymbol{u}} \frac{\left|\tilde{\boldsymbol{\imath}}_f(\boldsymbol{u})\tilde{\boldsymbol{h}}_d(\boldsymbol{\phi},\boldsymbol{u}) - \tilde{\boldsymbol{\imath}}_d(\boldsymbol{u})\tilde{\boldsymbol{h}}_f(\boldsymbol{\phi},\boldsymbol{u})\right|^2}{|\tilde{\boldsymbol{h}}_f(\boldsymbol{\phi},\boldsymbol{u})|^2 + |\tilde{\boldsymbol{h}}_d(\boldsymbol{\phi},\boldsymbol{u})|^2 + \epsilon}, \tag{3.3}$$

où $\tilde{\boldsymbol{h}}_f$ et $\tilde{\boldsymbol{h}}_d$ sont les fonctions de transfert des images focalisée et défocalisée respectivement, qui dépendent des aberrations inconnues $\boldsymbol{\phi}$, et ϵ correspond à une régularisation minimale évitant les explosions numériques. On remarque aisément que dans cette équation, le numérateur devient nul, au bruit près, lorsque les aberrations $\boldsymbol{\phi}$ sont égales à celles qui ont donné naissance aux images, et croît dès qu'elles s'en éloignent. Le dénominateur est moins sensible à la valeur des aberrations ; on peut donc envisager de le considérer comme un simple facteur de pondération entre fréquences et de remplacer J'_{MAPJ} par :

$$J"_{\text{MAPJ}}(\boldsymbol{\phi}) = \frac{1}{2}\sum_{\boldsymbol{u}} \frac{\left|\tilde{\boldsymbol{\imath}}_f(\boldsymbol{u})\tilde{\boldsymbol{h}}_d(\boldsymbol{\phi},\boldsymbol{u}) - \tilde{\boldsymbol{\imath}}_d(\boldsymbol{u})\tilde{\boldsymbol{h}}_f(\boldsymbol{\phi},\boldsymbol{u})\right|^2}{|\tilde{\boldsymbol{h}}_f(\boldsymbol{\phi}_0,\boldsymbol{u})|^2 + |\tilde{\boldsymbol{h}}_d(\boldsymbol{\phi}_0,\boldsymbol{u})|^2 + \epsilon} \tag{3.4}$$

où $\boldsymbol{\phi}_0$ est une phase fixée, par exemple la phase nulle, ou la dernière phase connues dans le cas d'une utilisation en boucle fermée. Le critère (3.4) est quadratique en $\tilde{\boldsymbol{h}}_f$ et $\tilde{\boldsymbol{h}}_d$. Dans l'approximation des faibles phases, il est classique de développer ces fonctions de transfert au premier ordre en fonction des aberrations, par exemple autour de la phase ϕ_0. Le critère (3.4) est alors quadratique vis-à-vis des aberrations, et la solution minimisant ce critère est alors analytique. Son coût est essentiellement la FFT des deux images, ce qui permet d'envisager une utilisation temps-réel à plusieurs dizaines de Hertz. Le lecteur intéressé trouvera plus de détails ainsi qu'une évaluation des performances de cette approche sur simulations dans [A8] et dans [A9, Annexe D page 229].

3.3 Mesure des aberrations statiques pour l'imagerie par optique adaptative

La mesure des aberrations instrumentales est un problème récurrent en Haute Résolution Angulaire et incontournable pour permettre leur correction, que celle-ci soit faite *a posteriori* numériquement ou en temps-réel par une correction physique – OA ou optique active. Dans le cas d'un instrument corrigé par OA, les aberrations différentielles entre la voie d'imagerie scientifique et la voie ASO sont à la fois délicates à mesurer car non vues par l'ASO et souvent dominantes en imagerie à Haute Dynamique. Ces aberrations différentielles, connues sous le nom de NCPA pour *Non-Common Path Aberrations*, ne peuvent être mesurées que par un analyseur proche du plan focal de l'instrument. La diversité de phase, qui peut utiliser le capteur d'imagerie scientifique lui-même, est donc un analyseur idéal pour cette tâche.

3.3.1 Mesure et correction des aberrations statiques sur NAOS-CONICA

Avec quelques collègues, nous avons utilisé la diversité de phase sur le système NAOS-CONICA du VLT dès son intégration afin de mesurer et de corriger les aberrations statiques de l'OA NAOS et de sa caméra CONICA. La méthode a été utilisée de manière opérationnelle au Chili, sur le VLT. Dans l'article [A31] nous avons tâché de lister toutes les limitations potentielles de cette technique et d'évaluer leur impact au moyen de simulations : méconnaissance partielle du système optique (distance exacte de défocalisation, facteur de sur-échantillonnage, forme de la pupille, etc.), qualité des images (RSB, structures résiduelles du fond, etc.), limitations de l'algorithme (largeur spectrale non modélisée, nécessité d'un recalage des images, etc). Dans l'article [A30] nous avons utilisé la diversité de phase pour mesurer et corriger les NCPA de NAOS-CONICA. Le très grand nombre de modes de l'instrument NAOS-CONICA correspond à un très grand nombre de combinaisons possibles pour les trois sources d'aberrations que sont : la lame dichroïque de NAOS, le filtre spectral de la caméra CONICA, et l'objectif de la caméra CONICA. Il était déraisonnable d'envisager d'étalonner individuellement chacun de ces modes. L'article [A30] détaille comment nous avons étalonné séparément l'ensemble NAOS-CONICA d'une part et la caméra d'autre part pour finalement démêler les trois sources d'aberrations, réduire notablement la combinatoire de l'étalonnage et aboutir à une méthode opérationnelle.

3.3.2 Diversité de phase en ligne : mesure des aberrations statiques seules sur le ciel

Une limitation de la technique de diversité de phase, en particulier pour la mesure des NCPA, est qu'elle utilisait jusqu'à ce jour des images suffisamment « courtes poses » pour que les aberrations du système restent constantes pendant la mesure. Ainsi par exemple, dans la méthode décrite ci-dessus, l'étalonnage est effectué de jour en utilisant des sources internes situées soit en entrée du système NAOS-CONICA soit en entrée de la caméra CONICA. Pour des applications très exigeantes en terme de qualité de front d'onde comme la détection d'exo-planètes avec SPHERE ou le futur système EPICS, il peut être utile voire indispensable de réaliser l'étalonnage plusieurs fois durant la nuit, sans interrompre l'observation scientifique.

Dans ce but, j'ai proposé récemment une extension de la technique de diversité de phase qui utilise des images *longues poses* corrigées par OA pour mesurer les aberrations quasi-statiques pendant l'observation scientifique. Le principe de la méthode est que, pour un temps de pose suffisamment long, l'effet de la turbulence résiduelle se moyenne en une composante convolutive de l'image et que la diversité de phase estime alors les seules aberrations statiques d'intérêt. Les avantages d'une telle procédure, comparée au traitement d'un ensemble de paires d'images courtes poses, sont nombreux : la séparation entre aberrations statiques et aberrations turbulentes résiduelles est effectuée par la longue pose elle-même et non numériquement ; seule une paire d'images doit être traitée ; l'estimation des aberrations bénéficie du fort RSB des images longues poses ; et enfin seules les aberrations statiques d'intérêt doivent être estimées. Les détails de la

méthode sont exposés dans l'article [A15, Annexe G page 257]. Elle devrait bientôt être validée expérimentalement sur le banc BOA de l'ONERA.

3.4 Analyse de front d'onde « sans analyseur »

Comme évoqué dans la section 2.3.3 et détaillé dans [L6, sect. 1.C], retrouver les aberrations d'un système optique à partir d'une seule image focale est très difficile. Cette opération, appelée *phase retrieval*, d'une part ne fonctionne qu'avec un objet ponctuel ou au moins connu, d'autre part donne une solution qui souffre d'une ambiguïté de signe et n'est *généralement* pas unique. Néanmoins, l'analyse de front d'onde à partir d'une seule image focale est conceptuellement très séduisante et pourrait avoir de nombreuses applications pratiques du fait de sa grande simplicité. Je mentionne donc dans les paragraphes qui suivent deux cas d'intérêt pratique avéré dans lesquels il a été possible de développer un étalonnage instrumental avec un seul plan image pour un télescope monolithique. Un troisième cas d'intérêt potentiel du « phase retrieval » concerne spécifiquement le cophasage d'un interféromètre et sera évoqué au paragraphe 3.5.2.

3.4.1 Étalonnage sur point source (SICLOPE)

Même lorsque l'objet observé est connu voire ponctuel, l'utilisation d'une image différente de l'image focalisée est indispensable pour lever l'ambiguïté de signe sur la partie paire de la phase recherchée – voir [L6, sect. 1.C] par exemple. L'ajout, par rapport à une image qui serait focalisée, d'une défocalisation ou de tout autre mode pair, supérieur en valeur absolue au mode correspondant dans l'aberration à estimer, suffit à lever cette ambiguïté de signe. Ceci peut être vu comme une forme de *porteuse* sur laquelle on met la phase inconnue, et permet alors de se passer de l'image focalisée. De plus, l'utilisation d'une porteuse autre que la défocalisation, par exemple un astigmatisme, permet de conserver une image plus contrastée, donc un meilleur RSB. Ce procédé, baptisé SICLOPE pour *Single Image with Calibrated Local Offset for Phase Estimation*, peut s'appliquer indifféremment à un télescope monolithique ou à un interféromètre et a fait l'objet d'un brevet [B2].

Nous avons développé, pour un capteur de type SICLOPE, un estimateur analytique fondé sur une estimation MV dans l'approximation des faibles phases[3] [A4]. Cet estimateur, inspiré de nos travaux similaires en diversité de phase (*cf.* [A9] et [B1]), est actuellement étudié dans l'équipe pour servir d'analyseur « bas ordres » (tip-tilt, défocalisation, etc.) pour le système d'OA ATLAS de l'ELT européen.

3.4.2 Étalonnage sur objet structuré : estimation de réponse instrumentale

Dans certaines applications, l'on cherche à étalonner un instrument imageur pour restaurer les images acquises mais l'on ne souhaite pas toujours le corriger de ses aberrations. C'est par

[3] Approximation raisonnable pour une utilisation en boucle fermée.

exemple le cas de télescopes d'observation de la Terre depuis satellite, pour lesquels une correction active peut être superflue s'ils sont assez stables et/ou trop complexe pour être envisagée. Dans ce type d'application, la quantité d'intérêt à estimer est la RI du système et non les aberrations, donc l'ambiguïté de signe sur la partie paire de la phase est sans importance. La RI peut être estimée à partir de l'image de motifs connus tels que des *bords de plage* (fonctions de Heavyside) dans plusieurs directions : limites entre champs pour des images basse résolution, bords de bâtiments, etc. J'ai, avec mon collègue Guy Le Besnerais, développé une méthode d'estimation de RI qui se fonde sur la même paramétrisation de la réponse que le *phase retrieval via* la phase pupillaire (équ. (2.1) en identifiant des bords de plages naturels ou artificiels [C72, C68]. Cette méthode peut également être utilisée pour estimer la RI d'un système d'imagerie en laboratoire comme le système INCA (Identification Nuit-jour de Cibles Aériennes, étude menée pour la DGA) [45].

3.5 Cophasage par diversité de phase

3.5.1 Introduction

Pour qu'un interféromètre, qu'il soit imageur ou non (*cf.* sect. 2.2.2), atteigne ses performances ultimes, les éléments de sa pupille, appelés sous-pupilles ici, doivent être cophasés avec une précision d'une petite fraction de la longueur d'onde. Un sous-système critique d'un interféromètre est donc le senseur de cophasage (SC), dont le but est de mesurer le positionnement relatif des sous-pupilles, en l'occurrence les pistons et tip-tilts différentiels entre sous-pupilles, qui sont les sources principales de dégradation du front d'onde.

La mesure de pistons et tip-tilts différentiels pour les grands interféromètres au sol a été étudiée en détail et démontrée expérimentalement. La plupart des dispositifs proposés sont fondés sur une recombinaison en plan pupille de la lumière provenant d'une paire donnée de sous-pupilles. Du fait de la recombinaison en plan pupille, le contraste des franges d'interférence décroît rapidement lorsque l'extension de l'objet observé augmente, ce qui rend de tels dispositifs inutilisables sur des scènes très étendues telles que la Terre vue de l'Espace. Par ailleurs, du fait de la recombinaison par paires de sous-pupilles, ces dispositifs engendrent une complexité qui croît déraisonnablement avec le nombre de sous-pupilles composant l'interféromètre.

Le fait que la technique de diversité de phase puisse être utilisée comme SC a été identifié très tôt [46]. De plus, contrairement aux dispositifs en plan pupille évoqués plus haut, la diversité de phase a deux propriétés remarquables : d'une part, elle est très appropriée pour un instrument possédant un grand nombre de sous-pupilles, parce que la complexité du senseur ne croît pas plus vite que le nombre de sous-pupilles et est quasiment indépendante de celui-ci ; ceci pourrait représenter un avantage décisif pour le SC des instruments de seconde génération du VLTI [C49] ou pour la mission DARWIN [C42]. D'autre part, elle peut être utilisée sur des objets très étendus comme la Terre observée depuis l'Espace [C51].

La première de ces propriétés et l'absence d'aberrations différentielles par rapport au plan focal d'imagerie (lorsqu'il y en a un) ou NCPA, *cf.* § 3.3, sont deux motivations fortes pour choisir la diversité de phase comme SC, même lorsque l'objet observé est non résolu.

La diversité de phase a fait l'objet de validations expérimentales au sol en tant que SC sur point source, en particulier pour le cophasage des segments du télescope Keck [47]. En ce qui concerne les instruments spatiaux, la diversité de phase a été retenue pour le cophasage fin du télescope segmenté JWST [48] à la suite de nombreuses études, en particulier [49, 50, 51]. Cette technique a également été choisie par l'ESA comme SC pour l'interféromètre DARWIN [C61].

L'utilisation de la diversité de phase comme SC sur source étendue a été validée expérimentalement [52, 42] afin de corriger les aberrations quasi-statiques en temps réel sur miroir segmenté. Plus récemment, la diversité de phase a permis de corriger en boucle fermée les aberrations quasi-statiques d'un interféromètre imageur à large bande spectrale composé de six télescopes [53, 54].

3.5.2 Cophasage d'interféromètre

sur objet étendu

L'ONERA a mené une étude de faisabilité de l'observation haute résolution permanente de la Terre depuis une orbite géostationnaire [C57, C51, C45], étude financée par la DGA et dont j'ai été le chef de projet. Dans le cadre de cette étude, surnommée SOTISE (Satellite d'Observation de la Terre par Interférométrie sur Scènes Etendues) nous avons identifié que la diversité de phase était la meilleure solution pour le SC puis nous avons conçu, réalisé et testé un prototype de capteur de cophasage avec son banc de test appelé BRISE (Banc Reconfigurable d'Interférométrie sur Sources Etendues).

Ce banc, illustré figure 3.1 et décrit en détail dans [C56, C43, C39], se compose d'un collimateur de diamètre 160 mm, en autocollimation sur un miroir segmenté à trois sous-pupilles commandables en piston et tip-tilt dans un diamètre circonscrit de 60 mm. Dans sa conception, nous avons limité au maximum les perturbations environnementales : turbulence, vibrations, dilatations thermiques. Ce banc inclut un capteur dit de référence, fondé également sur la diversité de phase mais fonctionnant sur point-source et à fort flux, implanté sur un trajet optique très proche du SC et enregistrant ses données sur le même capteur CCD. Le capteur de référence permet d'étalonner les mesures du SC. La figure 3.2 montre une image expérimentale produite par ces deux capteurs. Une répétabilité nanométrique des mesures de piston a été atteinte avec le capteur de co-phasage pour des scènes étendues représentatives de scènes terrestres en fort flux, *cf.* Fig. 3.3 (d'après [C45]).

Cette performance sur objets étendus peut aujourd'hui être atteinte en temps réel grâce aux algorithmes analytiques développés récemment et validés sur images expérimentales, *cf.* section 3.2 et [A9, Annexe D page 229].

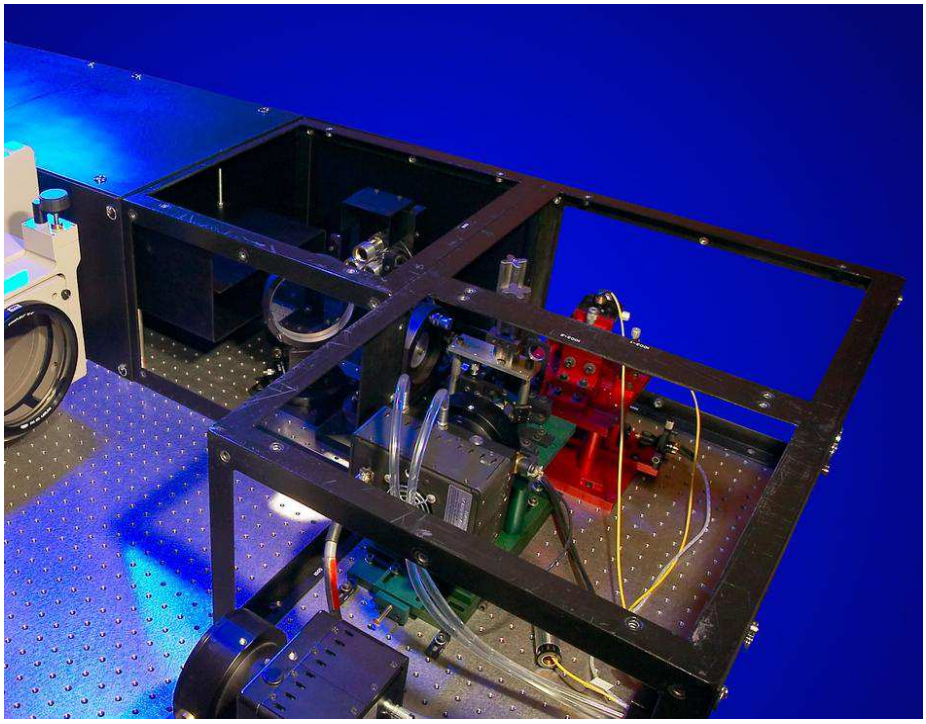


FIG. 3.1 – Photographie du banc BRISE. Le SC est en vert, le module contenant les sources en rouge. Le miroir segmenté à trois sous-pupilles est visible au fond.

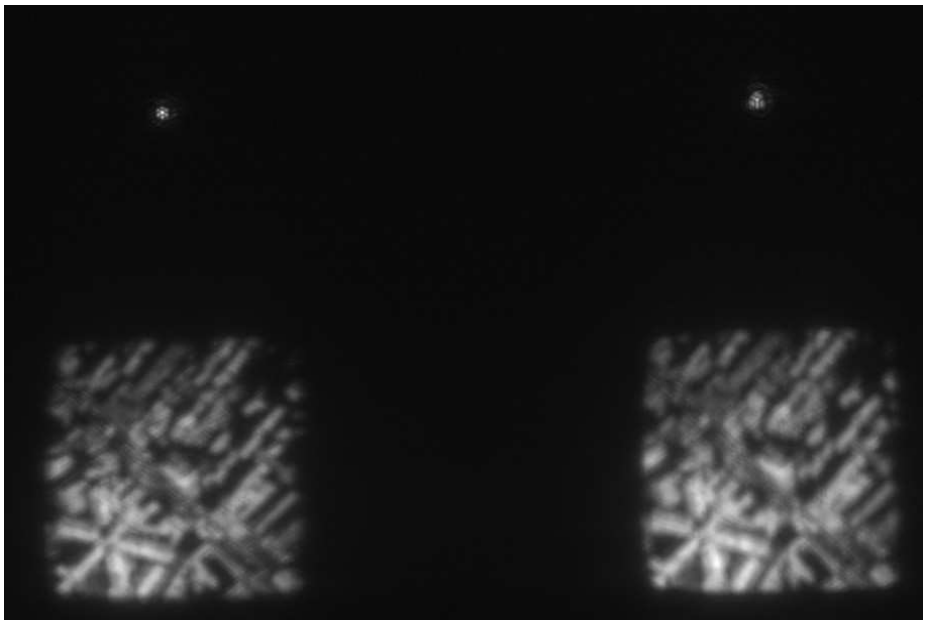


FIG. 3.2 – Images expérimentales du banc de test de cophasage BRISE, acquises simultanément. En haut, images d'un point source servant de référence (à g.: focalisé, à d. : défocalisé). En bas, images d'un objet étendu (scène urbaine, à g.: focalisée, à d. :défocalisée).

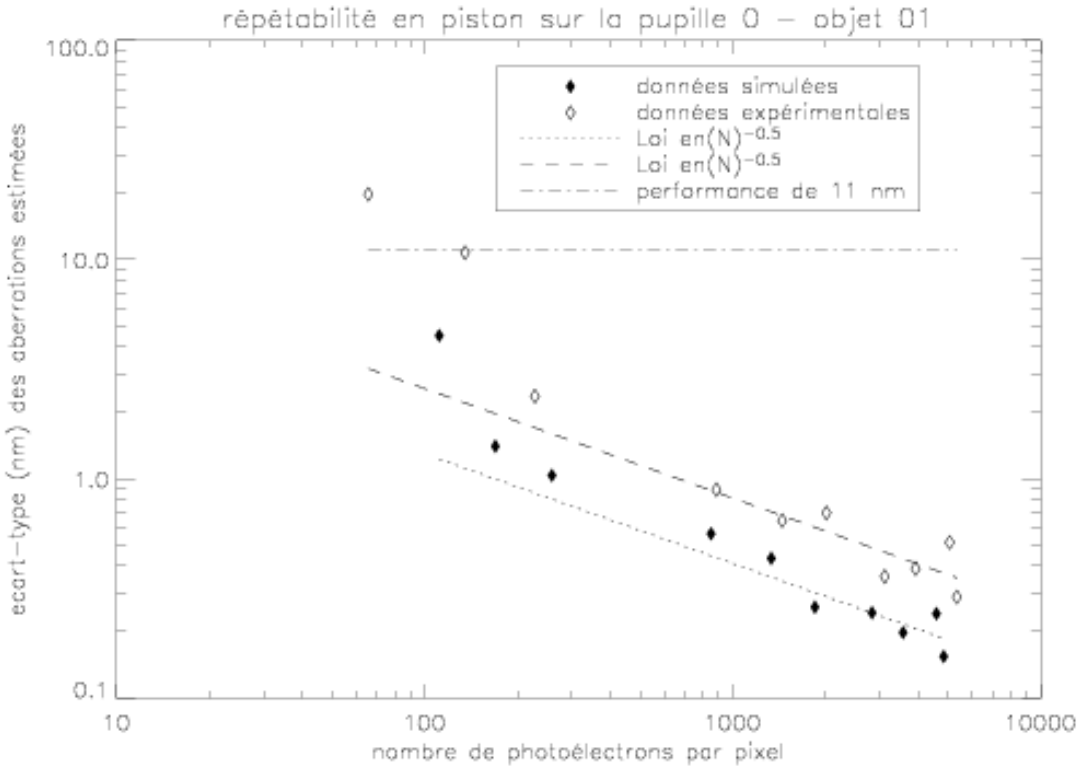


FIG. 3.3 – Répétabilité expérimentale de la mesure de piston par le capteur de co-phasage sur objet étendu et comparaison avec la simulation.

sur objet ponctuel

Pour un interféromètre observant un objet ponctuel comme une étoile non résolue, dès que l'agencement relatif des télescopes est approprié (configuration pupillaire dite « non redondante »), il est possible de mesurer pistons et tip-tilts différentiels avec une seule image plan focal, sans même recourir à l'ajout d'une aberration connue servant de porteuse comme pour SICLOPE (*cf.* section 3.4.1). Dans le cadre de la thèse de Fabien Baron [T8], j'ai initié le développement d'un estimateur analytique de « phase retrieval » fondé sur le maximum de vraisemblance pour cette application. Cet estimateur, baptisé FUSCHIA (Fast Unambiguous Sensor for CopHasing Interferometric Arrays), est exact si l'on ne recherche que les pistons, et recourt à l'approximation des faibles phases si l'on recherche également les tip-tilts [A17]. FUSCHIA a été appliqué avec succès à l'étude DWARF décrite ci-dessous.

Le concept retenu par l'ESA pour le senseur de cophasage de l'interféromètre spatial de la mission DARWIN a été le capteur plan focal proposé par notre équipe. Les spécifications de ce capteur, baptisé DWARF (DarWin AstRonomical Fringe sensor), sont sévères : il doit mesurer les 3 premiers modes de Zernike sur chacun des télescopes de DARWIN (6 dans la configuration initialement sélectionnée) avec une précision nanométrique et une cadence de 10 Hz, et mesurer également les aberrations individuelles de plus hauts ordres des télescopes composant l'interféromètre, jusqu'à l'aberration sphérique. Il s'agit donc autant d'un SC que d'un analyseur de surface d'onde. L'originalité du concept proposé par nos soins est de combiner les deux approches au sein d'un même capteur plan focal. Ce concept s'appuie, pour les trois premiers modes, sur le traitement direct de l'image plan focal par l'algorithme linéaire FUSCHIA [A17] et, pour les modes supérieurs, sur l'utilisation d'une image supplémentaire défocalisée et de la diversité de phase.

Les résultats expérimentaux ont été à la hauteur des attentes : d'une part la répétabilité nanométrique spécifiée pour le piston et le tip-tilt de DWARF ($0{,}75$ et $1{,}21$ nm respectivement) est atteinte à fort flux [A17] ; d'autre part l'estimation des hauts ordres par diversité de phase a également atteint sa spécification (10 nm)[C42] et les mesures ont pu être validées de manière croisée avec celles délivrées par un analyseur interférométrique commercial Zygo, visible en haut à gauche de la figure 3.1.

Chapitre 4

Travaux en imagerie

4.1 Introduction

Le traitement des images est un maillon essentiel de la chaîne d'acquisition et de traitement des systèmes d'observation HRA actuels et futurs pour obtenir l'information recherchée sur l'objet observé [L4] :

- en imagerie astronomique sans OA, le traitement conjoint des images et de mesures de front d'onde permet de compenser les effets de la turbulence (section 4.2) ;
- en imagerie astronomique par OA, le recalage (section 4.3) et la restauration (section 4.4) des images permettent de compenser les résidus de correction ou les effets de l'anisoplanétisme ;
- en imagerie coronographique par OA, aussi appelée imagerie à haute dynamique, des techniques d'estimation et de détection optimales permettent d'extraire l'information concernant la planète à détecter (section 4.5) ;
- en imagerie de la rétine par OA, la restauration des images à réaliser est tridimensionnelle, car le processus d'imagerie l'est (section 4.6) ;
- en interférométrie optique, classique (section 4.7) ou coronographique (section 4.8), contrairement au cas de la restauration d'image les données fournies par l'observation ne sont pas des images et sont tout simplement inexploitables sans un traitement approprié.

Les techniques mises en œuvre sont généralement fondées sur une approche bayésienne qui permet de prendre en compte non seulement les connaissances *a priori* sur la statistique du bruit de mesure mais aussi celles, mêmes qualitatives ou partielles, sur les paramètres recherchés – turbulence et/ou objet observé par exemple. Ces dernières permettent d'assurer ou de renforcer la robustesse de la solution obtenue vis-à-vis de bruits de mesure, inévitablement présents dans les données expérimentales.

Une caractéristique commune à ces problèmes de traitement, outre d'être mal posés, est que la réponse de l'instrument est souvent imparfaitement connue, et que cette méconnaissance partielle doit être prise en compte explicitement, et spécifiquement pour chaque problème, afin d'aboutir à une exploitation satisfaisante des données. En imagerie monopupille à travers la turbulence,

cette *déconvolution myope* prendra des formes différentes selon que l'on corrige la turbulence *a posteriori* par DECASO (section 4.2) ou en temps réel par OA (section 4.4).

4.2 Déconvolution par analyse de front d'onde

Le principe de cette technique d'imagerie a été rappelé dans la section 2.2.4. Le traitement des données de cette technique est un double problème inverse : estimation à partir des mesures de l'ASO des fronts d'ondes, et estimation de l'objet à partir des images et des mesures ASO.

Le traitement classique de ces données est séquentiel : on estime les fronts d'ondes à partir des mesures de l'ASO (*cf.* paragraphe 2.3.2), on calcule les réponses impulsionnelles instantanées correspondants aux fronts d'ondes estimés, puis on effectue une déconvolution multitrame non myope, c'est-à-dire en considérant comme vraies ces réponses impulsionnelles.

Un tel traitement ne permet pas d'obtenir des résultats satisfaisants en pratique sur des données expérimentales, et ce pour deux raisons : d'une part les mesures ASO sont bruitées, donc les réponses impulsionnelles qui s'en déduisent le sont également, d'autre part même en l'absence de bruit ces mesures sont biaisées par la présence d'aberrations différentielles (ou NCPA en Anglais) entre la voie ASO et la voie d'imagerie.

J'ai donc proposé, avec quelques collègues, une déconvolution myope, c'est-à-dire un traitement conjoint des données ASO et images, qui estime conjointement au sens du MAP l'objet observé et les fronts d'onde turbulents, et permet d'améliorer notablement l'estimation de l'objet observé. Les détails de la méthode sont donnés dans [A34, Annexe L page 317].

Ces traitements ont été appliqués à des images expérimentales de l'étoile double Capella enregistrées le 8 novembre 1990 avec un banc de l'ONERA appelé DECASO installé sur le télescope William Herschel (La Palma, îles Canaries) de 4,20 m de diamètre. La figure 4.1 présente les restaurations obtenues. A gauche, avec le traitement séquentiel, le caractère binaire de Capella est visible, mais reste noyé dans de fortes fluctuations. A droite, la déconvolution myope permet d'éliminer quasiment tous les artefacts des déconvolutions non myopes. Dans les deux cas, la même régularisation objet quadratique sous contrainte de positivité est utilisée, avec une DSP constante de valeur déduite du flux mesuré.

4.3 Recalage d'images

Le recalage des images est une problématique souvent rencontrée en imagerie, et en particulier en HRA. En imagerie de l'Espace depuis le sol dans l'infra-rouge sur NAOS-CONICA par exemple, on est amené à enregistrer un grand nombre de poses relativement courtes afin que le fond de ciel ne sature pas le détecteur. Dans d'autres contextes comme l'imagerie endo-atmosphérique, du fait des vibrations résiduelles, seules des poses très courtes permettent de conserver toute la résolution amenée par l'OA. Dans le cadre de la thèse de Damien Gratadour [T7], nous avons développé une méthode de recalage [A23, Annexe I page 279] fondée

FIG. 4.1 – Images expérimentales de Capella déconvoluées : à gauche, estimation des fronts d'ondes par MAP puis déconvolution quadratique ; à droite, déconvolution myope. Dans les deux cas, l'*a priori* utilisé est gaussien, de DSP constante déduite du flux mesuré, avec une contrainte de positivité. Conditions expérimentales : flux de $67\,500$ photons par image, temps de pose de 5 ms, D/r_0 de 13 et un RSB sur l'ASO de 5.

sur le maximum de vraisemblance (MV) adaptée à la problématique HRA c'est-à-dire, plus précisément, qui possède les caractéristiques suivantes :

- une précision sub-pixellique arbitrairement petite, limitée uniquement par le bruit, afin de conserver la résolution ultime du télescope ;
- la prise en compte, par le cadre MV, de la statistique du bruit, comportant une composante photonique poissonienne et une composante gaussienne liée au détecteur ;
- un recalage conjoint de l'ensemble des images, plutôt qu'un recalage deux à deux, lequel est sous-optimal en terme de performances à faible RSB ;
- la possibilité de prendre en compte des pixels morts.

Cette méthode et son application à un cas réel d'imagerie IR corrigée par OA d'une galaxie lointaine (Arp 220) sont détaillées dans [A23, Annexe I page 279]. La figure 4.2, extraite de cet article, illustre l'intérêt de ce recalage des images par MV et les performances de la méthode.

Elle est depuis utilisée pour de multiples applications, de l'imagerie de la rétine [T2] à l'imagerie IR endo-atmosphérique [45]. Elle a en particulier été utilisée dans le pré-traitement des données NAOS-CONICA qui ont révélé une planète géante autour de β-pictoris [A13], et sera vraisemblablement utilisée dans le projet SPHERE pour étalonner précisément le centrage de l'étoile sur le coronographe.

4.4 Restauration myope d'images longue pose corrigées par OA

4.4.1 Principe

La technique d'imagerie à travers la turbulence la plus efficace est l'OA, dont un bref rappel historique est fait paragraphe 2.2.4. Néanmoins les images longue pose corrigées par OA doivent

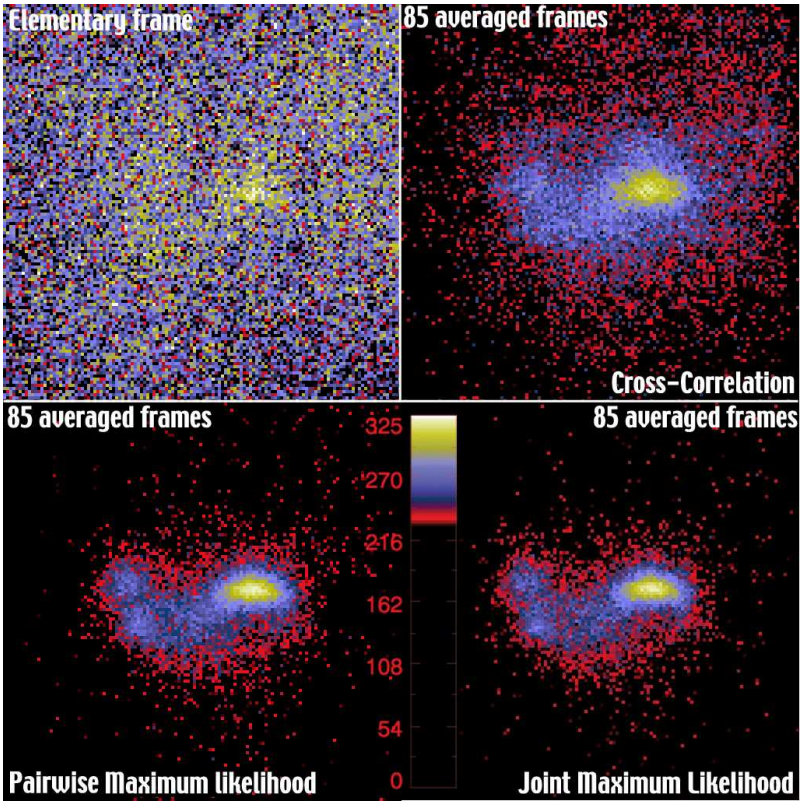


FIG. 4.2 – Image corrigée par OA en bande L obtenue avec NAOS-CONICA au VLT : en haut à gauche, une image élémentaire ; en haut à droite, résultat d'une corrélation croisée classique suivie d'une somme des 85 images ; en bas à gauche, résultat du recalage par la méthode MV développée appliquée aux 85 images prises deux à deux ; en bas à droite, résultat du recalage par la méthode MV développée appliquée conjointement aux 85 images [A23, Annexe I page 279].

être déconvoluées, car la correction réalisée n'est que partielle [22].

L'estimation bayésienne est un cadre naturel pour traiter le caractère mal posé de la déconvolution. Si l'on suppose la réponse impulsionnelle connue, l'objet estimé au sens du MAP est celui qui minimise un critère mixte composé d'un terme de fidélité aux données, noté par exemple J_i, et d'un terme de régularisation, noté par exemple J_o.

Afin de restaurer des objets à grande dynamique, fréquents en astronomie, le terme de fidélité aux données J_i doit incorporer une modélisation fine du bruit prenant en compte à la fois le bruit de photons et le bruit électronique. Ceci peut être réalisé en approximant le bruit de photons comme un bruit gaussien non stationnaire et aboutit à un critère J_i du type moindres carrés pondérés plutôt qu'à un terme de moindres carrés ordinaires (*cf.* [L4] par ex.).

Pour des objets à bords francs comme des satellites artificiels, des astéroïdes ou des planètes, un *a priori* gaussien, ou de manière équivalente un critère de régularisation quadratique, a tendance à lisser les bords et à introduire près de ceux-ci des oscillations parasites ou *ringing*. Une solution est alors d'utiliser un critère préservant les bords francs (ou *edge-preserving*) comme les critères dits quadratiques-linéaires, qui sont quadratiques pour les faibles sauts et linéaires pour les forts sauts. La partie quadratique assure un bon lissage du bruit et la partie linéaire annule la pénalisation des bords – voir les chapitres 6 et 10 de [2] pour plus de détails. J'ai proposé une version isotrope d'un tel critère [A34, Annexe L page 317], qui évite complètement les effets de bloc particulièrement gênants en astronomie.

Par ailleurs, pour de nombreuses raisons, on est souvent amené à considérer que la réponse impulsionnelle est imparfaitement connue. Effectuer une déconvolution *classique* c'est-à-dire en supposant la réponse impulsionnelle connue mais avec une réponse impulsionnelle fausse peut conduire à des résultats catastrophiques. À l'inverse, la déconvolution dite aveugle, où l'on minimise le même critère mais en recherchant simultanément objet et réponse impulsionnelle, est très instable, à l'instar des méthodes non régularisées. La *déconvolution myope* consiste à estimer conjointement l'objet $\boldsymbol{o}$ et la réponse impulsionnelle $\boldsymbol{h}$ dans un cadre bayésien avec une régularisation naturelle pour la réponse impulsionnelle et sans avoir à régler d'hyper-paramètre supplémentaire pour la réponse impulsionnelle.

L'estimateur MAP conjoint est donné par :

$$\begin{aligned}(\hat{\boldsymbol{o}}, \hat{\boldsymbol{h}}) &= \arg\max_{\boldsymbol{o},\boldsymbol{h}} p(\boldsymbol{o}, \boldsymbol{h} \mid \boldsymbol{i}) = \arg\max_{\boldsymbol{o},\boldsymbol{h}} p(\boldsymbol{i} \mid \boldsymbol{o}, \boldsymbol{h}) \times p(\boldsymbol{o}) \times p(\boldsymbol{h}) \\ &= \arg\min_{\boldsymbol{o},\boldsymbol{h}} \left(J_i(\boldsymbol{o}, \boldsymbol{h}) + J_o(\boldsymbol{o}) + J_h(\boldsymbol{h}) \right), \end{aligned} \tag{4.1}$$

où J_h est un critère de régularisation sur $\boldsymbol{h}$, qui introduit des contraintes sur la variabilité possible de la réponse impulsionnelle. La réponse impulsionnelle longue pose peut être considérée comme la somme d'un grand nombre de réponses courtes poses indépendantes, et donc modélisée par un *a priori* gaussien (tronqué aux valeurs positives). On considère de plus que la différence entre la réponse impulsionnelle et la réponse impulsionnelle moyenne est approximativement stationnaire ; la régularisation de la réponse impulsionnelle est alors une pénalisation quadratique de la fonction de transfert, indépendante entre fréquences [A40, A39, A27] :

La méthode de restauration MISTRAL, décrite dans [A27, Annexe K page 301], combine les trois ingrédients évoqués ci-dessus : la modélisation fine du bruit, la régularisation non quadratique et l'aspect myope.

Lors de la thèse de Damien Gratadour [T7], nous avons développé une estimation dite non-supervisée des hyper-paramètres de MISTRAL dans le cas d'une régularisation gaussienne *via* un modèle de spectre objet motivé physiquement [A20, annexe A].

La méthode MISTRAL a été utilisée par plusieurs équipes astronomiques dans le monde sur divers télescopes, en particulier [A32, A25]. Le paragraphe suivant présente quelques résultats obtenus avec cette méthode sur données expérimentales. Outre la restauration d'images anisoplanétiques mentionnée plus loin, deux axes au moins méritent d'être poursuivis pour prolonger ces travaux : d'une part envisager un estimateur bénéficiant de bonnes propriétés théoriques pour la déconvolution myope, d'autre part estimer les hyper-paramètres de la restauration afin d'aboutir à une restauration non supervisée. Ces deux axes peuvent être traités simultanément soit en adoptant un estimateur marginal (de la FEP ou de l'objet) comme nous l'avons fait en diversité de phase [A29], soit en adaptant à la HRA des travaux récents [55, 56] fondés sur l'estimateur de la moyenne *a posteriori*.

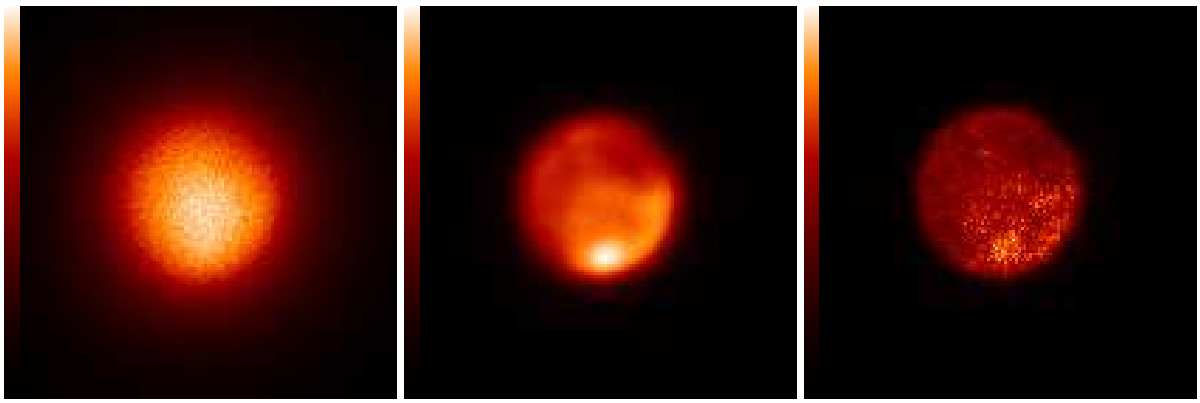
4.4.2 Restauration à partir de données astronomiques expérimentales

Restauration d'images de Ganymède

L'image de la figure 4.3a montre une longue pose corrigée par OA de Ganymède, satellite de Jupiter. Cette image a été enregistrée le 28/09/1997 sur le banc d'OA de l'ONERA installé sur le télescope de $1{,}52$ m de l'Observatoire de Haute-Provence. La longueur d'onde d'imagerie est $\lambda = 0{,}85\,\mu$m et le temps de pose 100 sec. Le flux total estimé est 8×10^7 photons et le rapport D/r_0 estimé est 23. Le champ total est de $7{,}9$ arcsec, dont seulement la moitié est montrée ici. La réponse impulsionnelle moyenne et sa variabilité ont été estimées à partir de l'enregistrement de cinquante images d'une étoile brillante située à proximité. Les figures 4.3b et c montrent les restaurations obtenues par l'algorithme de Richardson-Lucy (MV pour un bruit de Poisson), interrompu à 200 et 3 000 itérations respectivement[1]. Dans le premier cas, similaire à une restauration avec régularisation quadratique, l'image restaurée est assez floue et présente un léger « ringing », et dans le second cas, très similaire au résultat d'un filtrage inverse, le bruit domine la restauration.

L'image figure 4.4a illustre la déconvolution myope [A27, Annexe K page 301] avec *a priori* préservant les bords. La figure figure 4.4b montre une image synthétique large bande obtenue à partir de clichés d'une sonde spatiale NASA/JPL (voir `http://space.jpl.nasa.gov/`) passée à proximité de Ganymède. La comparaison montre que de nombreuses caractéristiques de Ganymède sont correctement restaurées. Une comparaison plus équitable consiste à examiner

[1] L'arrêt d'un algorithme non régularisé avant convergence est une méthode de régularisation encore répandue mais très *ad hoc*, *cf.* [2, chap. 2].



(a) image corrigée par OA (b) Richardson-Lucy, 200 it. (c) Richardson-Lucy, 3 000 it.

FIG. 4.3 – (a) Image corrigée de Ganymède, obtenue avec le banc d'OA de l'ONERA, le 28 septembre 1997. (b) Restauration par Richardson-Lucy interrompue à 200 itérations ; (c) idem à 3 000 itérations.

conjointement la déconvolution myope effectuée par MISTRAL avec l'image de la figure 4.4b convoluée par la réponse impulsionnelle parfaite d'un télescope de $1{,}52$ m, présentée figure 4.4c.

(a) déconvolution par MISTRAL (b) base de données JPL (NASA/JPL/Caltech) (c) image (b) + réponse impulsionnelle du télescope parfait

FIG. 4.4 – (a) Déconvolution par MISTRAL de l'image de Ganymède de la figure 4.3. (b) En comparaison, une image synthétique large bande obtenue grâce à la base de données NASA/JPL. (c) Même image synthétique convoluée par la réponse impulsionnelle parfaite d'un télescope de $1{,}52$ m de diamètre.

Restauration d'images de β-pictoris

Tout récemment, MISTRAL a également été utilisée pour déconvoluer les images NAOS-CONICA qui ont révélé une planète géante autour de β-pictoris, *cf.* [A13] et communiqué de

presse ESO sur http://www.eso.org/public/news/eso0842/.

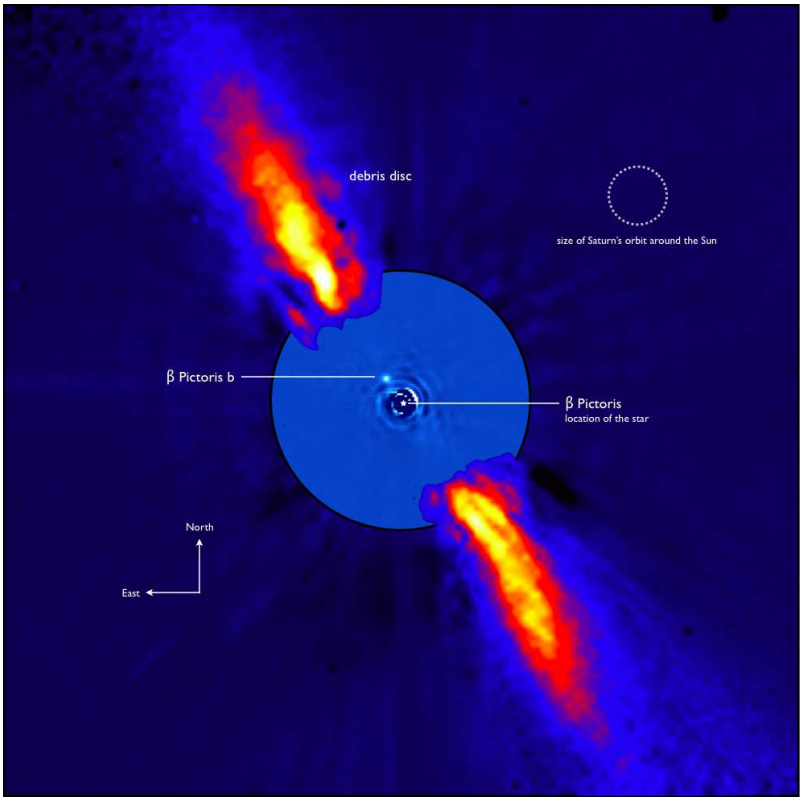


FIG. 4.5 – Image composite de l'environnement de l'étoile β-pictoris. Le centre de l'image est une image NAOS-CONICA déconvoluée à l'aide de MISTRAL. L'extérieur de l'image (poussière) provient du télescope de 3,6 m de l'ESO corrigé par l'OA du système ADONIS [A13].

4.4.3 Restauration d'images en présence d'anisoplanétisme

Les travaux de restauration d'images présentés ci-dessus s'appuient sur un modèle d'imagerie convolutif et leur application directe est donc limitée au domaine isoplanétique de l'OA.

Pour effectuer une restauration sur un champ plus grand il faut d'une part savoir modéliser le problème direct c'est-à-dire la variation de la FEP dans le champ, d'autre part gérer astucieusement le coût de calcul supplémentaire que représente cette variabilité. Lors de la thèse de Thierry Fusco, nous avons pu modéliser analytiquement cette variation de la FEP dans le champ, et nous avons mis en œuvre ce modèle dans un algorithme de restauration de champ d'étoiles [A36]. Ce type d'étude devra être étendu aux OAs à grand champ, en particulier l'OAMC ; pour une telle OA, les mesures de front d'onde seront multi-directionnelles [A33] et leurs statistiques pourront être utilisées pour estimer la variation de la FEP dans le champ.

4.5 Détection d'exoplanètes par imagerie coronographique

La détection directe et la caractérisation spectrale d'exo-planètes depuis le sol est un objectif majeur de l'astronomie actuelle. Cet objectif représente un défi technologique car même pour des planètes géantes et relativement chaudes, le rapport d'intensité, ou contraste, entre l'étoile hôte et sa planète peut être de l'ordre de 10^6 dans l'IR proche. Le consortium européen SPHERE, pour *Spectro-Polarimetry High-contrast Exoplanet Research*, construit actuellement un système éponyme de seconde génération pour le VLT dans ce but [C41]. Ce système combine une optique adaptative de haute performance appelée SAXO [57] qui concentre la lumière de l'étoile et un coronographe qui atténue fortement celle-ci et réduit donc significativement le bruit de photons.

Cette combinaison n'est malheureusement pas suffisante pour détecter les planètes d'intérêt pour les contrastes envisagés. La limitation principale est la présence dans l'image de tavelures ou *speckles* quasi-statiques dues aux aberrations statiques résiduelles, qui ont une taille du même ordre de grandeur qu'une planète (λ/D). Pour pouvoir distinguer le signal de l'éventuelle planète des speckles résiduels, il faut utiliser des informations supplémentaires. Ces informations peuvent prendre différentes formes selon l'instrument :

- information spectrale sur l'objet observé : le système SPHERE comprend en fait trois instruments, dont un imageur à deux canaux spectraux simultanés dénommé IRDIS. En faisant l'hypothèse de présence de méthane dans l'atmosphère de la planète recherchée et en choisissant judicieusement les longueurs d'ondes centrales, des techniques d'imagerie différentielle spectrale permettent d'éliminer en grande partie l'influence des aberrations commune aux deux canaux. Pour atteindre les contrastes visés cela n'est pas suffisant ;
- information temporelle : dans un système comme SPHERE, la pupille est stabilisée au cours de la nuit pour que les aberrations, donc les speckles, restent fixes. Par conséquent, du fait de la rotation terrestre, le champ tourne et toute planète avec lui. L'exploitation de cette différence de comportement temporel rend possible la séparation planète(s)/speckles. Nous avons développé, essentiellement dans le cadre de la thèse d'Alberto Cornia [T1], une méthode de détection traitant conjointement l'ensemble des images d'un objet observé et exploitant cette information temporelle de manière optimale [C26, C23, C19], [A10], après recombinaison éventuelle des deux canaux spectraux [C13, C14]. Pour plus de détails sur la méthode on pourra consulter [A10, Annexe E page 233] et [C14]. En utilisant conjointement ces informations spectrale et temporelle, ainsi que les informations disponibles sur la statistique du bruit et une contrainte de positivité sur le flux, nous atteignons, sur des données simulées de manière réaliste par le consortium SPHERE, une détection de planètes sans fausse alarme pour des planètes même proches de l'étoile ($4\lambda/D$) et un contraste étoile/planète de 10^6 [C13, C14], conformément aux spécifications du système SPHERE. La figure 4.6 illustre ces capacités de détection pour des séparations étoile-planète de $0{,}2$ à 1 arcsec.
  La méthode développée, appelée ANDROMEDA pour *ANgular DiffeRential OptiMal Exoplanet Detection Algorithm*, va être incluse dans le *pipeline* de traitement du système SPHERE et être disponible pour tous les utilisateurs du système.

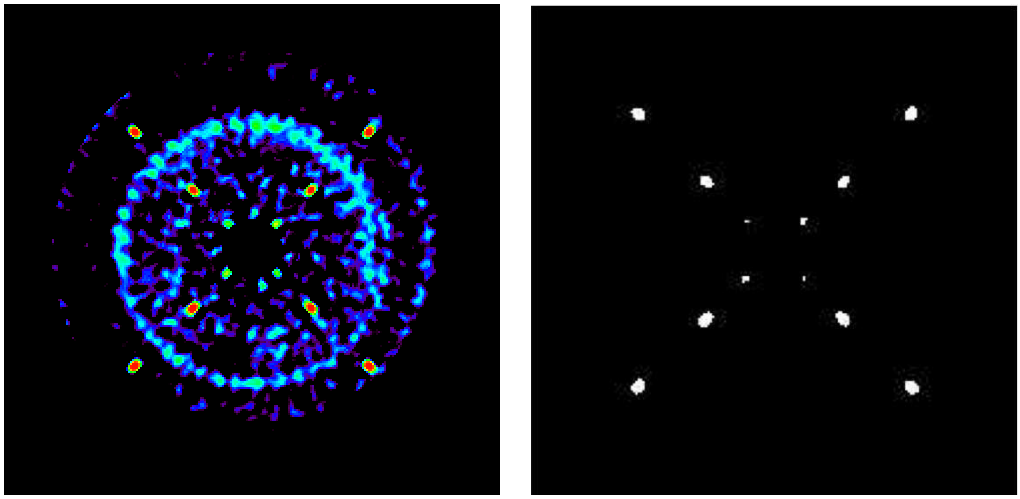


FIG. 4.6 – Détection d'exo-planètes à partir d'une série d'images dans deux canaux spectraux proches de $1{,}6\,\mu$m. Le temps de pose total est de 4 heures, les simulations sont réalisées par le consortium SPHERE en prenant en compte toutes les connaissances actuelles sur le système. Gauche : carte de vraisemblance des 12 planètes simulées, situées à $0{,}2$, $0{,}5$ et 1 arcsec de l'étoile. Droite : carte seuillée à 3 écarts-types. D'après [C14].

– information sur la réponse spectro-spatiale d'un instrument coronographique : une fois les planètes détectées avec IRDIS, elles seront caractérisées avec un IFS ou *Integral Field Spectrograph*, qui possède de nombreux canaux spectraux. Avec un tel instrument, on devrait pouvoir estimer de manière fiable à la fois le spectre objet et le champ de speckle. On exploitera pour cela d'une part la connaissance de l'évolution spectro-spatiale du champ de speckles, *cf.* [T4], [58] et [A5] et d'autre part la douceur spectrale éventuelle de l'objet observé. C'est l'objet de la thèse de Marie Ygouf (2009–2012), en collaboration avec l'IPAG, que je coencadre.
– informations apportées par un sous-système complémentaire : pour améliorer encore les capacités de détection d'un système comme SPHERE, il serait utile de mesurer les aberrations quasi-statiques en ligne, c'est-à-dire *pendant* la pose scientifique, afin soit de les corriger en temps réel, soit de prendre en compte dans le traitement les tavelures qu'elles créent au plan focal. La diversité de phase [L6, Annexe B page 133] permet de réaliser cette mesure des aberrations quasi-statiques seules, en longue pose sur le ciel malgré la turbulence atmosphérique résiduelle, *cf.* [A15, Annexe G page 257] et section 3.3.2, et ce éventuellement en temps-réel, *cf.* [A9, Annexe D page 229] et section 3.2. Le capteur peut être placé juste avant le coronographe afin de ne mesurer que les aberrations influençant directement l'efficacité du coronographe, comme envisagé pour la mise à jour du système SPHERE et étudié par l'équipe dans le cadre du contrat FP7 /JRA 1.
  Il pourrait également être constitué directement du plan focal scientifique, donc après le coronographe : j'ai proposé d'étendre la diversité de phase à l'imagerie coronographique grâce au modèle d'imagerie coronographique que Jean-François Sauvage et moi avons

développé à la fin de sa thèse [T4] et ensuite [A5]. Les premiers résultats de ce nouvel analyseur de front d'onde sont très encourageants [59].

4.6 Restauration d'images pour l'imagerie rétinienne

La détection précoce de pathologies rétiniennes aussi répandues que la DMLA, les glaucomes ou les rétinopathies diabétiques réclament une exploration *in situ* et *in vivo* du tissu rétinien à l'échelle cellulaire. Or l'examen direct depuis l'extérieur de l'œil souffre des aberrations optiques du segment antérieur (cornée et cristallin), qui limitent la résolution accessible. La mesure et la correction de ces aberrations sont possibles grâce à l'utilisation de l'OA. En imagerie plein champ, le caractère tridimensionnel de l'objet d'intérêt (la rétine) rend l'interprétation des images difficile puisque tous les plans qui constituent l'objet contribuent à la formation de chaque plan image. De plus, la correction par OA est toujours partielle.

4.6.1 Restauration 2D

Les imageurs plein champ actuels n'enregistrent qu'un seul plan image et leur RI est mal connue. Afin de restaurer correctement les images acquises par ces dispositifs malgré le manque d'information, nous avons développé, dans le cadre de la thèse de Leonardo Blanco (2009-2012), une méthode myope qui prend en compte le caractère 3D de l'imagerie en faisant l'approximation, raisonnable pour les photorécepteurs, que l'objet imagé est invariant par translation le long de l'axe optique. Ceci amène à rechercher un objet 2D et une RI qui est la combinaison linéaire des RIs associées à chaque plan. Nous avons montré théoriquement et vérifié par simulation que l'estimation conjointe conduit à un critère dégénéré et nous avons développé une estimation marginale inspirée de nos travaux en diversité de phase [L6], qui s'avère performante et en accord avec les propriétés attendues de l'estimateur MV. Cette méthode a été validée sur données expérimentales à la satisfaction des médecins et est actuellement testée sur plusieurs dizaines de patients dans le cadre du projet iPhot soutenu par l'ANR. La figure 4.7 montre un exemple de résultat sur les premières données utilisées, qui proviennent de l'imageur du LESIA [C2] installé au CIC de l'Hôpital des Quinze-Vingts. Les données de la campagne de test actuelle sont issues de l'imageur d'Imagine Eyes, également installé au CIC. Cette méthode fait l'objet d'une publication imminente [A2, Annexe C page 215].

4.6.2 Restauration 3D

Un dispositif ambitieux permettant de combiner la haute résolution latérale et la sélection d'une couche dans la rétine avec une bonne résolution longitudinale est en cours de développement, avec la collaboration de notre équipe [C22, C17] : il s'agit de la combinaison OA + OCT plein champ. Sans attendre qu'un tel dispositif soit opérationnel, une alternative consistera, dans un futur proche, à effectuer une déconvolution tridimensionnelle (3D) des images enregistrées

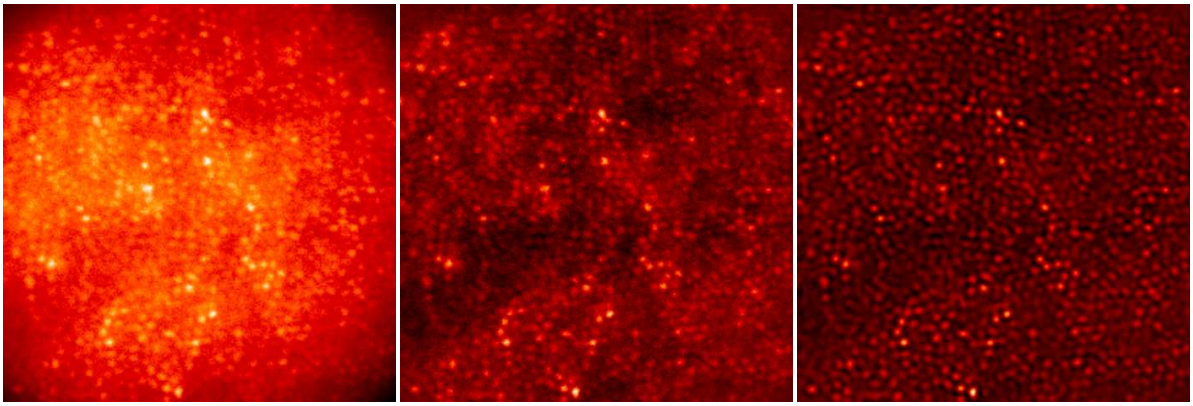


FIG. 4.7 – Gauche : image brute acquise sur l'imageur du LESIA au CIC (à $1{,}2°$ du centre de la fovéa). Centre : image après soustraction du fond estimé. Droite : image déconvoluée par notre méthode [A2].

par un imageur classique parcourant la rétine longitudinalement, afin d'une part de séparer numériquement les plans de l'objet et d'autre part d'améliorer la résolution latérale. Dans le cadre de la thèse de Guillaume Chenegros [T2], nous avons développé une méthode de déconvolution 3D [C38] et nous nous sommes intéressés à deux aspects importants dans ce contexte : pour donner des résultats satisfaisants, une méthode de déconvolution nécessite généralement, et particulièrement en 3D, d'une part une régularisation par un *a priori* adapté et un ajustement des paramètres de celui-ci (ou hyper-paramètres), d'autre part une bonne connaissance de la réponse impulsionnelle du système complet œil+instrument.

En ce qui concerne le premier aspect, nous avons proposé une régularisation prenant en compte le fait que les différents plans de l'objet observé peuvent être d'intensité et de contenu spectral très différents, et nous avons développé une technique d'estimation non supervisée (automatique) des hyper-paramètres [C18, C16]. Cette dernière, développée dans le cadre du projet INOVEO soutenu par le RNTS de l'ANR, permet d'envisager une utilisation efficace de la déconvolution 3D même par des utilisateurs peu familiers du traitement des images tels que médecins ou biologistes. Nous avons développé à l'ONERA un banc d'imagerie 3D sur lequel nous avons pu valider expérimentalement cette déconvolution 3D non supervisée sur un objet non biologique bien maîtrisé, actuellement constitué d'une règle graduée inclinée par rapport à l'axe optique. La figure 4.8 illustre le gain notable en résolution longitudinale apporté par la déconvolution, tout à fait compatible avec les performances visées pour l'imagerie rétinienne.

Le second aspect que nous avons traité est le fait que la réponse impulsionnelle du système œil+instrument est actuellement très mal connue. Nous avons développé une extension tridimensionnelle de la technique de diversité de phase qui permet d'estimer la réponse du système conjointement à l'objet d'intérêt, *cf.* [A18, Annexe H page 269]. Alors que ce type d'estimation conjointe a généralement de mauvaises propriétés statistiques, nous avons pu montrer que dans le contexte de la diversité de phase l'estimation conjointe conduisait à un estimateur consistant

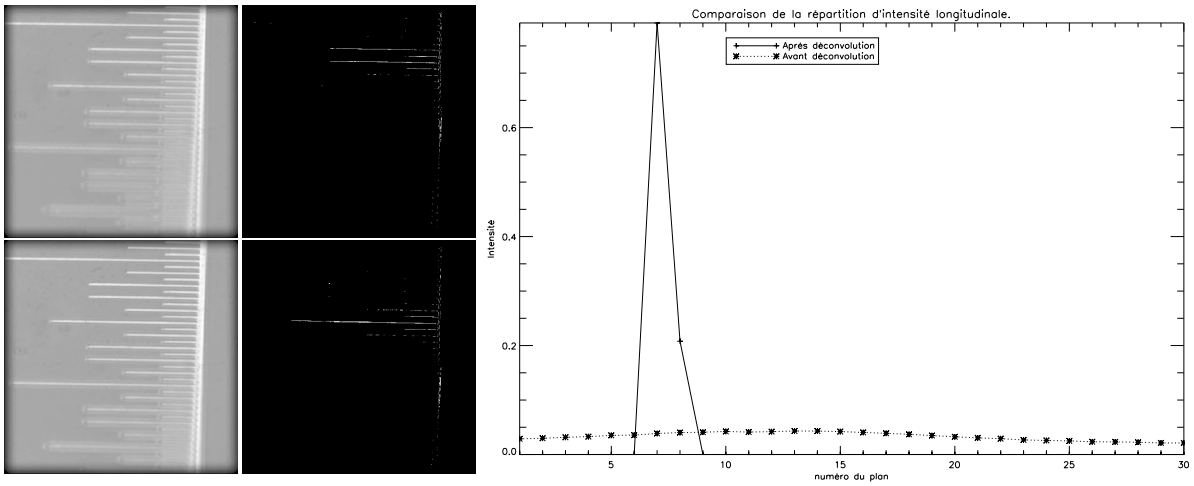


FIG. 4.8 – Validation expérimentale de la déconvolution 3D non supervisée. À gauche (haut et bas), deux images parmi trente d'une règle graduée, focalisées dans des plans distants de 14 µm. Au centre (haut et bas), les deux images correspondantes de la pile d'images déconvoluée. À droite, une coupe longitudinale dans l'un des traits de la règle avant et après déconvolution illustre le gain en résolution longitudinale ; largeur à mi-hauteur après déconvolution : 1,5 plans, soit 3 µm.

des aberrations [A21, Annexe J page 289].

Du fait du contexte opérationnel difficile (mouvements oculaires, clignements, sauts de focus de l'OA, etc.) une étape clé, avant de passer à l'application de la diversité de phase 3D à des images de patients, sera de valider cette technique sur des images expérimentales non biologiques.

Une perspective à plus long terme, lorsque le système OA + OCT plein champ sera opérationnel, serait d'appliquer la déconvolution 3D aux images issues de ce système et ainsi d'en améliorer encore la résolution, dans les trois dimensions.

4.7 Reconstruction d'images en interférométrie optique depuis le sol

Les objectifs scientifiques des futures missions astronomiques, notamment la détection d'exoplanètes ou l'étude des noyaux actifs de galaxies, nécessitent des résolutions angulaires de l'ordre de la milliseconde d'arc hors de portée des télescopes actuellement en fonctionnement. L'interférométrie optique est une solution permettant d'améliorer considérablement la résolution, donnée par l'espacement entre télescopes et non plus par le diamètre de ceux-ci.

Le principe de l'interférométrie optique a été rappelé dans le paragraphe 2.2.2. Le problème inverse de la reconstruction d'image à partir de données d'interféromètre optique au sol [L4] est difficile pour au moins trois raisons :

- D'abord à cause du faible nombre de données : chaque couple de télescopes mesure uniquement un coefficient de Fourier de l'objet. On a donc typiquement quelques dizaines, éventuellement quelques centaines de mesures sur plusieurs nuits d'observation. Pour reconstruire l'objet sur plusieurs milliers de pixels, il est donc nécessaire de régulariser l'estimation ;
- ensuite, parce qu'en raison de la turbulence atmosphérique, les interféromètres optiques actuels sont affectés par des phases turbulentes inconnues sur chaque télescope. Les échantillons de Fourier complexes mesurés sont donc multipliés par des phaseurs inconnus correspondant aux pistons turbulents différentiels sur chaque couple de télescopes, comme expliqué dans le paragraphe 2.2.3 ;
- enfin, parce que les codes développés jusqu'à très récemment ont été conçus pour la radioastronomie (domaine dans lequel l'interférométrie est une technique mature) et ne modélisent pas correctement la statistique du bruit dans des données optiques, visibles ou IR.

Dans le cadre de la thèse de Serge Meimon [T6], nous avons développé une méthode de reconstruction dénommée WISARD, décrite en détails dans [A7, Annexe F page 243], qui traite ces trois aspects. On pourra également consulter [L4] pour une synthèse de la méthode, et [A22] pour plus de détails sur le modèle de bruit.

Nous prenons en compte le manque d'information de phase en introduisant des paramètres d'aberrations du système, et reconstruisons une image en minimisant un critère joint original qui dépend à la fois de l'objet et des aberrations. Nous avons développé une stratégie de minimisation inspirée des techniques de radio-astronomie dites d'autocalibration, en tenant compte des spécificités de l'interférométrie optique. Cette méthode exploite notamment une approximation originale du modèle de bruit [A22] qui simplifie l'étape de minimisation du critère, tout en respectant les caractéristiques physiques du bruit.

Par ailleurs, des informations *a priori* sur la solution sont introduites afin de régulariser l'inversion. En particulier j'ai proposé avec mon collègue Éric Thiébaut (CRAL) un *a priori* dit de *soft support* très approprié pour l'interférométrie, qui permet de favoriser un support limité et permet de réaliser efficacement de l'extrapolation spectrale bien qu'il soit quadratique [A16]. Ceci est attesté figure 4.9 par la similitude des reconstructions obtenues entre un *a priori* linéaire-quadratique blanc et l'*a priori* de *soft support* proposé : toutes deux permettent de faire ressortir l'étoile centrale au milieu du disque de poussière (données issues du *Imaging Beauty Contest* 2004 [C54]).

Une comparaison entre les méthodes de type autocalibration, comme WISARD, et les méthodes de reconstruction bayésiennes classiques comme MIRA [60] a été menée avec des collègues, à la fois du point de vue théorique et du point de vue des résultats de reconstruction. Ceux-ci s'avèrent finalement très proches lorsque les *a priori* utilisés sont identiques, bien que le cadre conceptuel de ces méthodes soit assez différent [A16].

Dans le cadre de l'EII (European Interferometry Initiative), j'ai participé à un contrat, financé par le sixième PCRD de l'UE, visant à fournir à la communauté astronomique européenne des outils de traitement de données interférométriques. Avec Serge Meimon, nous avons entièrement

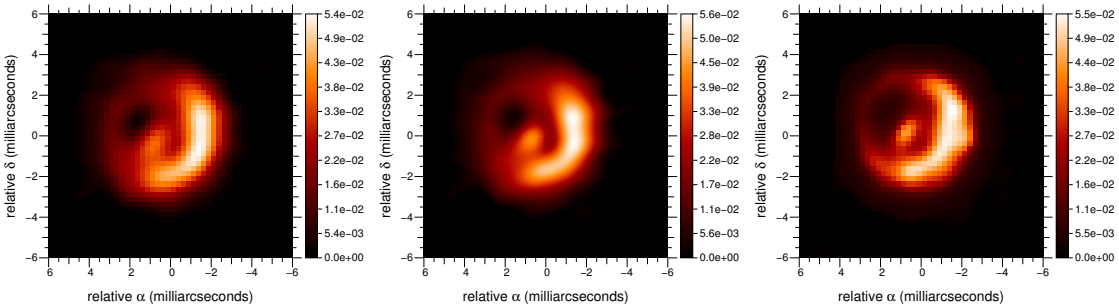


FIG. 4.9 – Reconstruction par WISARD d'une étoile entourée d'un disque de poussière sur données issues du *Imaging Beauty Contest* 2004. Gauche : *a priori* quadratique classique (DSP). Centre : *a priori soft support* [A16]. Droite : *a priori* linéaire-quadratique blanc [A7, Annexe F page 243].

ré-écrit le code WISARD pour en faire un code optimisé en temps de calcul et utilisable par un astronome non expert du traitement de signal [61]. Ce code a été mis sous une licence libre et livré, pour distribution à la communauté astronomique européenne, au JMMC (*Jean-Marie Mariotti Center*, centre français de coordination du traitement des données interférométriques). Il a été exploité sur des données expérimentales et a permis des interprétations astronomiques, notamment sur Arcturus [A14] et sur Bételgeuse [A6], *cf.* Fig. 4.10 et communiqué de presse sur `http://www.grandpublic.obspm.fr/Image-par-interferometrie-Des`.

4.8 Reconstruction d'images multispectrales en interférométrie coronographique depuis l'espace

Le lecteur peu familier de l'interférométrie coronographique, aussi appelée interférométrie à frange noire ou *nulling interferometry*, pourra trouver une introduction à cette technique dans ma contribution [L5] aux cours de l'École thématique CNRS de 2005 intitulée *Troisièmes Journées d'Imagerie Très Haute Dynamique et détection d'exoplanètes*, reproduite en Annexe A.

Mon collègue Éric Thiébaut (CRAL) et moi avons développé, dans le cadre d'un contrat avec TAS pour l'ESA intitulé *Reconstruction of Exo-Solar System Properties* ou RESSP, une méthode originale permettant de détecter et de caractériser spectralement des exo-planètes avec un interféromètre comme DARWIN. Une présentation pédagogique des grandes lignes de la mission DARWIN et de la méthode proposée est donnée dans [L5, Annexe A page 115]. Pour une présentation plus détaillée de la mission et de l'inversion on pourra consulter respectivement [A12] et [C36]. L'idée principale est d'utiliser toutes les informations *a priori* disponibles pour compenser la pauvreté des données.

En modulant la réponse instrumentale par des déphasages entre bras de l'interféromètre, comme imaginé par J.-M. Mariotti dès 1997 (*cf.* [62]), il est possible d'obtenir des cartes de transmissions asymétriques. En recombinant les données astucieusement et avec de telles cartes

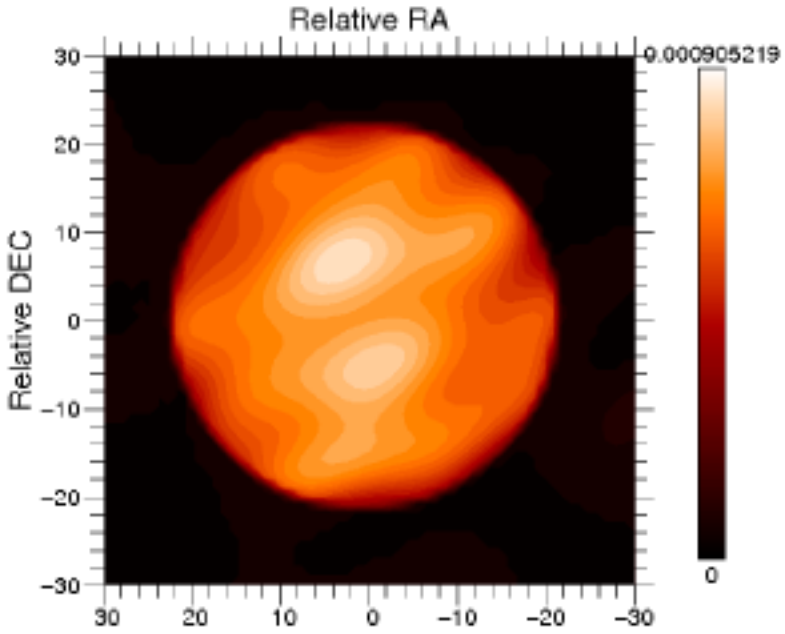


FIG. 4.10 – Surface de Bételgeuse reconstruite par WISARD dans le proche IR à $1{,}64\,\mu$m avec l'interféromètre IOTA. Champ 55 millisecondes d'angle. Ce résultat permet de mieux comprendre la structure et l'évolution des étoiles supergéantes.

de transmission, il est possible d'éliminer la contribution au signal mesuré des composantes de l'objet observé qui ont une distribution spatiale paire : fuites stellaires, lumière exozodiacale, et *a fortiori* lumière zodiacale et émission thermique de l'instrument (qui ont un niveau constant dans le champ). On peut alors ne rechercher, lors de la reconstruction d'image, que les planètes, modélisées par des Diracs, ce qui exprime toute notre information *a priori* spatiale.

Nous avons constaté que, dans les conditions de RSB envisagées, cette information spatiale ne suffisait malheureusement pas toujours à une détection non ambiguë. Nous avons montré que la détection était rendue possible en incorporant une information *a priori* supplémentaire de nature spectrale : l'ajout d'une contrainte sur la régularité des spectres de chaque planète permet d'améliorer non seulement l'estimation des spectres (Fig. 4.11), mais surtout la détection même des planètes (Fig. 4.12) [C36], [L5, Annexe A page 115].

4.9 Optimisation de la configuration pupillaire d'un interféromètre imageur

Dans un télescope monolithique, la surface collectrice et le pouvoir de résolution sont tous deux déterminés par le diamètre du télescope. Dans la conception d'un interféromètre imageur, le positionnement relatif des ouvertures qui composent celui-ci (sous-télescopes ou segments du miroir primaire), appelé configuration pupillaire, est un degré de liberté qui permet de décou-

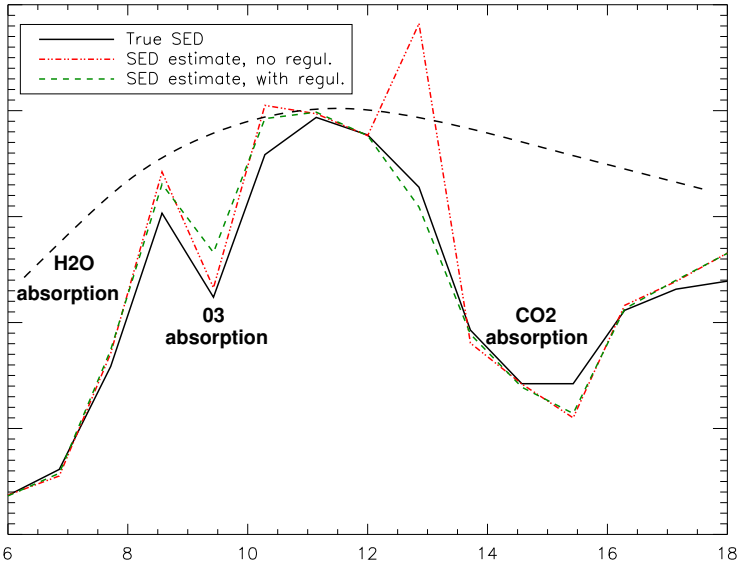
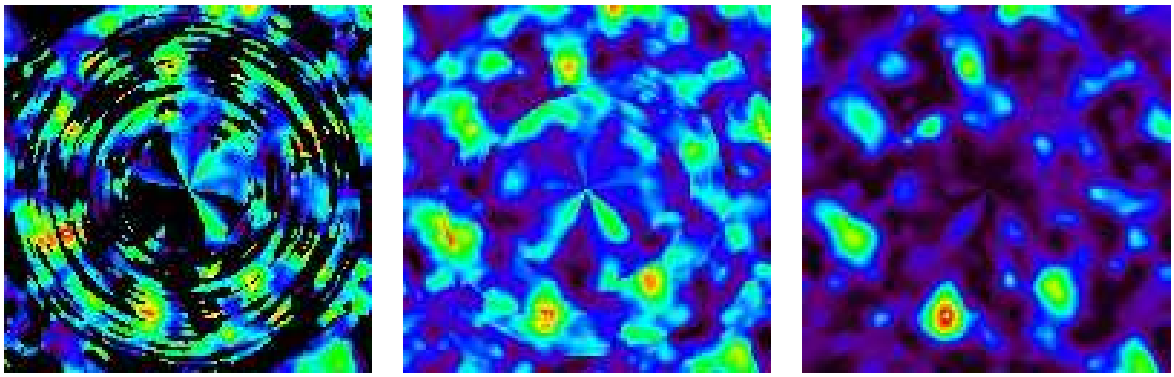


FIG. 4.11 – Spectre d'une Terre reconstruit à la position estimée de la planète : sans régularisation (trait mixte rouge) et avec régularisation (tirets verts), d'après [L5]. Ce spectre est à comparer au spectre vrai (trait plein noir), qui comporte trois bandes d'absorption caractéristiques par rapport au spectre du corps noir (tirets noirs).



sans contrainte | avec contrainte de positivité | positivité + régul. spectrale

FIG. 4.12 – Cartes de vraisemblance pour la position de la planète (d'après [L5] et [C36]). À gauche, vraisemblance seule ; au centre, vraisemblance sous contrainte de positivité des spectres ; à droite, MAP c'est-à-dire vraisemblance pénalisée par un critère de régularisation spectrale. Code couleur : noir correspond à une vraisemblance nulle, rouge à une vraisemblance maximale. La vraie position de la planète est en bas, légèrement à gauche, bien visible sur l'image de droite.

pler ces deux paramètres. L'optimisation de la configuration pupillaire est un aspect clé de la conception d'un interféromètre.

Pour un tel instrument (*cf.* section 2.2.2), le traitement à réaliser est en bonne approximation une déconvolution. Il est indispensable car la réponse impulsionnelle est bien plus irrégulière qu'avec un télescope monolithique du fait de la forme de la pupille.

J'ai développé une méthode d'optimisation de la configuration pupillaire dans la philosophie de la planification d'expérience, c'est-à-dire qui prend en compte l'ensemble de la chaîne d'acquisition et de traitement, et fournit la configuration donnant une erreur minimale d'estimation de l'objet, en moyenne, après restauration.

Cette méthode prend en compte la surface collectrice, le nombre de télescopes élémentaires ainsi que la fréquence spatiale maximale d'intérêt c'est-à-dire la résolution instrument, donnée par la mission. Elle est décrite dans [A41, Annexe M page 329]. Je l'ai étendue à un interféromètre 1D en rotation sur lui-même pour synthétiser une ouverture 2D [C57].

Elle a été appliquée à l'étude intitulée SOTISE d'un satellite d'observation de la Terre à haute résolution en orbite géostationnaire, étude dont j'ai été le chef de projet, et dont le lecteur intéressé trouvera une synthèse dans [C51] ou dans [C45].

Chapitre 5

Perspectives

Les perspectives de mes travaux sont organisées ci-dessous non pas par thématique (étalonnage d'instrument, c'est-à-dire ASO et cophasage) d'un côté et imagerie (restauration et reconstruction d'images de l'autre) mais par application. En effet, c'est souvent le besoin applicatif qui permet de guider les développements des traitements et leur donne tout leur sens.

Astronomie

Imagerie à haute dynamique et imagerie multi-spectrale

Les exigences draconniennes de l'imagerie à haute dynamique pour la détection et la caractérisation d'exo-planètes, domaine en pleine expansion, font passer les traitements d'un statut de *nice to have* à celui de sous-système essentiel du système, pour SPHERE par exemple et pour EPICS à l'avenir. Ces traitements incluent les deux axes de mes travaux, étalonnage éventuellement temps réel (par analyse de front d'onde post-focale) et imagerie *a posteriori* (restauration des images/détection).

La diversité de phase longue pose en ligne [A15, Annexe G page 257] est particulièrement prometteuse pour des systèmes d'imagerie à haute dynamique comme SPHERE. En effet, en plaçant un tel capteur juste avant le coronographe on mesurerait les aberrations quasi-statiques impactant directement l'efficacité du coronographe. Cette technique permettrait de corriger les aberrations quasi-statiques régulièrement pendant la nuit au lieu de le faire de jour seulement, donc potentiellement de gagner en détectivité c'est-à-dire de pouvoir détecter des exoplanètes plus faibles. Dans les années qui viennent je souhaite contribuer à ce qu'une telle technique soit mise en œuvre sur SPHERE. Pour le futur système EPICS de l'ELT européen, un intérêt supplémentaire de la diversité de phase longue pose en ligne serait d'assurer également et simultanément le cophasage fin des segments de ce télescope géant, *cf.* [C24].

L'extension « ultime » de la diversité de phase serait d'utiliser directement le plan focal scientifique, situé après le coronographe, comme senseur : ceci permettrait à la fois d'éviter toute aberration différentielle ou NCPA et d'éviter toute introduction de capteur supplémentaire. Dans

ce but, j'ai proposé d'étendre la diversité de phase à l'imagerie coronographique grâce au modèle d'imagerie coronographique que Jean-François Sauvage et moi avons développé [A5]. Les premiers résultats de ce nouvel analyseur de front d'onde sont très encourageants : en simulation nous avons validé l'estimation simultanée des aberrations en amont et en aval d'un coronographe parfait avec une précision nanométrique [59]. La validation expérimentale de cet analyseur sera réalisée lors de la thèse de Baptiste Paul, qui commence à la rentrée 2011.

La reconstruction d'image et la détection sont, tout particulièrement pour l'imagerie à haute dynamique, en lien étroit avec l'étalonnage de l'instrument. En effet, la limité de détectivité d'un tel système est donnée par les résidus d'étalonnage. Par conséquent, une reconstruction/détection myope c'est-à-dire estimant simultanément les résidus d'étalonnage (aberrations quasi-statiques donnant les *speckles*) devrait permettre d'améliorer significativement la détectivité de ces systèmes.

Par ailleurs, les spectro-imageurs ou IFSs sont aujourd'hui en passe de devenir des instruments aussi répandus sur les télescopes de 8-10 m que les imageurs simples dans la décennie précédente, et les traitements pour ces instruments sont absolument nécessaires du fait du grand volume de données que ceux-ci produisent. De plus, la richesse des données produites par un tel instrument permet d'envisager avec confiance le type de reconstruction myope évoqué plus haut.

Les instruments de type IFS pour l'imagerie à haute dynamique représentent donc une perspective de développement importante pour les traitements dans les années à venir. La thèse de Marie Ygouf, que je coencadre en collaboration étroite avec l'IPAG, a commencé fin 2009 sur ce sujet.

Les ELTs

La conception des ELTs, en particulier de l'ELT européen (E-ELT), est un défi instrumental majeur et est donc naturellement le cadre de nombreuses futures études. L'étalonnage d'un tel monstre sera, à n'en pas douter, source d'études originales et d'innovations. Dès à présent, deux besoins importants pourraient trouver une réponse appropriée grâce à l'analyse de front d'onde au voisinage du plan focal : le cophasage des segments de l'E-ELT et la mesure des premiers modes de la turbulence sur étoile naturelle [A4].

Imagerie grand champ

L'estimation de réponse impulsionnelle pour le traitement des images reste un domaine très incomplètement abordé sur les télescopes et instruments existants. Cette estimation deviendra essentielle pour l'interprétation des données dès que seront sur le ciel des systèmes grand champ donc ayant une réponse variable dans le champ.

Deux thèmes de la restauration d'images petit champ (c'est-à-dire isoplanétique) méritent à mon avis d'être à la fois poursuivis et étendus à l'imagerie grand champ : d'une part la déconvolution myope, d'autre part l'estimation des hyper-paramètres, ces deux aspects pouvant être

traités simultanément dans des approches marginalisées [A29] ou fondées sur l'estimation de l'espérance *a posteriori* [55, 56].

Imagerie de la rétine et bio-médicale

L'imagerie de la rétine connaît actuellement un développement sans précédent pour répondre à des besoins croissant très rapidement, en particulier du fait du vieillissement de la population dans les pays développés.

C'est un champ d'application passionnant des techniques de traitements et en particulier de la restauration des images. Une spécificité majeure de l'imagerie de la rétine, et plus généralement bio-médicale, est son caractère tri-dimensionnel. Plus précisément, la formation des images est 3D et un besoin pour les utilisateurs est de pouvoir obtenir soit une image 3D soit une image 2D résolue longitudinalement, c'est-à-dire possédant du « sectionnement optique ». La validation expérimentale de la déconvolution 3D sur des images de rétine, non myope d'abord puis myope ultérieurement [A18, Annexe H page 269], permettra de démontrer la capacité des traitements à effectuer un sectionnement optique pour cette application difficile. La déconvolution myope 3D, ou diversité de phase 3D, pourrait également trouver une application pertinente en microscopie pour l'imagerie bio-médicale hors ophtalmologie. Comme la diversité de phase classique, la diversité de phase 3D peut être envisagée d'une part comme ASO pour corriger en temps réel les aberrations des microscopes lorsqu'elles sont importantes (optique active), d'autre part comme une technique de mesure de RI et de déconvolution lorsque les aberrations sont modestes.

Sans attendre l'avènement de systèmes couplant l'OA et l'OCT plein champ, il me semble également intéressant d'envisager une technique hybride consistant à faire de l'imagerie corrigée par OA mais en modifiant le système d'imagerie afin de coder la profondeur dans la réponse impulsionnelle ; ce peut être par exemple réalisé en remettant au goût du jour une technique interférométrique en lumière spatialement incohérente comme l'holographie conoscopique, technique que j'ai modifiée pendant ma thèse pour en faciliter l'inversion [A44, A43] et pour laquelle j'ai obtenu la première reconstruction 3D [A42]. En effet, l'insertion d'un système conoscopique devant la caméra d'imagerie aboutit essentiellement à remplacer la RI des plans défocalisés par un système de franges circulaires à large spectre spatial, et nous avons pu vérifier par simulations que cette amélioration du contenu spectral 2D conduisait à une meilleure séparation des plans, ou résolution longitudinale, lors de la déconvolution 3D.

Une perspective à plus long terme, lorsque le couplage OA + OCT plein champ sera opérationnel, serait d'appliquer la déconvolution 3D aux images issues d'un tel système et ainsi d'en améliorer encore la résolution, dans les trois dimensions.

Autres applications

Imagerie depuis satellite

L'augmentation de la résolution des télescopes à bord de satellites passe par l'interférométrie et nécessitera le cophasage des éléments de ces instruments, que ceux-la soient des pétales d'un miroir primaire comme pour le JWST ou des télescopes *free-flyer* comme pour les missions DARWIN/TPF-I.

Les deux projets de l'équipe HRA que j'ai dirigés sur l'imagerie haute résolution par télescope multi-pupilles depuis une orbite géostationnaire ont permis d'identifier et d'étudier quelques points clés, notamment le cophasage. Des études industrielles ont poursuivi les nôtres et je compte maintenant, avec mes collègues concernés, mener des études en collaboration avec des industriels afin de continuer à lever les verrous technologiques des futurs télescopes spatiaux.

Systèmes lasers

Pour le contrôle et la focalisation de lasers de puissance, comme pour les télécoms laser (précompensation des effets de la turbulence), il est important de mesurer et de contrôler non seulement la phase mais également l'amplitude de l'onde. La diversité de phase peut être étendue en ce sens et être un analyseur pertinent pour ce type d'applications.

Enfin, une application émergente des techniques de cophasage est le cophasage de fibres optiques, et en particulier de lasers fibrés. Nous avons pu montrer, lors de travaux de thèse [T5], que la diversité de phase était une technique très prometteuse pour cophaser un grand nombre de fibres monomodes [B3] avec un seul capteur. Cette capacité pourrait être déterminante pour réaliser les lasers de puissance nécessaires pour des projets tel que l'*Extreme Light Infrastructure* ou ELI destinés à la physique de l'extrême.

Bibliographie

[1] P. Léna, D. Rouan, F. Lebrun, F. Mignard et D. Pelat, *L'observation en astrophysique*, EDP Sciences, Les Ulis, France (2008).

[2] J. Idier, rédacteur en chef, *Bayesian Approach to Inverse Problems*, Digital Signal and Image Processing Series. ISTE / John Wiley, London (2008).

[3] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill (1968).

[4] A. Maréchal et M. Françon, *Diffraction, Structure des images*, Masson, Paris (1970).

[5] M. Born et E. Wolf, *Principles of Optics*, Pergamon Press, Sixth (corrected) édition (1993).

[6] J.-M. Mariotti, *Introduction to Fourier Optics and Coherence*, Dans *Diffraction-limited imaging with very large telescopes*, sous la direction de D. M. Alloin et J.-M. Mariotti, vol. C-274 de *NATO ASI Series*, pp. 3–31. Kluwer Academic Publishers, Dordrecht (1989).

[7] F. Cassaing, *Analyse d'un instrument à synthèse d'ouverture optique : méthodes de cophasage et imagerie à haute résolution angulaire*, Thèse de doctorat, Université Paris XI Orsay (décembre 1997).

[8] J. W. Goodman, *Statistical optics*, John Wiley & Sons, New York (1985).

[9] F. Roddier, *The effects of atmospherical turbulence in optical astronomy*, Dans *Progress in Optics*, sous la direction de E. Wolf, vol. XIX, pp. 281–376. North Holland, Amsterdam (1981).

[10] R. J. Noll, *Zernike polynomials and atmospheric turbulence*, J. Opt. Soc. Am., 66 (3), pp. 207–211 (1976).

[11] D. L. Fried, *Statistics of a Geometric Representation of Wavefront Distortion*, J. Opt. Soc. Am., 55 (11), pp. 1427–1435 (1965).

[12] F. Roddier, J. M. Gilli et G. Lund, *On the origin of speckle boiling and its effects in stellar speckle interferometry*, J. of Optics (Paris), 13 (5), pp. 263–271 (1982).

[13] J.-M. Conan, G. Rousset et P.-Y. Madec, *Wave-front temporal spectra in high-resolution imaging through turbulence*, J. Opt. Soc. Am. A, 12 (12), pp. 1559–1570 (juillet 1995).

[14] A. Labeyrie, *Attainment of diffraction-limited resolution in large telescopes by Fourier analysing speckle patterns*, Astron. Astrophys., 6, pp. 85–87 (1970).

[15] K. T. Knox et B. J. Thompson, *Recovery of images from atmospherically degraded short exposure photographs*, Astrophys. J. Lett., 193, pp. L45–L48 (1974).

[16] G. Weigelt, *Modified astronomical speckle interferometry "speckle masking"*, Opt. Commun., 21 (1), pp. 55–59 (1977).

[17] J.-C. Fontanella, *Analyse de surface d'onde, déconvolution et optique active*, J. of Optics (Paris), 16 (6), pp. 257–268 (1985).

[18] J. Primot, G. Rousset et J.-C. Fontanella, *Deconvolution from wavefront sensing: a new technique for compensating turbulence-degraded images*, J. Opt. Soc. Am. A, 7 (9), pp. 1598–1608 (1990).

[19] F. Roddier, *Passive versus active methods in optical interferometry*, Dans *High-resolution imaging by interferometry, part II*, sous la direction de F. Merkle, no. 29 dans ESO Conference and Workshop Proceedings, pp. 565–574, Garching bei München, Germany (juillet 1988).

[20] G. Rousset, J.-C. Fontanella, P. Kern, P. Gigan, F. Rigaut, P. Léna, C. Boyer, P. Jagourel, J.-P. Gaffard et F. Merkle, *First diffraction-limited astronomical images with adaptive optics*, Astron. Astrophys., 230, pp. 29–32 (1990).

[21] F. Roddier, rédacteur en chef, *Adaptive Optics in Astronomy*, Cambridge University Press, Cambridge (1999).

[22] J.-M. Conan, *Étude de la correction partielle en optique adaptative*, Thèse de doctorat, Université Paris XI Orsay (octobre 1994).

[23] J.-M. Conan et G. Rousset, rédacteurs en chef, *Multi-Conjugate Adaptive Optics for Very Large Telescopes Dossier*, Vol. 6 fascicule 10 de *C. R. Physique, Académie des Sciences*. Elsevier, Paris (décembre 2005).

[24] K. Dohlen, J.-L. Beuzit, M. Feldt, D. Mouillet, P. Puget, J. Antichi, A. Baruffolo, P. Baudoz, A. Berton, A. Boccaletti, M. Carbillet, J. Charton, R. Claudi, M. Downing, C. Fabron, P. Feautrier, E. Fedrigo, T. Fusco, J.-L. Gach, R. Gratton, N. Hubin, M. Kasper, M. Langlois, A. Longmore, C. Moutou, C. Petit, J. Pragt, P. Rabou, G. Rousset, M. Saisse, H.-M. Schmid, E. Stadler, D. Stamm, M. Turatto, R. Waters et F. Wildi, *SPHERE: A planet finder instrument for the VLT*, Dans *Ground-based and Airborne Instrumentation for Astronomy*, sous la direction de I. S. McLean et M. Iye, vol. 6269. Proc. Soc. Photo-Opt. Instrum. Eng. (2006).

[25] M. Glanc, *Applications Ophtalmologiques de l'Optique Adaptative*, Thèse de doctorat, Univ. Paris XI (2002).

[26] M. Glanc, E. Gendron, F. Lacombe, D. Lafaille, J.-F. Le Gargasson et P. Léna, *Towards wide-field retinal imaging with adaptive optics*, Opt. Commun., 230, pp. 225–238 (2004).

[27] G. Rousset, *Wave-front sensors*, Dans Roddier [21], chap. 5, pp. 91–130.

[28] R. B. Shack et B. C. Plack, *Production and use of a lenticular Hartmann screen (abstract)*, J. Opt. Soc. Am., 61, p. 656 (1971).

[29] F. Roddier, *Curvature sensing and compensation: a new concept in adaptive optics*, Appl. Opt., 27 (7), pp. 1223–1225 (avril 1988).

[30] R. Ragazzoni, *Pupil plane wavefront sensing with an oscillating prism*, J. Mod. Opt., 43 (2), pp. 289–293 (1996).

[31] S. Esposito, A. Riccardi et al., *First light AO (FLAO) system for LBT: final integration and acceptance test results in Europe*, Dans *Astronomical Telescopes and Instrumentation / Adaptive Optics Systems II*, sous la direction de B. L. Ellerbroek, M. Hart, N. Hubin et P. L. Wizinowich, vol. 7736. Proc. Soc. Photo-Opt. Instrum. Eng. (juillet 2010).

[32] R. A. Gonsalves, *Phase retrieval and diversity in adaptive optics*, Opt. Eng., 21 (5), pp. 829–832 (1982).

[33] E. P. Wallner, *Optimal wave-front correction using slope measurements*, J. Opt. Soc. Am. A, 73 (12), pp. pp 1771–1776 (décembre 1983).

[34] P. A. Bakut, V. E. Kirakosyants, V. A. Loginov, C. J. Solomon et J. C. Dainty, *Optimal wavefront reconstruction from a Shack-Hartmann sensor by use of a Bayesian algorithm*, Opt. Commun., 109, pp. 10–15 (juin 1994).

[35] R. J. Sasiela et J. G. Mooney, *An optical phase reconstructor based on using a multiplier-accumulator approach*, Dans *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. 551, pp. 170–176. Proc. Soc. Photo-Opt. Instrum. Eng. (1985).

[36] C. Petit, J.-M. Conan, C. Kulcsar, H.-F. Raynaud, T. Fusco, J. Montri et D. Rabaud, *Optimal Control for Multi-conjugate Adaptive Optics*, C. R. Physique 6, pp. 1059–1069 (2005).

[37] R. W. Gerchberg et W. O. Saxton, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik, 35, pp. 237–246 (1972).

[38] R. A. Gonsalves, *Phase retrieval from modulus data*, J. Opt. Soc. Am., 66 (9), pp. 961–964 (1976).

[39] R. G. Paxman, T. J. Schulz et J. R. Fienup, *Joint estimation of object and aberrations by using phase diversity*, J. Opt. Soc. Am. A, 9 (7), pp. 1072–1085 (1992).

[40] M. G. Löfdahl, A. L. Duncan et G. B. Scharmer, *Fast phase diversity wavefront sensor for mirror control*, Dans *Adaptive Optical System Technologies*, sous la direction de D. Bonaccini et R. K. Tyson, vol. 3353, pp. 952–963. Soc. Photo-Opt. Instrum. Eng. (mars 1998).

[41] C. R. Vogel, T. Chan et R. Plemmons, *Fast algorithms for phase-diversity-based blind deconvolution*, Dans *Adaptive Optical System Technologies*, sous la direction de D. Bonaccini et R. K. Tyson, vol. 3353, pp. 994–1005. Soc. Photo-Opt. Instrum. Eng. (mars 1998).

[42] R. L. Kendrick, D. S. Acton et A. L. Duncan, *Phase-diversity wave-front sensor for imaging systems*, Appl. Opt., 33 (27), pp. 6533–6546 (septembre 1994).

[43] G. B. Scharmer, *Object-independent fast phase-diversity*, Dans *High resolution solar physics: Theory, observations and techniques*, sous la direction de T. R. Rimmele, K. S. Balasubramaniam et R. R. Radick, vol. 183 de *Astron. Soc. Pacific Conf. Series*, pp. 330–341 (1999).

[44] M. G. Löfdahl et G. B. Scharmer, *A predictor approach to closed-loop phase-diversity wavefront sensing*, Dans *UV, optical and IR space telescopes and instruments*, sous la direction de J. B. Breckinridge et P. Jakobsen, vol. 4013, pp. 737–748. Soc. Photo-Opt. Instrum. Eng. (2000).

[45] B. Fleury, A. Bonnefois, L. Mugnier et N. Vedrenne, *INCA passif TC3, Avenant N°2 : Travaux complémentaires de dépouillement et d'évaluation*, Rapport Technique de Synthèse ONERA n^o 1/14713 DOTA (décembre 2009).

[46] R. G. Paxman et J. R. Fienup, *Optical misalignment sensing and image reconstruction using phase diversity*, J. Opt. Soc. Am. A, 5 (6), pp. 914–923 (1988).

[47] M. G. Löfdahl, R. L. Kendrick, A. Harwit, K. E. Mitchell, A. L. Duncan, J. H. Seldin, R. G. Paxman et D. S. Acton, *Phase diversity experiment to measure piston misalignement on the segmented primary mirror of the Keck II Telescope*, Dans *Space Telescopes and Instruments V*, sous la direction de P. Y. Bely et J. B. Breckinridge, vol. 3356, pp. 1190–1201. Soc. Photo-Opt. Instrum. Eng. (1998).

[48] *James Webb Space telescope*, http://www.jwst.nasa.gov/.

[49] D. Redding et al., *Wavefront sensing and control for a Next Generation Space Telescope*, Dans *Space Telescopes and Instruments V*, sous la direction de P. Y. Bely et J. B. Breckinridge, vol. 3356 (2), pp. 758–772. Soc. Photo-Opt. Instrum. Eng. (1998).

[50] D. A. Carrara, B. J. Thelen et R. G. Paxman, *Aberration correction of segmented-aperture telescopes by using phase diversity*, Dans *Image reconstruction from incomplete data*, sous la direction de M. A. Fiddy et R. P. Millane, vol. 4123, pp. 56–63. Soc. Photo-Opt. Instrum. Eng. (2000).

[51] L. H. Lee, G. Vasudevan et E. H. Smith, *Point-by-point approach to phase-diverse phase retrieval*, Dans *IR space telescopes and Instruments*, sous la direction de J. C. Mather, vol. 4850, pp. 441–452. Soc. Photo-Opt. Instrum. Eng. (2003).

[52] R. L. Kendrick, D. S. Acton et A. L. Duncan, *Experimental results from the Lockheed phase diversity test facility*, Dans *Image reconstruction and restoration*, sous la direction de T. J. Schulz et D. L. Snyder, vol. 2302, pp. 312–322. Soc. Photo-Opt. Instrum. Eng. (1994).

[53] V. Zarifis et al., *The Multi Aperture Imaging Array*, Dans *Working on the Fringe: optical and IR interferometry from ground and space*, sous la direction de S. Unwin et R. Stachnik, vol. 194 de *Astron. Soc. Pacific Conf. Series*, pp. 278–285, Dana Point (mai 1999).

[54] J. H. Seldin, R. G. Paxman, V. G. Zarifis, L. Benson et R. E. Stone, *Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity*, Dans *Imaging technology and telescopes*, no. 4091 dans Proc. Soc. Photo-Opt. Instrum. Eng. (juillet 2000).

[55] F. Orieux, J.-F. Giovannelli et T. Rodet, *Bayesian estimation of regularization and PSF parameters for Wiener-Hunt deconvolution*, J. Opt. Soc. Am. A, 27 (7), pp. 1593–1607 (juillet 2010).

[56] J.-F. Giovannelli, *Unsupervised Bayesian convex deconvolution based on a field with an explicit partition function*, IEEE Trans. Image Processing, 17 (1), pp. 16–23 (janvier 2008).

[57] T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper et P. Puget, *High order Adaptive Optics requirements for direct detection of Extra-solar planets. Application to the SPHERE instrument.*, Opt. Express, 14 (17), pp. 7515–7534 (2006).

[58] N. Thatte, R. Abuter, M. Tecza, E. L. Nielsen, F. J. Clarke et L. M. Close, *Very high contrast integral field spectroscopy of AB Doradus C: 9 mag contrast at 0.2" without a coronagraph using spectral deconvolution*, Mon. Not. R. Astr. Soc. (2007).

[59] L. Mugnier, J.-F. Sauvage, R. Villecroze et T. Fusco, *Coronagraphic phase diversity: a simple focal plane sensor for high contrast imaging*, Rapport Technique ONERA n^o 1/14127 DOTA (avril 2011), 7th Framework Programme of the EU.

[60] E. Thiébaut, P. J. V. Garcia et R. Foy, *Imaging with Amber/VLTI: the case of microjets*, Astrophys. Space. Sci., 286, pp. 171–176 (2003).

[61] L. Mugnier et S. Meimon, WISARD *software documentation*, Rapport Technique ONERA (2007), European Interferometry Initiative, Joint Research Action 4, 6th Framework Programme of the EU.

[62] B. Mennesson, *Interférométrie stellaire dans l'infrarouge thermique : observations d'environnements circumstellaires par optique guidée monomode et contributions à la mission spatiale DARWIN*, Thèse de doctorat, Univ. Paris VII, Paris (décembre 1999).

Troisième partie

Liste des publications

Contributions à des ouvrages de synthèse

NB : deux contributions récentes dont je suis premier auteur, [L5] et [L6], sont reproduites dans les annexes A et B, page 115 et suivantes.

[L1] L. Blanc-Féraud, L. Mugnier et A. Jalobeanu, *Blind Image Deconvolution*, Dans *Inverse Problems in Vision and 3D Tomography*, sous la direction de A. Mohammad-Djafari, chap. 3, pp. 97–121. ISTE / John Wiley, London (2010).

[L2] L. Blanc-Féraud, L. Mugnier et A. Jalobeanu, *Déconvolution aveugle d'image*, Dans *Problèmes inverses en imagerie et en vision*, sous la direction de A. Mohammad-Djafari, chap. 3, pp. 107–132. Hermes, Paris (2009).

[L3] L. Mugnier, *Des données à la connaissance de l'objet : le problème inverse*, Dans *L'observation en astrophysique*, sous la direction de P. Léna, D. Rouan, F. Lebrun, F. Mignard et D. Pelat, chap. 9, section 6, pp. 591–613. EDP Sciences, Les Ulis, France (2008).

[L4] L. M. Mugnier, G. Le Besnerais et S. Meimon, *Inversion in Optical Imaging through Atmospheric Turbulence*, Dans *Bayesian Approach to Inverse Problems*, sous la direction de J. Idier, Digital Signal and Image Processing Series, chap. 10, pp. 243–283. ISTE / John Wiley, London (2008).

[L5] L. M. Mugnier, E. Thiébaut et A. Belu, *Data processing in nulling interferometry: case of the Darwin mission*, Dans *Astronomy with High Contrast Imaging III*, sous la direction de C. Aime, M. Carbillet et A. Ferrari, EAS Publications Series, pp. 69–84. EDP Sciences, Les Ulis, France (2006).

[L6] L. M. Mugnier, A. Blanc et J. Idier, *Phase Diversity: a Technique for Wave-Front Sensing and for Diffraction-Limited Imaging*, Dans *Advances in Imaging and Electron Physics*, sous la direction de P. Hawkes, vol. 141, chap. 1, pp. 1–76. Elsevier (2006).

[L7] L. Mugnier et G. Le Besnerais, *Problèmes inverses en imagerie optique à travers la turbulence*, Dans *Approche bayésienne pour les problèmes inverses*, sous la direction de J. Idier, chap. 10, pp. 241–270. Hermes, Paris (2001).

Brevets

[B1] F. Cassaing, I. Mocoeur et L. Mugnier, *Procédé d'estimation d'au moins une déformation du front d'onde d'un système optique ou d'un objet observé par le système optique et dispositif associé*, Brevet Onera WO 2009/010493 (Déposé le 19 juillet 2007), Numéro de soumission 1000014362.

[B2] F. Cassaing et L. Mugnier, *Procédé et dispositif de mesure d'au moins une déformation d'une surface d'onde*, Brevet Onera FR2916045 (A1) (Déposé le 11 mai 2007), Numéro de soumission 1000010994.

[B3] S. Demoustier, A. Brignon, J.-P. Huignard, L. Mugnier et J. Primot, *Source laser à recombinaison cohérente de faisceaux*, Brevet Thales (Déposé le 12 août 2005), Numéro d'enregistrement national 05 08542.

Thèse de Doctorat et thèses (co)encadrées

[T1] A. Cornia, *High-contrast differential image processing for extrasolar planet detection*, Thèse de doctorat, École Doctorale d'Astronomie et d'Astrophysique d'Île de France (décembre 2010).

[T2] G. Chenegros, *Restauration d'images de la rétine corrigées par optique adaptative*, Thèse de doctorat, Université Paris VII (novembre 2008).

[T3] I. Mocoeur, *Analyse de front d'onde en plan focal: développement d'algorithmes temps-réel et application au cophasage de télescopes multipupilles imageurs*, Thèse de doctorat, Université Paris XI (juin 2008).

[T4] J.-F. Sauvage, *Calibration et méthodes d'inversion en imagerie haute dynamique pour la détection directe d'exoplanétes.*, Thèse de doctorat, Université Paris VII (décembre 2007).

[T5] S. Demoustier, *Recombinaison cohérente de fibres laser*, Thèse de doctorat, Université Paris XI, Orsay (décembre 2006).

[T6] S. Meimon, *Reconstruction d'images astronomiques en interférométrie optique*, Thèse de doctorat, Université Paris Sud (2005).

[T7] D. Gratadour, *Optique adaptative, traitement d'image et étude des noyaux actifs de galaxie*, Thèse de doctorat, Ecole Doctorale d'Astronomie et d'Astrophysique d'Ile de France (2005).

[T8] F. Baron, *Définition et test d'un capteur de cophasage sur télescope multipupilles: application à la détection d'exoplanétes et à l'observation de la Terre*, Thèse de doctorat, Ecole Doctorale d'Astronomie et d'Astrophysique d'Ile de France (2005).

[T9] B. Le Roux, *Commande optimale en optique adaptative classique et conjuguée*, Thèse de doctorat, Université de Nice-Sophia Antipolis (octobre 2003).

[T10] A. Blanc, *Identification de réponse impulsionnelle et restauration d'images : apports de la diversité de phase*, Thèse de doctorat, Université Paris XI Orsay (juillet 2002).

[T11] T. Fusco, *Correction partielle et anisoplanétisme en Optique Adaptative : traitement* a posteriori *et Optique Adaptative Multiconjuguée*, Thèse de doctorat, Université de Nice-Sophia Antipolis, Nice France (octobre 2000).

[T12] L. Meynadier, *Analyse de surface d'onde pour le contréle actif d'un télescope spatial*, Thèse de doctorat, Université de Nice-Sophia Antipolis (1997).

[T13] L. Mugnier, *Vers une inversion des hologrammes conoscopiques*, Thèse de doctorat, École Nationale Supérieure des Télécommunications, Paris (novembre 1992).

Articles de revues à comité de lecture

NB : une sélection d'articles dont je suis premier ou second auteur est reproduite dans les annexes C à M, page 215 et suivantes.

[A1] F. H. Vincent, T. Paumard, G. Perrin, L. Mugnier, F. Eisenhauer et S. Gillessen, *Performance of astrometric detection of a hotspot orbiting on the innermost stable circular orbit of the Galactic Centre black hole*, Mon. Not. R. Astr. Soc. (2011).

[A2] L. Blanco et L. M. Mugnier, *Marginal blind deconvolution of adaptive optics retinal images*, Opt. Express, 19 (23), pp. 23227–23239 (novembre 2011).

[A3] F. Gillard, S. Lefebvre, Y. Ferrec, L. Mugnier, S. Rommeluère, C. Benoit, N. Guérineau et J. Taboury, *Inverse problem approaches for stationary Fourier transform spectrometers*, Opt. Lett., 36 (13), pp. 2444–2446 (juillet 2011).

[A4] S. Meimon, T. Fusco et L. M. Mugnier, *LIFT: a focal-plane wavefront sensor for real-time low-order sensing on faint sources*, Opt. Lett., 35 (18), pp. 3036–3038 (2010).

[A5] J.-F. Sauvage, L. M. Mugnier, G. Rousset et T. Fusco, *Analytical expression of long-exposure adaptive-optics-corrected coronagraphic image. First application to exoplanet detection*, J. Opt. Soc. Am. A, 27 (11), pp. A157–A170 (novembre 2010).

[A6] X. Haubois, G. Perrin, S. Lacour, T. Verhoelst, S. Meimon, L. Mugnier, E. Thiébaut, J.-P. Berger, S. T. Ridgway, J. D. Monnier, R. Millan-Gabet et W. Traub, *Imaging the spotty surface of Betelgeuse in the H band*, Astron. Astrophys., 508 (2), pp. 923–932 (2009).

[A7] S. Meimon, L. M. Mugnier et G. Le Besnerais, *Self-calibration approach for optical long-baseline interferometry imaging*, J. Opt. Soc. Am. A, 26 (1), pp. 108–120 (2009).

[A8] I. Mocœur, L. Mugnier et F. Cassaing, *Estimation des défauts d'alignement d'un instrument multipupille par Diversité de Phase temps-réel*, Traitement du Signal, 26 (1), pp. 67–76 (2009).

[A9] I. Mocœur, L. M. Mugnier et F. Cassaing, *Analytical solution to the phase-diversity problem for real-time wavefront sensing*, Opt. Lett., 34 (22), pp. 3487–3489 (novembre 2009).

[A10] L. M. Mugnier, A. Cornia, J.-F. Sauvage, G. Rousset, T. Fusco et N. Védrenne, *Optimal Method for Exoplanet Detection by Angular Differential Imaging*, J. Opt. Soc. Am. A, 26 (6), pp. 1326–1334 (juin 2009).

[A11] C. S. Cockell, T. Herbst, A. Léger, O. Absil, C. Beichman, W. Benz, A. Brack, B. Chazelas, A. Chelli, H. Cottin, V. Coudé du Foresto, W. Danchi, D. Defrère, J. den Herder, C. Eiroa, M. Fridlund, T. Henning, K. Johnston, L. Kaltenegger, L. Labadie, H. Lammer, R. Launhardt, P. Lawson, O. P. Lay, R. Liseau, S. R. Martin, D. Mawet, D. Mourard, C. Moutou, L. Mugnier, F. Paresce, A. Quirrenbach, Y. Rabbia, H. J. A. Rottgering, D. Rouan, N. Santos, F. Selsis, E. Serabyn, F. Westall, G. White, M. Ollivier et P. Bordé, *Darwin—an experimental astronomy mission to search for extrasolar planets*, Experimental Astronomy, 23, pp. 435–461 (mars 2009).

[A12] C. S. Cockell, A. Léger, M. Fridlund, T. M. Herbst, L. Kaltenegger, O. Absil, C. Beichman, W. Benz, M. Blanc, A. Brack, A. Chelli, L. Colangeli, H. Cottin, F. Coudé du Foresto, W. C. Danchi, D. Defrère, J. den Herder, C. Eiroa, J. Greaves, T. Henning, K. J. Johnston, H. Jones, L. Labadie, H. Lammer, R. Launhardt, P. Lawson, O. P. Lay, J. LeDuigou, R. Liseau, F. Malbet, S. R. Martin, D. Mawet, D. Mourard, C. Moutou, L. M. Mugnier, M. Ollivier, F. Paresce, A. Quirrenbach, Y. D. Rabbia, J. A. Raven, H. J. A. Rottgering, D. Rouan, N. C. Santos, F. Selsis, E. Serabyn, H. Shibai, M. Tamura, E. Thiébaut, F. Westall et G. J. White, *Darwin: A Mission to Detect and Search for Life on Extrasolar Planets*, Astrobiology, 9 (1), pp. 1–22 (mars 2009).

[A13] A.-M. Lagrange, D. Gratadour, G. Chauvin, T. Fusco, D. Ehrenreich, D. Mouillet, G. Rousset, D. Rouan, F. Allard, É. Gendron, J. Charton, L. Mugnier, P. Rabou, J. Montri et F. Lacombe, *A probable giant planet imaged in the β Pictoris disk. VLT/NaCo deep L'-band imaging*, Astron. Astrophys., 493, pp. L21–L25 (janvier 2009).

[A14] S. Lacour, S. Meimon, E. Thiébaut, G. Perrin, T. Verhoelst, E. Pedretti, P. A. Schuller, L. Mugnier, J. Monnier, J. P. Berger, X. Haubois, A. Poncelet, G. Le Besnerais, K. Eriksson, R. Millan-Gabet, S. Ragland, M. Lacasse et W. Traub, *The limb darkened Arcturus: Imaging with the IOTA/IONIC interferometer*, Astron. Astrophys., 485, pp. 561–570 (2008).

[A15] L. M. Mugnier, J.-F. Sauvage, T. Fusco, A. Cornia et S. Dandy, *On-Line Long-Exposure Phase Diversity: a Powerful Tool for Sensing Quasi-Static Aberrations of Extreme Adaptive Optics Imaging Systems*, Opt. Express, 16 (22), pp. 18406–18416 (octobre 2008).

[A16] G. Le Besnerais, S. Lacour, L. M. Mugnier, E. Thiébaut, G. Perrin et S. Meimon, *Advanced Imaging Methods for Long-Baseline Optical Interferometry*, IEEE Journal of Selected Topics in Signal Processing, 2 (5), pp. 767–780 (octobre 2008).

[A17] F. Baron, I. Mocoeur, F. Cassaing et L. M. Mugnier, *Unambiguous phase retrieval as a cophasing sensor for phased array telescopes: derivation of an analytical estimator.*, J. Opt. Soc. Am. A, 25 (5), pp. 1000–1015 (mai 2008).

[A18] G. Chenegros, L. M. Mugnier, F. Lacombe et M. Glanc, *3D Phase Diversity: a Myopic Deconvolution Method for Short-Exposure Images. Application to Retinal Imaging.*, J. Opt. Soc. Am. A, 24 (5), pp. 1349–1357 (mai 2007).

[A19] J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, J. Antici, 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 et F. Wildi, *SPHERE: A 'Planet Finder' Instrument for the VLT*, The Messenger, 125, p. 29 (septembre 2006).

[A20] D. Gratadour, D. Rouan, L. M. Mugnier, T. Fusco, Y. Clénet, E. Gendron et F. Lacombe, *Near-IR AO dissection of the core of NGC 1068 with NaCo*, Astron. Astrophys., 446 (3), pp. 813–825 (février 2006).

[A21] J. Idier, L. Mugnier et A. Blanc, *Statistical behavior of joint least square estimation in the phase diversity context*, IEEE Trans. Image Processing, 14 (12), pp. 2107–2116 (décembre 2005).

[A22] S. Meimon, L. M. Mugnier et G. Le Besnerais, *A convex approximation of the likelihood in optical interferometry*, J. Opt. Soc. Am. A (novembre 2005).

[A23] D. Gratadour, L. M. Mugnier et D. Rouan, *Sub-pixel image registration with a maximum likelihood estimator*, Astron. Astrophys., 443, pp. 357–365 (novembre 2005).

[A24] S. Meimon, L. M. Mugnier et G. Le Besnerais, *Reconstruction method for weak-phase optical interferometry*, Opt. Lett., 30 (14), pp. 1809–1811 (juillet 2005).

[A25] A. D. Storrs, C. Dunne, J.-M. Conan, L. Mugnier, B. P. Weiss et B. Zellner, *A Closer Look at Main Belt Asteroids 1: WF/PC Images*, Icarus, 173 (2), pp. 409–416 (février 2005).

[A26] D. Rouan, F. Lacombe, E. Gendron, D. Gratadour, Y. Clénet, A.-M. Lagrange, D. Mouillet, C. Boisson, G. Rousset, T. Fusco, L. Mugnier, N. Thatte, R. Genzel, P. Gigan, R. Arsenault et P. Kern, *Hot Very Small Dust Grains in NGC 1068 seen in jet induced structures thanks to VLT/NACO adaptive optics*, Astron. Astrophys., 417, pp. L1–L4 (2004).

[A27] L. M. Mugnier, T. Fusco et J.-M. Conan, *MISTRAL: a Myopic Edge-Preserving Image Restoration Method, with Application to Astronomical Adaptive-Optics-Corrected Long-Exposure Images.*, J. Opt. Soc. Am. A, 21 (10), pp. 1841–1854 (octobre 2004).

[A28] B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. M. Mugnier et T. Fusco, *Optimal control law for classical and Multiconjugate Adaptive Optics*, J. Opt. Soc. Am. A, 21 (7) (juillet 2004).

[A29] A. Blanc, L. M. Mugnier et J. Idier, *Marginal estimation of aberrations and image restoration by use of phase diversity*, J. Opt. Soc. Am. A, 20 (6), pp. 1035–1045 (2003).

[A30] M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset et R. Lenzen, *Calibration of NAOS and CONICA static aberrations. Experimental results*, Astron. Astrophys., 399, pp. 385–394 (2003).

[A31] A. Blanc, T. Fusco, M. Hartung, L. M. Mugnier et G. Rousset, *Calibration of NAOS and CONICA static aberrations. Application of the phase diversity technique*, Astron. Astrophys., 399, pp. 373–383 (2003).

[A32] A. Coustenis, E. Gendron, O. Lai, J.-P. Véran, J. Woillez, M. Combes, L. Vapillon, T. Fusco, L. Mugnier et P. Rannou, *Images of Titan at 1.3 and 1.6 micron with adaptive optics at the CFHT*, Icarus, 154, pp. 501–515 (2001).

[A33] T. Fusco, J.-M. Conan, G. Rousset, L. M. Mugnier et V. Michau, *Optimal wavefront reconstruction strategies for Multiconjugate Adaptive Optics*, J. Opt. Soc. Am. A, 18 (10), pp. 2527–2538 (octobre 2001).

[A34] L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau et S. Salem, *Myopic deconvolution from wavefront sensing*, J. Opt. Soc. Am. A, 18, pp. 862–872 (avril 2001).

[A35] G. Rousset, L. M. Mugnier, F. Cassaing et B. Sorrente, *Imaging with Multi-Aperture Optical Telescopes and an application*, C. R. Acad. Sci. Paris, Série IV, tome 2 (1), pp. 17–25 (janvier 2001).

[A36] T. Fusco, J.-M. Conan, L. Mugnier, V. Michau et G. Rousset, *Characterization of adaptive optics point spread function for anisoplanatic imaging. Application to stellar field deconvolution.*, Astron. Astrophys. Suppl. Ser., 142, pp. 149–156 (2000).

[A37] T. Fusco, J.-M. Conan, V. Michau, L. Mugnier et G. Rousset, *Efficient phase estimation for large field of view adaptive optics*, Opt. Lett., 24 (21), pp. 1472–1474 (novembre 1999).

[A38] L. Meynadier, V. Michau, M.-T. Velluet, J.-M. Conan, L. M. Mugnier et G. Rousset, *Noise propagation in wave-front sensing with phase diversity*, Appl. Opt., 38 (23), pp. 4967–4979 (août 1999).

[A39] T. Fusco, J.-P. Véran, J.-M. Conan et L. Mugnier, *Myopic deconvolution method for adaptive optics images of stellar fields*, Astron. Astrophys. Suppl. Ser., 134, pp. 1–10 (janvier 1999).

[A40] J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau et G. Rousset, *Myopic Deconvolution of Adaptive Optics Images by use of Object and Point Spread Function Power Spectra*, Appl. Opt., 37 (21), pp. 4614–4622 (juillet 1998).

[A41] L. M. Mugnier, G. Rousset et F. Cassaing, *Aperture Configuration Optimality Criterion for Phased Arrays of Optical Telescopes*, J. Opt. Soc. Am. A, 13 (12), pp. 2367–2374 (décembre 1996).

[A42] L. M. Mugnier, *Conoscopic holography: towards three-dimensional reconstructions of opaque objects*, Appl. Opt., 34 (8), pp. 1363–1371 (1995).

[A43] L. M. Mugnier, G. Y. Sirat et D. Charlot, *Conoscopic holography: two-dimensional reconstructions*, Opt. Lett., 18 (1), pp. 66–68 (janvier 1993).

[A44] L. M. Mugnier et G. Y. Sirat, *On-axis conoscopic holography without a conjugate image*, Opt. Lett., 17 (4), pp. 294–296 (1992).

Communications dans des conférences, articles grand public

[C1] M. Ygouf, L. M. Mugnier, D. Mouillet, T. Fusco et J.-L. Beuzit, *Restauration myope d'images coronographiques pour la détection d'exoplanètes*, Dans *23ème Colloque sur le Traitement du Signal et des Images* (2011), Date conférence : septembre 2011, Bordeaux (France).

[C2] L. Blanco, L. Mugnier et M. Glanc, *Myopic deconvolution of adaptive optics retina images*, Dans *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XVIII*, sous la direction de J.-A. Conchello, C. J. Cogswell, T. Wilson et T. G. Brown, vol. 7904. Proc. Soc. Photo-Opt. Instrum. Eng. (2011), Date conférence : janvier 2011, San Francisco, CA, USA.

[C3] F. de la Barrière, S. Derelle, G. Druart, N. Guérineau, S. Rommeluère, L. Mugnier, N. Baier, B. Cloix, O. Gravrand, N. Lhermet et G. Destefanis, *Modulation Transfer Function Measurement of Infrared Focal Plane Arrays with Small Fill Factors*, Dans *II-VI Materials for Detectors & sources*, sous la direction de S. Sivananthan et N. K. Dhar. II-VI Workshop, Journal of Electronics Materials (2011).

[C4] F. H. Vincent, T. Paumard, G. Perrin, L. Mugnier, F. Eisenhauer et S. Gillessen, *Astrometric study of the complex environment of Sgr A* in imaging mode with the VLTI/GRAVITY instrument*, Dans *The Galactic Center: a Window to the Nuclear Environment of Disk Galaxies*, sous la direction de M. Morris, D. Q. Wang et F. Yuan, vol. 439, p. 275. Astron. Soc. Pacific Conf. Series (2011), Date conférence : 19-23 octobre 2009, Shanghai, China.

[C5] L. Blanco et L. M. Mugnier, *Blind deconvolution of adaptive optics retinal images*, Dans *Adaptive Optics for Industry and Medicine* (juin 2011), Date conférence : juin 2011, Murcia (Spain).

[C6] L. Blanco, L. M. Mugnier et M. Glanc, *Déconvolution myope d'images de la rétine humaine corrigées par optique adaptative*, Dans *6ème édition des journées "imagerie optique non conventionnelle"* (mars 2011), Date conférence : mars 2011, Paris (France).

[C7] C. Robert, J. Voyez, N. Védrenne et L. Mugnier, *C_n^2 profile from Shack-Hartmann data with CO-SLIDAR data processing*, Dans *"Comprehensive characterization of astronomical sites"*. Sternberg Astronomical Institute, arXiv (janvier 2011), http://adsabs.harvard.edu/abs/2011arXiv1101.3924R.

[C8] M. Ygouf, L. Mugnier, J.-F. Sauvage, T. Fusco, D. Mouillet et J.-L. Beuzit, *Approximate analytical expression for AO-corrected coronagraphic imaging in preparation of exoplanet signal extraction*, Dans *SF2A 2010 : Semaine de l'Astrophysique Francaise*, sous

la direction de S. Boissier, M. Heydari-Malayeri, R. Samadi et D. Valls-Gabaud, Paris, France (2010).

[C9] A. Vigan, C. Moutou, M. Langlois, D. Mouillet, K. Dohlen, A. Boccaletti, M. Carbillet, I. Smith, A. Ferrari, L. Mugnier et C. Thalmann, *Comparison of methods for detection and characterization of exoplanets with SPHERE/IRDIS*, Dans *Ground-based and Airborne Instrumentation for Astronomy III*, sous la direction de I. S. McLean, S. K. Ramsay et H. Takami, vol. 7735. Proc. Soc. Photo-Opt. Instrum. Eng. (2010).

[C10] J.-F. Sauvage, L. Mugnier, R. Villecroze, G. Rousset et T. Fusco, *Analytical expression for long exposure coronagraphic imaging*, Dans *In the Spirit of Bernard Lyot: Direct Detection of Planets and Circumstellar Disks* (2010), Date conférence : octobre 2010. Abstract and presentation: http://lyot2010.lesia.obspm.fr/abstract_160.html.

[C11] L. M. Mugnier, J.-F. Sauvage, T. Fusco, S. Dandy et A. Cornia, *Sensing Quasi-Static Aberrations of Adaptive Optics Systems On-Line with Long-Exposure Phase Diversity*, Dans *Adaptive Optics for Extremely Large Telescopes (AO4ELT) 2009*, Les Ulis, France, EDP Sciences (2010), Date conférence : 22-26juin 2009, Paris.

[C12] S. Dandy, J.-F. Sauvage, T. Fusco et L. Mugnier, *Optimized phase diversity sensor for wideband analysis on long-exposure AO corrected images: theory, simulations, and experimental validations*, Dans *Adaptive Optics Systems II*, sous la direction de B. L. Ellerbroek, M. Hart, N. Hubin et P. L. Wizinowich, vol. 7736. Proc. Soc. Photo-Opt. Instrum. Eng. (2010).

[C13] A. Cornia, L. M. Mugnier, M. Carbillet, J.-F. Sauvage, A. Boccaletti, N. Védrenne, D. Mouillet, G. Rousset et T. Fusco, *Optimal Method for Exoplanet Detection by Spectral and Angular Differential Imaging*, Dans *Adaptive Optics for Extremely Large Telescopes (AO4ELT) 2009*, Les Ulis, France, EDP Sciences (2010), Date conférence : 22-26juin 2009, Paris.

[C14] A. Cornia, L. M. Mugnier, D. Mouillet, A. Vigan, A. Eggenberger, G. Rousset, A. Boccaletti, M. Carbillet, K. Dohlen, T. Fusco, J. Carson et G. Montagnier, *Optimal method for exoplanet detection by spectral and angular differential imaging*, Dans *Adaptive Optics Systems II*, sous la direction de B. L. Ellerbroek, M. Hart, N. Hubin et P. L. Wizinowich, vol. 7736, p. 7736 1E. Proc. Soc. Photo-Opt. Instrum. Eng. (2010).

[C15] A. Cornia, L. M. Mugnier, J.-F. Sauvage, N. Védrenne, M. Carbillet, A. Boccaletti, G. Rousset et T. Fusco, *Méthode de type Maximum de Vraisemblance pour la détection d'exoplanêtes par Imagerie Différentielle spectrale et angulaire*, Dans *22ème Colloque sur le Traitement du Signal et des Images*. GRETSI (2009), Date conférence : septembre 8-11, Dijon (France).

[C16] G. Chenegros, L. M. Mugnier, C. Alhenc-Gelas, M. Glanc, F. Lacombe et M. Nicolas, *Unsupervised 3D deconvolution method for retinal imaging: principle and preliminary validation on experimental data*, Dans *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XVI*, sous la direction de J.-A. Conchello, C. J. Cogswell et T. Wilson, vol. 7184. Proc. Soc. Photo-Opt. Instrum. Eng. (2009), Date conférence : janvier 2009, San Jose, CA, USA.

[C17] M. Blavier, L. Blanco, M. Glanc, F. Pouplard, S. Tick, I. Maksimovic, L. Mugnier, G. Chenegros, G. Rousset, F. Lacombe, M. Pâques, J.-F. L. Gargasson et J.-A. Sahel, *Adding the third dimension on adaptive optics retina imager thanks to full-field optical coherence tomography*, Dans *Ophthalmic Technologies XIX*, sous la direction de F. Manns, P. G. Söderberg et A. Ho, vol. 7163. Proc. Soc. Photo-Opt. Instrum. Eng. (2009), Date conférence : janvier 2009, San Jose, CA, USA.

[C18] L. Mugnier, G. Chenegros, M. Glanc, M. Nicolas et F. Lacombe, *Restauration d'images 3D en vue de l'imagerie de la rétine*, Dans *Journées Recherche Industrie Optique Adaptative* (2008).

[C19] L. M. Mugnier, A. Cornia, J.-F. Sauvage, N. Védrenne, T. Fusco et G. Rousset, *Maximum likelihood-based method for angular differential imaging*, Dans *Adaptive Optics Systems*, sous la direction de N. Hubin, C. E. Max et P. L. Wizinowich, vol. 7015. Proc. Soc. Photo-Opt. Instrum. Eng. (2008), Date conférence : juin 2008, Marseille.

[C20] E. Delavaquerie, F. Cassaing, L. Mugnier, I. Mocoeur, F. Baron et V. Michau, *Mesures en plan focal pour l'analyse de front d'onde ou l'imagerie.*, Dans *Journées Recherche Industrie Optique Adaptative* (2008).

[C21] E. Delavaquerie, S. Meimon, F. Cassaing, I. Mocœur, T. Fusco, L. Mugnier et V. Michau, *Cophasing segmented mirrors using a phase diversity algorithm: preliminary results*, Dans *Optical Complex Systems* (2008).

[C22] L. Blanco, M. Blavier, M. Glanc, F. Pouplard, S. Tick, I. Maksimovic, G. Chenegros, L. Mugnier, F. Lacombe, G. Rousset, M. Pâques, J.-F. L. Gargasson et J.-A. Sahel, *First steps toward 3D high resolution imaging using adaptive optics and full-field optical coherence tomography*, Dans *1st Canterbury Workshop on Optical Coherence Tomography and Adaptive Optics*, sous la direction de A. Podoleanu, vol. 7139. Proc. Soc. Photo-Opt. Instrum. Eng. (2008), Date conférence : 2008, Canterbury, UK.

[C23] A. Cornia, L. Mugnier, J.-F. Sauvage, T. Fusco, G. Rousset et N. Védrenne, *Maximum likelihood-based method for angular differential imaging*, Dans *SF2A-2008: Semaine de l'Astrophysique Francaise*, sous la direction de C. Charbonnel, F. Combes et R. Samadi, pp. 65–68, Paris, France (juillet 2008).

[C24] S. Meimon, E. Delavaquerie, F. Cassaing, T. Fusco, L. M. Mugnier et V. Michau, *Phasing segmented telescopes with long-exposure phase diversity images*, Dans *Ground-based and Airborne Telescopes II*, sous la direction de L. M. Stepp et R. Gilmozzi, vol. 7012 de *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* (août 2008).

[C25] C. Robert, T. Fusco, J.-F. Sauvage et L. Mugnier, *Improvement of phase diversity algorithm for non-common path calibration in extreme AO context*, Dans *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7015 de *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* (juillet 2008).

[C26] J.-F. Sauvage, L. Mugnier, A. Woelfflé, T. Fusco et G. Rousset, *Multi-Channel Algorithm for Exoplanets Detection by Angular Differential Imaging*, Dans *Semaine de l'astrophysique française*, sous la direction de J. Bouvier, A. Chalabaev et C. Charbonnel. SF2A, EDP Sciences (2007), Date conférence : juillet 2–6, 2007, Grenoble (France).

[C27] L. M. Mugnier, J.-F. Sauvage, A. Woelffle, T. Fusco et G. Rousset, *Algorithme multi-canaux pour la détection d'exo-planètes en imagerie différentielle angulaire*, Dans *21ième Colloque sur le Traitement du Signal et des Images*. GRETSI (2007), Date conférence : septembre 11-14, 2007, Troyes (France).

[C28] L. M. Mugnier, J.-F. Sauvage, T. Fusco et G. Rousset, *Multi-Channel Planet Detection Algorithm for Angular Differential Imaging*, Dans *Adaptive Optics: Analysis and Methods*. OSA (2007), Date conférence : juin 18–20, 2007, Vancouver (Canada).

[C29] I. Mocœur, L. M. Mugnier et F. Cassaing, *Amélioration des performances d'imagerie d'un télescope multi pupilles par Diversité de Phase temps-réel*, Dans *21ième Colloque sur le Traitement du Signal et des Images*. GRETSI (2007), Date conférence : Sept 11-14, 2007, Troyes (France).

[C30] M. Glanc, L. Blanco, L. Vabre, F. Lacombe, P. Puget, G. Rousset, G. Chenegros, L. Mugnier, M. Pâques, J.-F. Le Gargasson et A. J. Sahel, *First Adaptive Optics images with the upgraded Quinze-Vingts Hospital retinal imager*, Dans *Adaptive Optics: Analysis and Methods*. OSA (2007), Date conférence : juin 18–20, 2007, Vancouver (Canada).

[C31] G. Chenegros, L. Mugnier et F. Lacombe, *3D Myopic Deconvolution Method for Adaptive Optics Corrected Retinal Images*, Dans *Semaine de l'astrophysique française*, sous la direction de J. Bouvier, A. Chalabaev et C. Charbonnel. SF2A, EDP Sciences (2007), Date conférence : juillet 2–6, 2007, Grenoble (France).

[C32] G. Chenegros, L. M. Mugnier, F. Lacombe et M. Glanc, *Phase estimation and Retinal Image Restoration by 3D Phase Diversity.*, Dans *Adaptive Optics: Analysis and Methods*. OSA (2007), Date conférence : juin 18–20, 2007, Vancouver (Canada).

[C33] F. Cassaing, L. M. Mugnier, I. Mocoeur, B. Fleury et V. Michau, *ASO focaux et télescopes multipupilles : un gain en compacité et/ou en complexité*, Dans *Journées Recherche Industrie de l'Optique Adaptative*. AFOP & SFO (2007), Date conférence : novembre 21–22, 2007, Arcachon (France).

[C34] G. Chenegros, L. Mugnier et F. Lacombe, *Restauration myope d'images 3D par diversité de phase*, Dans *21ième Colloque sur le Traitement du Signal et des Images*, pp. 21–23. GRETSI (2007), Date conférence : septembre 11–14, 2007, Troyes (France).

[C35] C. Verinaud, M. Kasper, J.-L. Beuzit, N. Yaitskova, V. Korkiakoski, K. Dohlen, P. Baudoz, T. Fusco, L. Mugnier et N. Thatte, *EPICS Performance Evaluation through Analytical and Numerical Modeling*, Dans *Proceedings of the conference In the Spirit of Bernard Lyot: The Direct Detection of Planets and Circumstellar Disks in the 21st Century. June 04 - 08, 2007. University of California, Berkeley, CA, USA. Edited by Paul Kalas.*, sous la direction de P. Kalas (juin 2007), Date conférence : juin 2007,.

[C36] E. Thiébaut et L. Mugnier, *Maximum a posteriori planet detection and characterization with a nulling interferometer*, Dans *IAUC 200, Direct Imaging of Exoplanets: Science & Techniques* (2006), Date conférence : octobre 2005, Nice, France.

[C37] J.-F. Sauvage, L. Mugnier, T. Fusco et G. Rousset, *Post processing of differential images for direct extrasolar planet detection from the ground*, Dans *Advances in Adaptive Optics II*, sous la direction de L. Ellerbroek B. et D. Bonaccini Calia, vol. 6272. Proc. Soc. Photo-Opt. Instrum. Eng. (2006).

[C38] G. Chenegros, L. M. Mugnier, F. Lacombe et M. Glanc, *3D deconvolution of adaptive-optics corrected retinal images*, Dans *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIII*, sous la direction de J.-A. Conchello, C. J. Cogswell et T. Wilson, vol. 6090. Proc. Soc. Photo-Opt. Instrum. Eng. (2006), Date conférence : janvier 2006, San Jose, CA, USA.

[C39] F. Cassaing, B. Sorrente, L. Mugnier, G. Rousset, V. Michau, I. Mocœur et F. Baron, *BRISE: a multipurpose bench for cophasing sensors*, Dans *Advances in stellar interferometry*, sous la direction de J. D. Monnier et M. Schöller, vol. 6268. Proc. Soc. Photo-Opt. Instrum. Eng. (2006), Date conférence : juin 2006, Orlando, USA.

[C40] F. Cassaing, B. Sorrente, L. Mugnier, G. Rousset, V. Michau, I. Mocœur et F. Baron, *BRISE: a multi-purpose bench for cophasing sensors*, Dans *Visions for InfraRed Astronomy*, sous la direction de V. C. du Foresto, D. Rouan et G. Rousset, vol. 6 [1-4] de *Instrumentation Mesure Métrologie*. Lavoisier (2006), Date conférence : mars 2006, Paris.

[C41] J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, J. Antici, 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 et F. Wildi, *SPHERE: A 'Planet Finder' Instrument for the VLT*, Vol. 125, pp. 29–+ (septembre 2006).

[C42] I. Mocœur, F. Cassaing, F. Baron, L. Mugnier, S. Hofer et H. Thiele, *Darwin fringe sensor: experimental results on the BRISE bench*, Dans *Advances in stellar interferometry*, sous la direction de J. D. Monnier, M. Schöller et W. C. Danchi, vol. 6268. Proc. Soc. Photo-Opt. Instrum. Eng. (juillet 2006), Date conférence : juin 2006, Orlando, USA.

[C43] B. Sorrente, F. Cassaing, I. Mocœur, L. Mugnier, F. Baron et G. Rousset, *BRISE : bench for testing cophasing sensors in multi-telescope interferometry*, Dans *1st International Conference on Optical Complex Systems*, Marseille, France (2005).

[C44] L. Mugnier, F. Cassaing, G. Rousset, F. Baron, V. Michau, I. Mocœur, B. Sorrente et M.-T. Velluet, *Permanent High-Resolution Earth Observation with Multiple Aperture Optical Telescopes*, Dans *Military space: questions in Europe*. AAAF (2005), Date conférence : avril 2005.

[C45] L. Mugnier, F. Cassaing, G. Rousset, F. Baron, V. Michau, I. Mocœur, B. Sorrente et M.-T. Velluet, *Continuous High-Resolution Earth Observation with Multiple Aperture Optical Telescopes*, Dans *Proceedings of the OPTRO 2005 International Symposium*, Paris, France, AAAF (2005), Date conférence : mai 2005.

[C46] I. Mocœur, F. Cassaing, F. Baron, L. M. Mugnier, G. Rousset, B. Sorrente et A. Blanc, *Multi-telescope interferometer cophasing for astronomy*, Dans *Semaine de l'astrophysique française*. SF2A, EDP Sciences (2005).

[C47] S. Meimon, L. M. Mugnier et G. Le Besnerais, *A novel reconstruction method for weak-phase optical interferometry*, Vol. SMB 4. Optical Society of America, OSA (2005), Date conférence : juin 2005, Charlotte, USA.

[C48] S. Demoustier, A. Brignon, E. Lallier, J.-P. Huignard, J.-L. Meyzonnette, J. Primot et L. Mugnier, *Recombinaison cohérente de fibres laser*, Dans *Bulletin POLOQ*, no. 2005-3, Paris, Groupe de Prospective Orientée sur les Lasers et l'OptroniQue, DGA/D4S/MRIS (2005).

[C49] F. Cassaing, F. Baron, B. Fleury, I. Mocœur, L. M. Mugnier, G. Rousset et B. Sorrente, *Multiple-beam fringe tracking for the VLTI*, Dans *The Power of Optical/IR Interferometry: Recent Scientific Results and 2nd Generation VLTI Instrumentation*, sous la direction de A. Chelli et F. Delplancke., ESO Astrophysics Symposia. Springer Verlag (2005).

[C50] J. Idier, L. Mugnier et G. Le Besnerais, *Traitement de données expérimentales et problèmes inverses*, Dans *La Lettre de l'AAAF, num 4*, pp. 10–12. Association Aéronautique & Astronautique de France (avril 2005).

[C51] L. Mugnier, G. Rousset, F. Cassaing, B. Sorrente et M.-T. Velluet, *Observation haute résolution permanente de la Terre par synthèse d'ouverture optique*, Dans *Bulletin POLOQ*, no. 2005-1, pp. 9–16, Paris, Groupe de Prospective Orientée sur les Lasers et l'OptroniQue, DGA/D4S/MRIS (mars 2005).

[C52] S. C. Meimon, E. Thiébaut, L. M. Mugnier et G. Le Besnerais, *Reconstruction d'images en interférométrie stellaire*, Dans *Scientific Highlights 2004*, sous la direction de F. Combes, D.Barret, T. Contini, F. Meynadier et L. Pagani, EDP Sciences, SF2A (2004).

[C53] S. C. Meimon, L. M. Mugnier et G. Le Besnerais, *A novel method of reconstruction for weak-phase optical interferometry*, Dans *New frontiers in stellar interferometry*, sous la direction de W. A. Traub, vol. 5491, pp. 909–919. Proc. Soc. Photo-Opt. Instrum. Eng. (2004), Date conférence : juin 2004, Glasgow, UK.

[C54] P. R. Lawson, W. D. Cotton, C. A. Hummel, J. D. Monnier, M. Zhao, J. S. Young, H. Thorsteinsson, S. C. Meimon, L. Mugnier, G. Le Besnerais, E. Thiébaut et P. G. Tuthill, *An interferometric imaging beauty contest*, Dans *New frontiers in stellar interferometry*, sous la direction de W. A. Traub, vol. 5491, pp. 886–899. Proc. Soc. Photo-Opt. Instrum. Eng. (2004), Date conférence : juin 2004, Glasgow, UK.

[C55] D. Gratadour, L. M. Mugnier et D. Rouan, *Image Centering with a Maximum Likelihood Estimator: Application to Infrared Astronomical imaging*, Dans *ASP Conf. Ser. 314: Astronomical Data Analysis Software and Systems (ADASS) XIII*, pp. 558–+ (2004).

[C56] B. Sorrente, F. Cassaing, F. Baron, C. Coudrain, B. Fleury, F. Mendez, V. Michau, L. Mugnier, G. Rousset, L. Rousset-Rouvière et M.-T. Velluet, *Multiple-Aperture Optical Telescopes: cophasing sensor testbed*, Dans *5th International Conference On Space Optics*, vol. SP-554, pp. 479–484, Toulouse, France, CNES/ESA, ESA (2004).

[C57] L. Mugnier, F. Cassaing, B. Sorrente, F. Baron, M.-T. Velluet, V. Michau et G. Rousset, *Multiple-Aperture Optical Telescopes: some key issues for Earth observation from a GEO orbit*, Dans *5th International Conference On Space Optics*, vol. SP-554, pp. 181–187, Toulouse, France, CNES/ESA, ESA (2004).

[C58] E. Schmidt, F. Cassaing, M. Barillot, F. Baron, S. Hofer, L. M. Mugnier et G. Rousset, *Entwicklung des Wellenfrontsensors DWARF für die DARWIN-Mission der ESA*, Dans *DGaO Jahrestagung*. Deutschen Gesellschaft für Angewandte Optik (2003), Date conférence : juin 2003.

[C59] T. Fusco, L. M. Mugnier, J.-M. Conan, F. Marchis, G. Chauvin, G. Rousset, A.-M. Lagrange et D. Mouillet, *Deconvolution of astronomical adaptive optics images*, Dans *Astronomy with High Contrast Imaging: From Planetary Systems to Active Galactic Nuclei*, sous la direction de C. Aime et R. Soummer, vol. 8 de *European Astronomical Society publication series* (2003), Date conférence : mai 2002.

[C60] E. Schmidt, F. Cassaing, S. Hofer, M. Barillot, F. Baron, L. M. Mugnier, G. Rousset et T. Stuffler, *DARWIN Fringe Sensor (DWARF): Breadboard Development*, Dans *Towards Other Earths*, vol. SP-539, pp. 575–577. ESA (2003), Date conférence : avril 2003.

[C61] F. Cassaing, F. Baron, E. Schmidt, S. Hofer, L. M. Mugnier, M. Barillot, G. Rousset, T. Stuffler et Y. Salvadé, *DARWIN Fringe Sensor (DWARF): Concept Study*, Dans *Towards Other Earths*, vol. SP-539, pp. 389–392. ESA (2003), Date conférence : avril 2003.

[C62] S. Meimon, L. Mugnier et G. Le Besnerais, *Approximations convexes de critères pour la synthèse de Fourier optique*, Dans *19ième Colloque sur le Traitement du Signal et des Images*, sous la direction de J.-M. Chassery et C. Jutten. GRETSI (septembre 2003).

[C63] F. Baron, F. Cassaing et L. Mugnier, *Alignement des pupilles d'un télescope multi-pupilles*, Dans *19ième Colloque sur le Traitement du Signal et des Images*, sous la direction de J.-M. Chassery et C. Jutten. GRETSI (septembre 2003).

[C64] B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. M. Mugnier et T. Fusco, *Optimal control law for multiconjugate adaptive optics*, Dans *Adaptive Optical System Technology II*, sous la direction de P. L. Wizinowich et D. Bonaccini, vol. 4839, Hawaii, USA, Proc. Soc. Photo-Opt. Instrum. Eng. (2002).

[C65] M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset et R. Lenzen, *Calibration of CONICA static aberrations by phase diversity*, Dans *Instrumental Design and Performance for Optical/Infrared Ground-Based Telescopes*, sous la direction de M. Iye et A. F. M. Moorwood, vol. 4841, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (2002), Date conférence : août 2002, Waikoloa, Hawaii.

[C66] T. Fusco, L. M. Mugnier et J.-M. Conan, *MISTRAL, a deconvolution algorithm for astronomical adaptive optics images*, Dans *Scientific Highlights 2002*, sous la direction de F. Combes et D.Barret, EDP Sciences, SF2A (2002).

[C67] T. Fusco, L. M. Mugnier, J.-M. Conan, F. Marchis, G. Chauvin, G. Rousset, A.-M. Lagrange, D. Mouillet et F. Roddier, *Deconvolution of astronomical images obtained from ground-based telescopes with adaptive optics*, Dans *Adaptive Optical System Technology II*, sous la direction de P. L. Wizinowich et D. Bonaccini, vol. 4839, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng. (2002).

[C68] L. M. Mugnier et G. L. Besnerais, *Optical Transfer Function Identification from Satellite Images*, Dans *Earth Observing Systems VI*, sous la direction de W. L. Barnes, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (2001).

[C69] L. M. Mugnier, T. Fusco, J.-M. Conan, V. Michau et G. Rousset, *Deconvolution of adaptive optics corrrected images*, Dans *Scientific Highlights 2001*, sous la direction de F. Combes, D.Barret et F. Thévenin, pp. 593–596, EDP Sciences, SF2A (2001).

[C70] J.-M. Conan, B. Le Roux, D. Bello, T. Fusco, L. M. Mugnier, V. Michau et G. Rousset, *MultiConjugate Adaptive Optics: performance with optimal wavefront reconstruction*,

Dans *Scientific Highlights 2001*, sous la direction de F. Combes, D.Barret et F. Thévenin, pp. 541–544, EDP Sciences, SF2A (2001).

[C71] J.-M. Conan, T. Fusco, L. M. Mugnier, V. Michau, G. Rousset et P.-Y. Madec, *Imagerie haute résolution grand champ par optique adaptative multiconjuguée*, Dans *Ateliers de l'optique en astronomie*, sous la direction de J.-P. Lemonnier, M. Ferrari et P. Kern, pp. 31–36, Grenoble, France, INSU/CNRS (2001).

[C72] G. L. Besnerais et L. M. Mugnier, *Transfer Function Estimation for Spaceborne Telescopes*, Dans *Proceedings of the International Conference on Image Processing*, pp. 826–829, Los Alamitos, California, IEEE, IEEE Computer Society (octobre 2001).

[C73] T. Fusco, J.-M. Conan, V. Michau, G. Rousset et L. Mugnier, *Isoplanatic angle and optimal guide star separation for multiconjugate adaptive optics*, Dans *Adaptive Optical Systems Technology*, sous la direction de P. Wizinowich, vol. 4007, pp. 1044–1055, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng. (2000).

[C74] J.-M. Conan, T. Fusco, L. Mugnier, F. Marchis, C. Roddier et F. Roddier, *Deconvolution of adaptive optics images: from theory to practice*, Dans *Adaptive Optical Systems Technology*, sous la direction de P. Wizinowich, vol. 4007, pp. 913–924, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (2000).

[C75] A. Blanc, J. Idier et L. M. Mugnier, *Novel estimator for the aberrations of a space telescope by phase diversity*, Dans *UV, Optical, and IR Space Telescopes and Instruments*, sous la direction de J. B. Breckinridge et P. Jakobsen, vol. 4013, pp. 728–736, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (2000).

[C76] A. Blanc, L. M. Mugnier et J. Idier, *Marginal estimator for the aberrations of a space telescope by phase diversity*, Dans *4th International Conference On Space Optics*, pp. 77–86, Toulouse, France, CNES, CNES (décembre 2000).

[C77] L. M. Mugnier, F. Cassaing, G. Rousset et B. Sorrente, *Earth observation from a high orbit: pushing the limits with synthetic aperture optics*, Dans *Space-based observation techniques*, Samos, Greece, NATO/RTO-SET (octobre 2000).

[C78] T. Fusco, J.-M. Conan, L. Mugnier, V. Michau et G. Rousset, *Post-processing for anisoplanatic adaptive optics corrected images*, Dans *Propagation through the Atmosphere IV*, sous la direction de M. Roggemann, pp. 108–119, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng. (août 2000).

[C79] T. Fusco, J.-M. Conan, V. Michau, L. Mugnier et G. Rousset, *Optimal phase reconstruction in large field of view: application to multiconjugate adaptive optics systems*, Dans *Propagation through the Atmosphere IV*, sous la direction de M. Roggemann, vol. 4125, pp. 65–76, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng. (août 2000).

[C80] J.-M. Conan, T. Fusco, L. M. Mugnier et F. Marchis, *MISTRAL: Myopic Deconvolution Method Applied to ADONIS and simulated VLT-NAOS Images*, ESO Messenger, 99, pp. 38–45 (mars 2000).

[C81] T. Fusco, J.-M. Conan, L. M. Mugnier, V. Michau et J.-P. Véran, *Caractérisation et traitement d'images astronomiques à réponse impulsionnelle variable dans le champ*, Dans *17ième Colloque sur le Traitement du Signal et des Images*. GRETSI (1999).

[C82] L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau et S. Salem, *Regularized multiframe myopic deconvolution from wavefront sensing*, Dans *Propagation through the Atmosphere III*, sous la direction de M. C. Roggemann et L. R. Bissonnette, vol. 3763, pp. 134–144, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (1999).

[C83] T. Fusco, J.-M. Conan, V. Michau, L. M. Mugnier et G. Rousset, *Phase estimation for large field of view: application to multiconjugate adaptive optics*, Dans *Propagation through the Atmosphere III*, sous la direction de M. C. Roggemann et L. R. Bissonnette, vol. 3763, pp. 125–133, Bellingham, Washington, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE (1999).

[C84] A. Blanc, L. M. Mugnier et J. Idier, *Estimation des aberrations d'un telescope optique par diversité de phase*, Dans *17ième Colloque sur le Traitement du Signal et des Images*, vol. 1, pp. 63–66. GRETSI (septembre 1999).

[C85] C. Robert, J.-M. Conan, L. Mugnier, V. Michau et G. Rousset, *A comparative study of modal wavefront reconstructions for the Shack-Hartmann*, Dans *International Workshop on Wavefront Sensing and its Applications*, Canterbury,England, University of Kent (juillet 1999).

[C86] J.-M. Conan, T. Fusco, L. M. Mugnier, E. Kersalé et V. Michau, *Deconvolution of adaptive optics images with imprecise knowledge of the point spread function: results on astronomical objects.*, Dans *Astronomy with adaptive optics: present results and future programs*, sous la direction de D. Bonaccini, vol. 56 de *ESO Conference and Workshop Proceedings*, pp. 121–132, Garching bei München, Germany, ESO/OSA (février 1999).

[C87] L. M. Mugnier, J.-M. Conan, T. Fusco et V. Michau, *Joint Maximum a Posteriori Estimation of Object and PSF for Turbulence Degraded Images*, Dans *Bayesian Inference for Inverse problems*, vol. 3459, pp. 50–61, San Diego, CA (USA), Proc. Soc. Photo-Opt. Instrum. Eng. (juillet 1998).

[C88] F. Cassaing, L. M. Mugnier, G. Rousset et B. Sorrente, *Key aspects in the design of a synthetic aperture optics space telescope for wide field imaging*, Dans *Catching the perfect wave*, sous la direction de N. Duric, vol. 174 de *Pub. Astron. Soc. Pacific*, Albuquerque (juin 1998).

[C89] F. Cassaing, L. Mugnier, G. Rousset et B. Sorrente, *Éléments-clés de la conception d'un instrument spatial à synthèse d'ouverture optique*, Dans *International Conference on Space Optics*, Toulouse (France), CNES (décembre 1997).

[C90] L. M. Mugnier, J.-M. Conan, V. Michau et G. Rousset, *Imagerie à travers la turbulence par déconvolution myope multi-trame*, Dans *Seizième Colloque sur le Traitement du Signal et des Images*, sous la direction de J.-M. Chassery et C. Jutten, pp. 567–570. GRETSI (septembre 1997).

[C91] J.-M. Conan, L. Mugnier, T. Fusco, V. Michau et G. Rousset, *Deconvolution of adaptive optics images using the object autocorrelation and positivity*, Dans *Optical Science, Engineering and Instrumentation*, vol. 3126, pp. 56–67, San Diego, CA (USA), Proc. Soc. Photo-Opt. Instrum. Eng. (juillet 1997).

[C92] J. M. Conan, V. Michau, L. Mugnier, P. Y. Madec et G. Rousset, *Post-processing of adaptive optics corrected images*, Dans *Science with THEMIS*, pp. 201–203, Paris, France, Observatoire de Paris (1996).

[C93] L. M. Mugnier et G. Rousset, *Pupil Configuration Optimality Criterion in Synthetic Aperture Optics*, Dans *Spaceborne Interferometry II*, sous la direction de R. D. Reasenberg, vol. 2477, pp. 124–131. Proc. Soc. Photo-Opt. Instrum. Eng. (1995), TP 1995-42.

[C94] L. M. Mugnier, D. Charlot et G. Y. Sirat, *Recent results towards a three-dimensional conoscopic camera*, SPIE International Technical Working Group on Holography Newsletter (1993).

[C95] L. M. Mugnier et G. Y. Sirat, *Reconstruction of a three-dimensional object from its conoscopic hologram*, Dans *Inverse Problems in Scattering and Imaging*, sous la direction de M. A. Fiddy, vol. 1767, pp. 287–298, San-Diego, Proc. Soc. Photo-Opt. Instrum. Eng. (décembre 1992).

Quatrième partie

Annexes

Les annexes qui suivent reproduisent une sélection de publications dans lesquelles je suis premier ou second auteur.

Annexe A

Contribution [L5] à *Astronomy with High Contrast Imaging III*, 2006

Astronomy with High Contrast III
Editors : will be set by the publisher
EAS Publications Series, Vol. ?, 2006

DATA PROCESSING IN NULLING INTERFEROMETRY: CASE OF THE DARWIN MISSION

Laurent Mugnier[1], Eric Thiébaut[2] and Adrian Belu[3]

Abstract. This paper first gives an introduction to the broad features of the Darwin mission of ESA and then describes in some detail the data processing that is necessary to detect planets and spectrally characterize them. The proposed data processing method is validated by means of simulated data.

in *Astronomy with High Contrast Imaging III*, EAS Publications Series, EDP Sciences, Les Ulis, France, 2006.

1 Introduction

This paper gives an introduction to the Darwin corner-stone mission of the European Space Agency (ESA) and then describes the data processing that is necessary to detect planets and spectrally characterize them. More details on Darwin's instrumental concepts can be found in the previous edition of this School (Rabbia, 2004).

The prospect of planets like our own harboring life can be traced back to ancient Greek and medieval scholars. The confirmation of the first extra-solar planet in 1995 (Mayor and Queloz, 1995) conferred new levels of credibility to this prospect, and has oriented noteworthy efforts of the astronomical community, through the last decade, towards the goal of detecting such planets. The possible methods for this endeavour can be grouped into two classes: indirect or direct detection. Indirect methods consist in searching for the planet's influence on its parent star: one measures the star's wobble around the center of mass of the two-body system. This is done either through astrometry, *i.e.,* variations of the star's position on the sky, or by radial (*i.e.,* "along the line of sight") velocimetry, which consists in looking for a periodic Doppler shift in the star's spectrum lines. Most

[1] ONERA, Dép^{t} d'Optique Théorique et Appliquée, B.P. 72, 92322 Châtillon cedex, France
[2] CRAL, URA 300, 9 av. Charles André, 69561 Saint Genis Laval Cedex, France
[3] Laboratoire Universitaire d'Astrophysique de Nice, Parc Valrose 06100 Nice, France

DOI: (will be inserted later)

of the 170 exoplanets detected today have been detected through this method; planets detected this way are large and relatively close to the star (Schneider, 2004). Another indirect method likely to reveal planets of the size of the Earth in the near future is the observation of their transit in front of the star (Basri et al., 2005).

In direct methods, it is the photons from the planet itself (rather than from the star), that are measured. This can be done either by direct imaging (see, e.g., (Boccaletti et al., 2005)) or by interferometry, as proposed for Darwin by Léger *et al.* (1993). The latter technique requires a somewhat sophisticated numerical reconstruction of the scene, which is studied in Sect. 5.

2 Darwin's essentials

2.1 Aims

The aim of the Darwin mission is to screen for the presence of planets down to a fraction of an Earth mass around a pre-established list of nearby target stars. By the time of Darwin launch, the frequency of telluric (or rocky) planets is likely to be known thanks to earlier missions, especially transit ones, such as Kepler and Eddington. Depending on this frequency, a fraction of Darwin observing time will be devoted to the spectroscopic follow-up of interesting planets, in order to detect possible molecular absorption features in their atmospheres. The presence of certain molecules detailed below is likely to be associated with the presence of life—they are called bio-markers or astro-biosignatures.

2.2 Basic drivers

A star like our Sun is roughly a 5 000 K black-body, and a planet with liquid water at its surface (this is the definition of a *habitable planet*) is a 300 K one (Fig. 1). The flux contrast between the star and the planet is thus "only" 10^6 at 10μm, whereas it is about 10^9 in the visible, hence the choice of thermal infrared wavelengths for the scientific sensors of Darwin. Nevertheless, 10^6 is still a huge contrast, which must be further reduced to allow for direct detection. This is possible by selectively suppressing the star's photons: this is called coronagraphy, and will be explained in Section 3.

In order to distinguish photons coming from the planet and photons coming from its star, the instrument's angular resolution must be better than their angular separation. Not surprisingly, radiative equilibrium considerations yield a value of one Astronomical Unit (AU) for the typical distance from the star at which physical conditions for liquid water are encountered; this is the mean distance of the so-called *habitable zone*. For a target at a typical distance of 10 parsecs (pc, $1\,\mathrm{pc} \simeq 3$ light years) from us, the angular separation is $\theta \approx 1\,\mathrm{AU}/10\,\mathrm{pc} = 0.1\,\mathrm{arcsec} \simeq 0.5\,10^{-6}\,\mathrm{rd}$. For an optical instrument, the resolution limit given by diffraction is of the order of λ/B, with B the maximum distance between points in the instrument's aperture; B is the diameter of the mirror for a monolithic

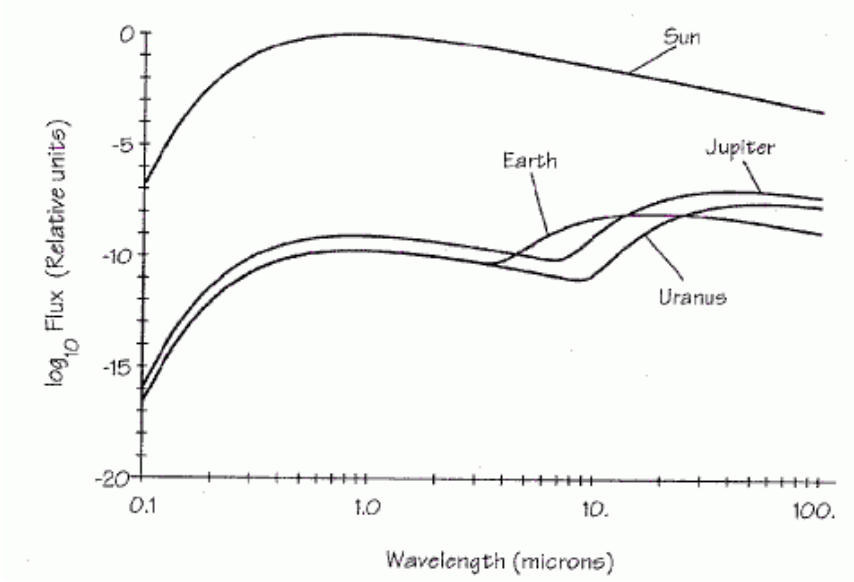


Fig. 1. Compared fluxes of Sun and solar system planets, from (Burke, 1992).

telescope or the separation between apertures for an interferometer[1]. At $\lambda = 10\,\mu$m this calls for tens of meter-sized instrument, hence the choice of an interferometric free-flyer design for Darwin.

The exact choice of the wavelengths is performed by selecting interesting biosignatures within the thermal infrared domain. The methodological choice of searching for "life as we know it" has been one of the first established consensus; this means liquid water and oxygen. The presence of water vapor in the atmosphere (Fig. 2) is an indicator of a liquid water source (water is quickly dissociated by stellar radiation into atomic oxygen and hydrogen, the latter escaping into space). The water bands on the sides of the 6–18 μm selected waveband are thus an indicator of habitability.

How do we know if that world is *inhabited* (*i.e.,* effectively harbors life)? The presence of carbon dioxide (CO_2) on a telluric planet is an indicator of emerged continents and of volcanic activity, both of which would rapidly consume any non-permanent source of oxygen (Kastings, 1997). And indeed, the 20 % of oxygen on Earth is of biological origin. Molecular oxygen (O_2) cannot be detected in the chosen spectral band, but ozone (O_3) can (Léger et al., 1994), and fortunately is a byproduct of the combination of oxygen molecules with atomic oxygen from water photolysis; it is thus an indirect marker of oxygen.

[1]For a coronagraphic interferometer, the resolution is slightly better and is actually $\lambda/(2B)$, as explained in Sect. 3.

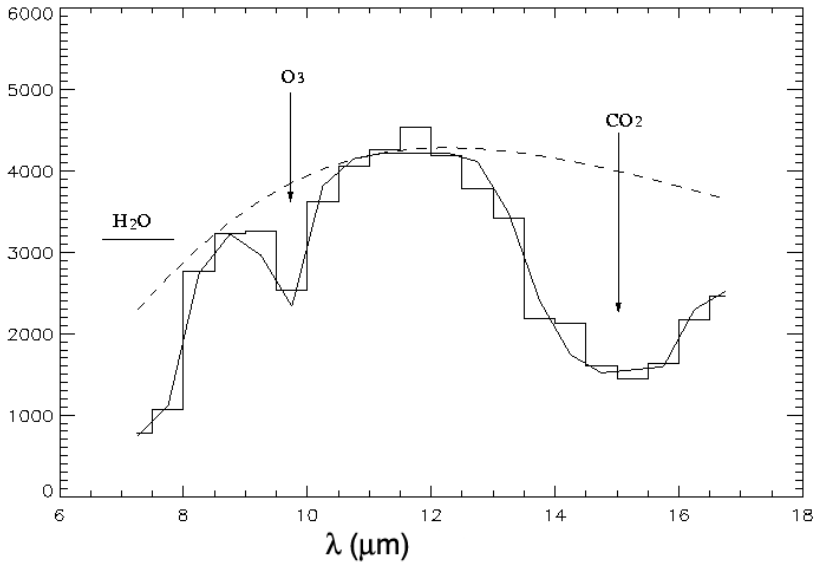


Fig. 2. Absorption features of the Earth's atmosphere (courtesy M. Ollivier), in photo-electrons per resolution element. The dashed line is the spectrum of a black-body around 300 K.

In short, the simultaneous presence of H_2O, CO_2 and O_3 is believed to be a reliable indicator of life.

2.3 Light sources

The planets range widely in terms of photometry. The reference is an exo-earth at 10pc, which shines with less than 0.1 photons/s/m^2/μm. The other, and predominant, sources of light are the star, the so-called exo-zodiacal light (see below), the zodiacal light and the instrument thermal background.

Because the star has a non-zero angular extension, it can not be totally masked by the coronagraph, as detailed in the next section.

A major source of unwanted light is the light emitted by the micrometer sized dust grains present in a planetary system (Kuchner et al., 1998), whether ours or the observed one. The dust cloud of our solar system is called zodiacal (because, like all the planets, it lies in the ecliptic plane) and that of an exo-solar system is called exo-zodiacal. Analysis of the IR excess in the stars light has shown that these exo-zodiacal dust clouds are not uncommon, and they are often thousands of times denser than ours. And the integrated flux of an exo-zodiacal cloud similar to ours already represents 1 000 times the flux of the Earth. Additionally, even though an instrument observing a target at several parsecs intercepts only a small

fraction of the light from our own zodiacal dust cloud, the latter is a problem for such a mission, especially for satellites near the Earth where a higher density of dust is trapped in resonant orbit with the Earth (Dermott et al., 1994). That is why the Darwin mission will operate 1/100th AU further away from the Sun, at Lagrange point L2.

Finally, the interstellar medium (molecules or dust) and the instrument thermal emission also add to the uniform background of the zodiacal light.

3 Basics of nulling interferometry

Let us consider an ideal two-telescope interferometer observing an object of specific intensity $I(\boldsymbol{\theta})$ in the sky, where $\boldsymbol{\theta}$ denotes 2D angular coordinates. The interferometer combines the electro-magnetic fields impinging on the telescopes' apertures and records the corresponding intensity, which is the (time-averaged) square of the sum of these complex fields. In the following we shall assume that the telescopes are identical and of diameter much smaller than their separation $\mathbf{B} = \mathbf{T}_2 - \mathbf{T}_1$, where $\mathbf{T}_1$ and $\mathbf{T}_2$ denote the telescopes' positions.

As is well-known (Young's experiment), the intensity A measured for a point-source of intensity I in the direction $\boldsymbol{\theta}$ is the sum of the intensities that would be recorded by each aperture plus an interference term, which depends on the optical path difference $\boldsymbol{\theta} \cdot \mathbf{B}$ between the two arms of the interferometer (see Fig. 3):

$$A = I \left(1 + \cos \left[\frac{2\pi}{\lambda} \boldsymbol{\theta} \cdot \mathbf{B} \right] \right), \tag{3.1}$$

where λ is the imaging wavelength.

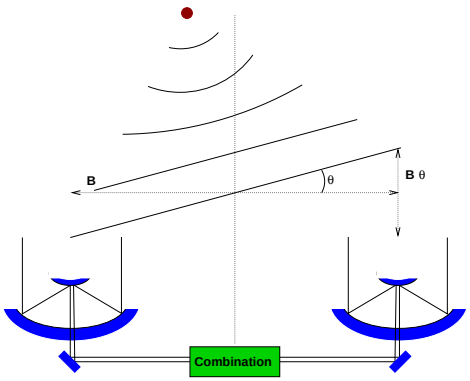


Fig. 3. Schematic 1D view of a two-telescope interferometer: for an off-axis point source at a small angle θ, the optical path difference between the two arms is θB.

For a (spatially incoherent) extended object $I(\boldsymbol{\theta})$, the measured intensity is

simply the sum of such terms for each object point:

$$A = \iint I(\boldsymbol{\theta}) \left(1 + \cos\left[2\pi\boldsymbol{\theta} \cdot \frac{\mathbf{B}}{\lambda}\right]\right) \mathrm{d}\theta \tag{3.2}$$

One can easily recognize the real part of a Fourier transformation in the oscillating term of Eq. (3.2), which shows that the recorded interference contains information on the Fourier transform of the observed object at the spatial (more precisely, angular) frequency $\mathbf{B}/\lambda$.

While this classical interpretation is useful for conventional interferometry, where the aim is to reconstruct the object (whether as a pixel map or described by only a few parameters), the following viewpoint is more fruitful for nulling interferometry: the data can be simply viewed as the total intensity of the object modulated (*i.e.,* multiplied) by a *transmission map* $R(\boldsymbol{\theta}) = 1 + \cos\left[2\pi\boldsymbol{\theta} \cdot \frac{\mathbf{B}}{\lambda}\right]$. For the two-telescope interferometer considered so far, this map is a sinusoid of period $\lambda/\mathbf{B}$, and the transmission is maximum on-axis. In order to cancel the contribution of the star to the recorded intensity, it is judicious to insert a π phase shift between the two arms of the interferometer, as suggested by Bracewell (Bracewell, 1978). Indeed, the transmission map then becomes $R(\boldsymbol{\theta}) = 1 + \cos\left[2\pi\boldsymbol{\theta} \cdot \frac{\mathbf{B}}{\lambda} + \pi\right] = 1 - \cos\left[2\pi\boldsymbol{\theta} \cdot \frac{\mathbf{B}}{\lambda}\right]$. A schematic view of the modified instrument, called Bracewell interferometer, along with the corresponding transmission map, is shown on Fig. 4.

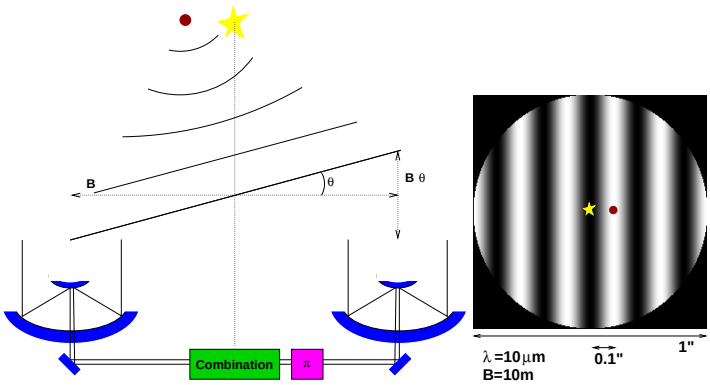


Fig. 4. Schematic 1D view of a Bracewell nulling interferometer: the light of an on axis point-source is cancelled, while the transmission is maximum for an off-axis point source at angular position $\lambda/(2B)$.

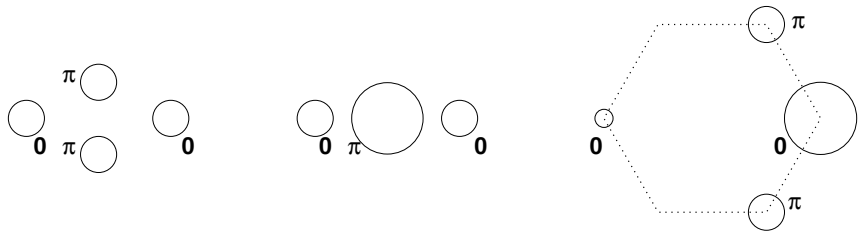


Fig. 5. Variants of the Angel cross nulling interferometer. From left to right: Angel cross, degenerated Angel cross (DAC) and generalized Angel cross (GAC). Values indicate the achromatic phase shifts introduced prior to the recombination.

4 Interferometer configurations

4.1 Reduction of the stellar leakage

As shown in the previous section, a nulling interferometer performs *coherent combination* of the light from different telescopes with proper phase shifts to achieve destructive interferences in the direction of the star. This coherent destructive combination is required to hide most of the light from the star and to obtain the contrast needed for the detection of much fainter sources (the planets). However, the destructive combination by the nulling interferometer cannot completely hide the star emission except exactly on the line of sight (for an ideal combiner), the other parts of the star limb are only attenuated as a function of the angular distance θ to the center of the field of view. This incomplete extinction of the star emission is termed *stellar leakage*.

For a simple Bracewell interferometer, the transmission scales as θ^2 around the line of sight. The higher the power of θ in the center of the transmission map, the better the extinction of the central star and the higher the planet-star contrast achievable at a given angular separation from the star. For instance, Angel (1990) proposed to coherently combine the outputs of two Bracewell interferometers with a π phase shift between the two outputs to get a θ^4 transmission map (see Fig. 5).

In principle, more complex interferometers involving more telescopes can achieve a transmission map scaling as arbitrary powers of θ. Practical considerations such as cost and spacecraft limitations will nevertheless strongly limit the number of telescopes for a real spatial nulling interferometer. For these reasons, three telescope configurations are still under study.

In practice, due to the contrast level of $\sim 10^6 - 10^9$ required for planet detection, sufficient stellar extinction cannot be achieved solely by means of coherent combination. Besides, a real instrument is subject to jitter and pointing errors which yield larger leakage since the star is not always exactly on-axis.

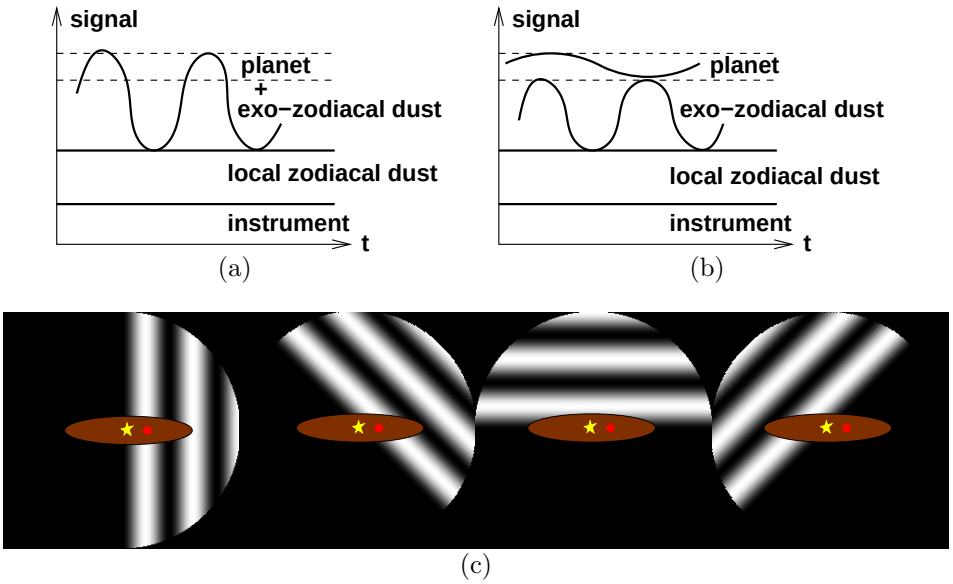


Fig. 6. (a): Modulation of the recorded signal as the interferometer is rotated for, e.g., a Bracewell interferometer. (b): same as (a) but for an asymmetric transmission map. (c): effects of an asymmetric transmission map; the exo-zodiacal emission (the grayed ellipse) and the light from the off-axis planet are modulated differently as the interferometer is rotated (from (Rabbia, 2004)).

4.2 Suppression of spurious light sources

In addition to stellar leakage, the exo-zodiacal dust emission, the zodiacal dust emission and the instrumental thermal emission are unwanted signals that reduce the ability to detect planetary emission.

The last two of these four spurious light sources have a constant value in the field of view. They can thus be distinguished from the planetary signal by rotating the interferometric array and demodulating the signal, as illustrated in Fig. 6a.

It is possible to also disentangle the first two spurious light sources from the planetary signal by designing the array so that it has an asymmetric transmission map, as illustrated in Fig. 6b and c. Indeed, the stellar leakage as well as the exo-zodiacal dust emission are expected to be symmetrically distributed with respect to the central star, whereas an off-axis planet is asymmetric, so that they are modulated differently. To this end, variants of the Angel cross have been introduced to obtain asymmetric transmission maps (see Fig. 5). Then, the signal from symmetrically distributed sources, such as the exo-zodiacal dust emission and the stellar leakage, can be removed by proper demodulation of the signal recorded during the rotation of the array—see (Mennesson and Mariotti, 1997) and Section 5.

4.3 Internal vs *external modulation*

To improve their planet detection capabilities, nulling interferometers have to perform some kind of modulation and to combine the intensities measured for different configurations. We have seen how the rotation of the interferometric array, called external modulation, can be used with asymmetric transmission maps to get rid of symmetrically distributed sources. However external modulation requires the rotation the instrument without changing the relative position of the flying telescopes. This operation is costly in terms of energy (thus reducing the mission lifetime) and implies a complex metrology. As a result, interferometer rotation is a slow operation (typically several hours for a complete revolution). Since external modulation is only effective if the level and the spatial distribution of all emission sources remain stable during the operation, this adds severe constraints to the allowed fluctuations of the levels of the instrumental thermal emission and the local zodiacal emission, but also to the exo-zodiacal, star and planetary emission level and spatial distribution (in particular no planet should move too much during the total observing time).

To overcome this problem, J.-M. Mariotti proposed, in 1997, the technique of *internal modulation* (Mennesson, 1999; Mennesson et al., 2005), which achieves different effective configurations by introducing a variable phase shift between the outputs of, at least, two nulling interferometers. More precisely, it is possible to properly choose the phase shifts and the fraction of light from each telescope involved in the different simultaneous nulling combinations to obtain asymmetric transmission maps. Figure 8 shows synthetic transmission maps obtained by means of internal modulation for a Robin-Laurance nulling interferometer. With such an interferometer, some telescopes are involved in several coherent combinations; in Fig. 7, the relative size of the pupils for the three GAC's indicates the fraction of light taken from the corresponding telescope.

Internal modulation has the appealing property that it can be performed quickly, thus relaxing some stability requirements compared to external modulation. But since transmission maps are very inhomogeneous with respect to the sky direction, the instrument must be rotated to achieve similar sensitivity for all possible planet positions. The stability constraints for this instrumental rotation are however much less severe than for external modulation. In principle, as long as the average transmission correctly samples the field of view, the instrumental motion need not be a perfect solid rotation.

Unlike nulling combination, which cancel the light in a particular direction by destructive interferences, the linear combination of detected intensities for different modulations (or configurations) is an *incoherent combination* since it is done after detection. Removing the contribution of symmetrically distributed sources by means of modulation and incoherent combination is therefore only effective for the signal part of the data, the symmetrically distributed sources still contribute to the data noise.

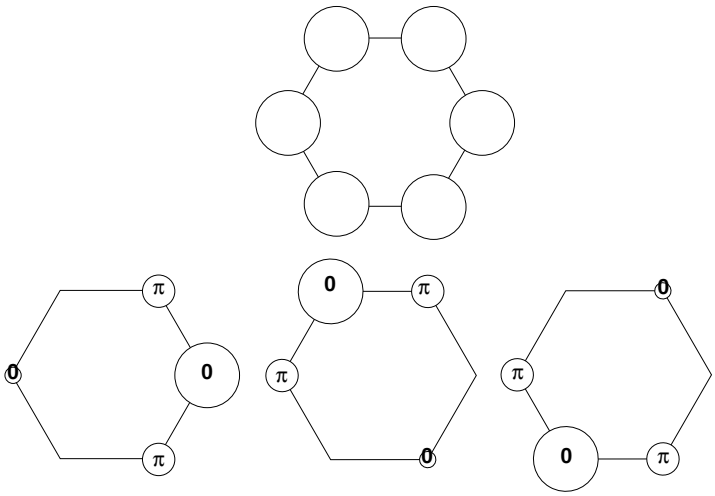


Fig. 7. Robin-Laurance interferometer. Top: geometrical setup of the six identical telescopes in the Robin-Laurance interferometer. Bottom: the three Generalized Angel crosses (GAC) interferometric configurations which can be simultaneously built from a Robin-Laurance interferometer.

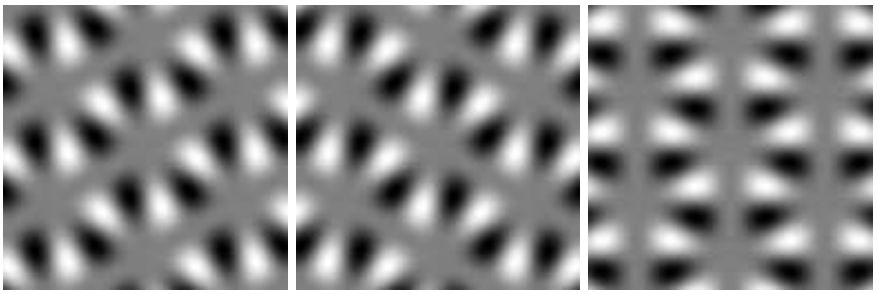


Fig. 8. Transmission maps in a Robin-Laurance interferometer after combination by internal modulation.

5 Planet detection and characterization

5.1 Data model

A straightforward generalization of the data model presented in Sect. 3 to an arbitrary transmission map yields the following one:

$$A_{t,\lambda} = \iint R_{t,\lambda}(\boldsymbol{\theta})\, I_\lambda(\boldsymbol{\theta})\, \mathrm{d}^2\boldsymbol{\theta}, \tag{5.1}$$

where $A_{t,\lambda}$ is the recorded amplitude at time t and for the effective wavelength λ, $R_{t,\lambda}(\boldsymbol{\theta})$ is the response of the instrument, *i.e.,* the transmission map at time t and for wavelength λ as a function of angular coordinates $\boldsymbol{\theta}$, and $I_\lambda(\boldsymbol{\theta})$ is the brightness distribution of the observed object (or scene) for each wavelength. In this model, we assume that the scene does not vary during the whole data acquisition. Doing otherwise would make the inversion prohibitively complicated and probably not robust.

As explained in Sect. 4, the Darwin mission will use both external modulation (*i.e.,* rotation of the interferometric array with time) and internal modulation (*i.e.,* modification of the transmission map *via* phase shifts in the beam combination), so one should bear in mind that index t in Eq. (5.1) actually codes for both the position in time and for the configuration of the phase shifts of the array.

Because the data model of Eq. (5.1) is linear, any linear combination of the data gives "synthetic" data that follow the same model, with a "synthetic" transmission map that is the corresponding linear combination of transmission maps. When using internal modulation it is possible to find coefficients for such a linear combination that make the synthetic transmission map an odd function with respect to the angular direction (Absil, 2001). Additionally, it is trivial to show from Eq. (5.1) that the data corresponding to an odd R and an even object's brightness distribution is zero (apart from the object's contribution to the noise). This is of paramount importance as all components of the observed scenes except the planets should be even within a good approximation, as mentioned above.

This means that we can, in our data model, retain only the contribution of the planets, and search only for the planets in the inversion. The object's brightness distribution for each wavelength thus reads:

$$I_\lambda(\boldsymbol{\theta}) = \sum_{i=1}^{N_{\mathrm{src}}} F_{i,\lambda}\,\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_i), \tag{5.2}$$

where N_{src} is the number of planets and $F_{i,\lambda}$ is the Spectral Energy Distribution (SED) of the i-th planet.

5.2 Inversion method and simulation results

Combining the data model of Eq. (5.1) with the object's model of Eq. (5.2) and accounting for detection noise yields:

$$A_{t,\lambda} = \sum_{i=1}^{N_{\mathrm{src}}} R_{t,\lambda}(\boldsymbol{\theta}_i) F_{i,\lambda} + n_{t,\lambda}, \tag{5.3}$$

where $n_{t,\lambda}$ is assumed to be an independent Gaussian noise, whose variance $\sigma^2_{t,\lambda}$ can be estimated from the data and is assumed known in the following.

The problem at hand is to estimate the positions $\boldsymbol{\theta}_i$ and the SED's $F_{i,\lambda}$ of the planets, assuming that their number N_{src} is known—estimating the number of planets is a difficult task that is outside the scope of this contribution. Let us denote by $(\boldsymbol{\theta}, \mathbf{F}) \triangleq (\{\boldsymbol{\theta}_i\}_{i=1}^{N_{\mathrm{src}}}, \{\{F_{i,\lambda}\}_{\lambda=\lambda_{\min}}^{\lambda_{\max}}\}_{i=1}^{N_{\mathrm{src}}})$ this set of parameters.

A well-known approach for such a detection task is correlation. One could for instance consider correlating the data recorded at a given wavelength with the noiseless data model obtained for a point-source at all possible locations (Mennesson and Mariotti, 1997; Angel and Woolf, 1997). The location yielding the highest correlation would be a probable position for a planet. Yet, let alone the fact that such a scheme is not statistically appropriate for an instrument that is not shift-invariant, it neither makes use of statistical information on the noise, nor provide a way to optimally combine the data at several wavelengths.

The Maximum Likelihood (ML) approach has none of these shortcomings; it defines the solution $(\hat{\boldsymbol{\theta}}, \hat{\mathbf{F}})$ as the set of parameters that jointly maximizes the likelihood of the data $\mathbf{A}$, which is computed by making use of the instrument and the noise models (Thiébaut and Mugnier, 2005):

$$p(\mathbf{A}|\boldsymbol{\theta}, \mathbf{F}) \propto \exp -\frac{1}{2} \left\{ \sum_{t,\lambda} \frac{1}{\sigma_{t,\lambda}^2} \left(A_{t,\lambda} - \sum_{i=1}^{N_{\mathrm{src}}} R_{t,\lambda}(\boldsymbol{\theta}_i) F_{i,\lambda} \right)^2 \right\}. \tag{5.4}$$

In practice, this maximization is performed by minimizing a criterion or cost function $J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F})$ which is the neg-log-likelihood:

$$J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F}) \triangleq \sum_{t,\lambda} \frac{1}{\sigma_{t,\lambda}^2} \left(A_{t,\lambda} - \sum_{i=1}^{N_{\mathrm{src}}} R_{t,\lambda}(\boldsymbol{\theta}_i) F_{i,\lambda} \right)^2 \tag{5.5}$$

and measures the discrepancy between the actual data $A_{t,\lambda}$ and the model of the data for the current set of parameters $(\boldsymbol{\theta}, \mathbf{F})$.

For given planet positions, the optimal SED's of all planets is obtained by minimizing J_{data} with respect to the $F_{i,\lambda}$'s. This is done by solving:

$$\frac{\partial J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F})}{\partial F_{i,\lambda}} = 0 \quad \forall i, \forall \lambda \,. \tag{5.6}$$

Because the criterion J_{data} of Eq. (5.5) is quadratic with respect to the SED's, Equation (5.6) is a set of linear equations. Its solution, denoted in the following by $\hat{\mathbf{F}}(\boldsymbol{\theta})$, is analytical, given by a simple matrix inversion, and depends on the data and on the considered planet positions.

If we replace the SED's by $\hat{\mathbf{F}}$ in the cost function $J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F})$ of Eq. (5.5) we obtain a "new" cost function with many less parameters, which only depends (explicitly) on the assumed planet positions:

$$J_{\mathrm{data}}^{\dagger}(\boldsymbol{\theta}) = J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F})_{|\mathbf{F}=\hat{\mathbf{F}}(\boldsymbol{\theta})} \,. \tag{5.7}$$

It must be noted that this new cost function $J_{\mathrm{data}}^{\dagger}(\boldsymbol{\theta})$ is nothing but the original ML cost function $J_{\mathrm{data}}(\boldsymbol{\theta}, \mathbf{F})$ that has been optimized on a subset of its parameters (namely, on the SED's). Consequently, the former has exactly the same minima as the latter and is significantly simpler to optimize.

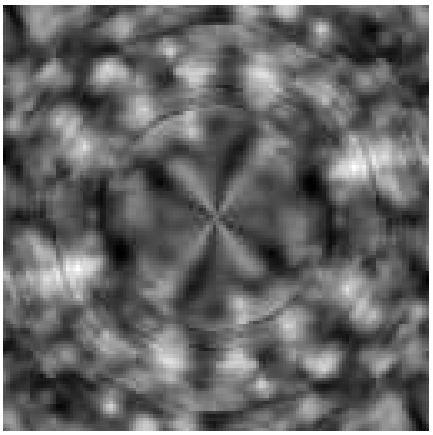


Fig. 9. $J^{\dagger}_{\text{data}}$ cost function for a single planet. The white color indicates where the cost function is minimal, *i.e.,* where the planet is the most likely located.

The minimum of $J^{\dagger}_{\text{data}}$ provides the most likely set of planet positions. As shown by Fig. 9, $J^{\dagger}_{\text{data}}$ is multi-modal, hence its deepest minimum cannot be found by a descent optimization algorithm: *global optimization is required.* To find the global minimum, the most obvious approach is to sample the field of view and map the cost function $J^{\dagger}_{\text{data}}$ onto the grid of planet positions. However if the grid has N_{grid} cells, this requires the solving of Eq. (5.6) and the computation of $J^{\dagger}_{\text{data}}$ for ${N_{\text{grid}}}^{N_{\text{src}}}$ cases. Such a global search is therefore limited to a modest number of planets on a reasonably small grid. In the single planet case (see Fig. 9), $J^{\dagger}_{\text{data}}$ is a 2-D pseudo-image of the field of view where local minima corresponds to the most likely locations of the planet. Obviously, this pseudo-image is even, which constitutes a sign ambiguity on the planet position. This ambiguity is due to the fact that the transmission map R is odd so that for any $(\boldsymbol{\theta}_1, \mathbf{F}_1)$, $(-\boldsymbol{\theta}_1, -\mathbf{F}_1)$ has exactly the same likelihood. Simply selecting the SED that is mostly positive removes the ambiguity. The fact that the SED of each planet should actually be positive at all wavelengths suggests an enhancement to the SED estimation, which consists in minimizing Eq. (5.5) with respect to the SED's under a positivity constraint. The solution is no longer analytical but improves the planet detection, as seen on Fig. 10.

An additional improvement on the SED estimation, which also impacts beneficially on the robustness of the planet detection, is to regularize the SED estimation by going from a ML estimation to a penalized ML estimation, also known as Maximum *a posteriori* (MAP) estimation—see, e.g. (Idier, 2001) for background on MAP estimation. In the present case, this consists in favoring smooth SED's among all SED's that are compatible with the data—see (Thiébaut and Mugnier, 2005) for details. Fig. 11 shows that such a regularization tends to eliminate local

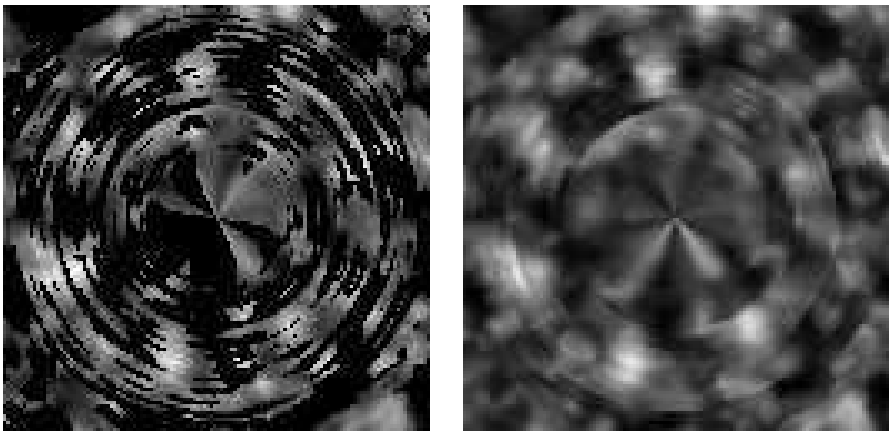


Fig. 10. $J^{\dagger}_{\mathrm{data}}$ cost function for a single planet. Left: cost function of Fig. 9 computed for an unconstrained SED, with black color where the estimated flux (total of the estimated SED) is negative. Right: cost function computed for a non-negative SED. The true planet position is correctly detected only on the right figure and is the white spot close to the bottom, slightly to the left.

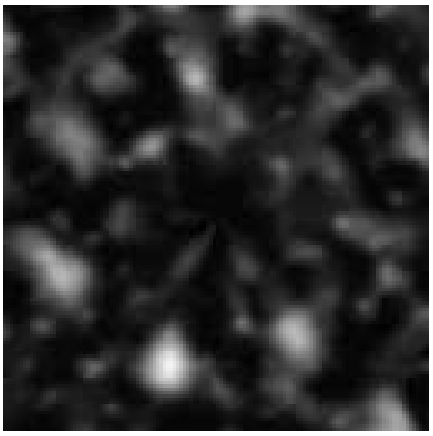


Fig. 11. Cost function for a single planet with a positivity constraint on the SED estimation and spectral regularization. The true planet position is correctly detected and is the white spot close to the bottom, slightly to the left.

minima of the cost function. It also improves the robustness to noise of the SED estimation, as shown in Fig. 12 for a two-planet example where the second planet has an Earth-like spectrum: the latter is correctly estimated only if both positivity and smoothness constraints are applied during the SED estimation.

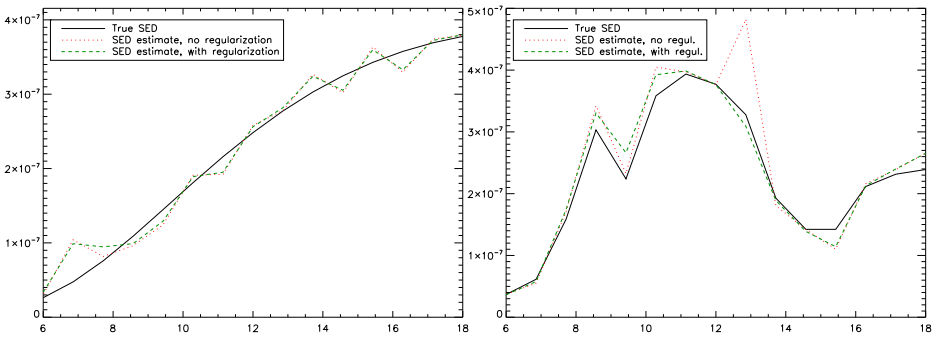


Fig. 12. Estimated SED's for a two-planet case. Continuous line: true SED's; dotted line: estimated SED's without spectral regularization; dashed line: estimated SED's with spectral regularization.

References

Absil, O.: 2001, *Nulling Interferometry with* Darwin*: Detection and Characterization of Earth-like Exoplanets*, Rapport de fin d'études d'ingénieur civil physicien de l'Université de Liège, Faculté des Sciences Appliquées

Angel, J. R. P.: 1990, in P. Bely, C. J. Burrows, and G. D. Illingworth (eds.), *Proc. NASA/STSI workshop on the Next Generation Space Telescope*, p. 81, Space telescope Science Institute, Baltimore, MD, USA

Angel, J. R. P. and Woolf, N. J.: 1997, *Astrophys. J.* **475**, 373

Basri, G., Borucki, W. J., and Koch, D.: 2005, *New Astronomy Review* **49**, 478

Boccaletti, A., Baudoz, P., Baudrand, J., Reess, J. M., and Rouan, D.: 2005, *Advances in Space Research* **36**, 1099

Bracewell, R. N.: 1978, *Nature* **274**, 780

Burke, B. F.: 1992, *TOPS, Towards Other Planetary Systems*, Technical report, NASA

Dermott, S. F., Jayaraman, S., Xu, Y. L., Gustafson, B. A. S., and Liou, J. C.: 1994, *Nature* **369**, 719

Idier, J. (ed.): 2001, *Approche bayésienne pour les problèmes inverses*, Hermès, Paris

Kastings, J. F.: 1997, *Origins of Life and Evolution of the Biosphere* pp 291–307

Kuchner, M. J., Brown, M. E., and Koresko, C. D.: 1998, *The Publications of the Astronomical Society of the Pacific* **110**, 1336

Léger et al.: 1993, *The DARWIN Mission Concept*, proposal to the Horizon 2000+ program, ESA

Léger, A., Pirre, M., and Marceau, F. J.: 1994, *Advances in Space Research* **14**, 117

Mayor, M. and Queloz, D.: 1995, *Nature* **378**, 355

Mennesson, B.: 1999, *Ph.D. thesis*, Univ. Paris VII, Paris

Mennesson, B., Léger, A., and Ollivier, M.: 2005, *Icarus* **178**, 570

Mennesson, B. and Mariotti, J.-M.: 1997, *Icarus* **128**, 202

Rabbia, Y.: 2004, in C. Aime and R. Soummer (eds.), *Astronomy with High Contrast Imaging II: Instrumentation for Coronography and Nulling Interferometry*, Vol. 12 of *EAS Publications Series*, pp 215–234, EDP Sciences, Les Ulis, France

Schneider, J.: 2004, *Exoplanet Encyclopedia*, http://cfa-www.harvard.edu/planets/

Thiébaut, E. and Mugnier, L.: 2005, in *IAUC 200, Direct Imaging of Exoplanets: Science & Techniques*, Conference date: Oct. 2005, Nice, France

Annexe B

Contribution [L6] à *Advances in Imaging and Electron Physics*, 2006

Contribution to
ADVANCES IN IMAGING & ELECTRON PHYSICS
Vol. 141, pages 1–76, edited by Peter Hawkes,
Elsevier, 2006

Phase Diversity: a technique for Wave-Front Sensing and for Diffraction-Limited Imaging

Laurent M. Mugnier, Amandine Blanc and Jérôme Idier *

Final version. Last compiled on January 19, 2011

*L. M. M. is with ONERA/DOTA, BP 72, 92322 Châtillon cedex, France. A. B. was with ONERA/DOTA at the time this work was done. J. I. is with IRCCyN/ADTSI, 1 rue de la Noe, BP 92101, 44321 Nantes Cedex 3, France.

ISSN 1076-5670/05

Abstract

The theoretical angular resolution of an optical imaging instrument such as a telescope is given by the ratio of the imaging wavelength lambda over the aperture diameter D of the instrument.

For real-world instrument, optical aberrations often prevent this so-called diffraction-limit resolution lambda/D from being achieved. These aberrations may arise both from the instrument itself and from the propagation medium of the light. The aberrations can be compensated either during the image acquisition by real-time techniques or a posteriori, i.e., by post-processing. Most of these techniques require the measurement of the aberrations, also called wave-front, by a wave-front sensor (WFS).

The focal-plane family of sensors was born from the very natural idea that an image of a given object contains information not only about the object, but also about the wave-front. A focal-plane sensor thus requires little or no optics other than the imaging sensor; it is also the only way to be sensitive to all aberrations down to the focal plane.

The first practical method for wave-front sensing from focal-plane data was proposed by Gerchberg and Saxton (1972). This so-called "phase-retrieval" method has two major limitations. Firstly, it only works with a point source. Secondly, there is generally a sign ambiguity in the recovered phase, i.e., the solution is not unique, as will be detailed below. Gonsalves (1982) showed that by using a second image with an additional known phase variation with respect to the first image (such as defocus), it is possible to estimate the unknown phase even when the object is extended and unknown. The presence of this second image additionally removes the above-mentioned sign ambiguity of the solution. This technique is referred to as "phase diversity"

This contribution attempts to provide a survey of the phase diversity technique, with an emphasis on its wave-front sensing capabilities.

Section 1 gives an introduction to the image formation for the considered instruments (i.e. those working with spatially incoherent light, such as telescopes), reviews the sources of image degradation, and states the inverse (estimation) problem to be solved in phase diversity. Section 2 reviews the domains of application of phase diversity. Then, Sections 3 and 4 review the wave-front estimation methods associated with this technique and their properties, while Section 5 examines the possible object estimation (i.e., image restoration) methods. Section 6 gives some background on the various minimization algorithms that have been used for phase diversity. Section 7 illustrates the use of phase diversity on experimental data for wave-front sensing. Finally, Sections 8 and 9 highlight two fields of phase diversity wave-front sensing that have witnessed noteworthy advances: Section 8 reviews the methods used to estimate the large-amplitude aberrations that one faces when imaging through turbulence, and proposes a novel approach for this difficult problem. And Section 9 reviews the developments of phase diversity for a recent application: the phasing (also called cophasing) of multi-aperture telescopes.

Contents

1 Introduction and problem statement

1.1 Context

The theoretical angular resolution of an optical imaging instrument such as a telescope is given by the ratio of the imaging wavelength λ over the aperture diameter D of the instrument.

For real-world instrument, optical aberrations often prevent this so-called diffraction-limit resolution λ/D from being achieved. These aberrations may arise both from the instrument itself and from the propagation medium of the light. When observing Space from the ground, the aberrations are predominantly due to atmospheric turbulence: inhomogeneities of air temperature induce inhomogeneities of refraction index.

The aberrations can be compensated either during the image acquisition by real-time techniques or *a posteriori*, i.e., by post-processing. Adaptive optics (AO) is a technique to compensate in real-time for turbulence-induced aberrations (Roddier, 1999). Most of these techniques require the measurement of the aberrations, also called wave-front, by a wave-front sensor (WFS).

There is today a large number of WFSs, which are thoroughly reviewed in (Rousset, 1999) and can be classified into two families: focal-plane sensors and pupil-plane sensors. Today's AO systems use either Shack-Hartmann WFSs (Shack and Plack, 1971) or Curvature WFSs (Roddier, 1988), which both divert part of the incoming light by means of a (dichroic) beam-splitter into some auxiliary optics and belong to the second family. For AO they both have the appealing property that they work with broad-band light (because they are well described by geometrical optics) and that the relationship between the unknown wave-front and the recorded data is linear, so that it can be inverted in real-time.

The focal-plane family of sensors was born from the very natural idea that an image of a given object contains information not only about the object, but also about the wave-front. A focal-plane sensor thus requires little or no optics other than the imaging sensor; it is also the only way to be sensitive to all aberrations down to the focal plane. The first practical method for wave-front sensing from focal-plane data was proposed by Gerchberg and Saxton (1972) in the electron microscopy context and later re-discovered by Gonsalves (1976). If the pupil (or aperture) function of the imaging system is known, this method only requires one focal-plane image of a point source in order to estimate the aberrations, which are coded in the phase of the pupil transmittance. It finds the aberrations that are the most compatible ones with the known constraints in the pupil plane (known aperture) and in the focal plane (measured image). The original implementation uses projections: it works by imposing the known constraints on the wave's complex amplitude alternatively in the two domains until convergence. The connection between this projection-based algorithm and the minimization of a least-square functional of the unknown aberrations was later made by Fienup (1982). This so-called "phase-retrieval" method has two major limitations. Firstly, it only works with a

point source. Secondly, there is generally a sign ambiguity in the recovered phase, i.e., the solution is not unique, as will be detailed below.

Gonsalves (1982) showed that by using a second image with an additional known phase variation with respect to the first image (such as defocus), it is possible to estimate the unknown phase even when the object is extended and unknown. The presence of this second image additionally removes the above-mentioned sign ambiguity of the solution. This technique is referred to as "phase diversity" by analogy with a technique used in wireless telecommunications. The idea of using two images of the same object with a known finite relative defocus in order to determine phase information from intensity measurements can actually be traced back to the work of Misell (1973), again in the electron microscopy context.

This contribution attempts to provide a (necessarily incomplete) survey of the phase diversity technique, with an emphasis on its wave-front sensing capabilities. In most of what follows, we shall consider a single-aperture[1] optical imaging instrument working with spatially incoherent light, such as a telescope. The remainder of this section gives an introduction to the image formation for such an instrument, reviews the sources of image degradation, and states the inverse (estimation) problem to be solved.

1.2 Image formation

Image formation is well-described by the scalar theory of diffraction, presented in detail in reference books such as (Goodman, 1968, Born and Wolf, 1993). It can be modeled by a convolution of the observed object by the instrument's point spread function (PSF), at least within the so-called isoplanatic patch of the instrument. At visible wavelengths, this patch is typically of the order of one degree when considering only the aberrations of the telescope itself and of the order of a few arcseconds (1 arcsec = 1/3600°) for a telescope observing Space through atmospheric turbulence.

1.2.1 PSF of a telescope

The PSF of a telescope or of the "telescope + atmosphere" system at an imaging wavelength λ is the square modulus of the (inverse) Fourier transform of the complex amplitude ψ, where $\psi = \mathcal{P}\exp(j\varphi)$ is the electromagnetic field in the instrument pupil (or aperture) when the observed object is a point source:

$$h^{\text{opt}}(x,y) = \left|\text{FT}^{-1}\left(\mathcal{P}(\lambda u,\lambda v)\, e^{j\varphi(\lambda u,\lambda v)}\right)\right|^2 (x,y). \tag{1}$$

In this expression, the Fourier transform models the transformation of the electromagnetic field between infinity and the focal plane, the square modulus is due to the quadratic detection, i.e., the detection of the field's intensity, and x and y are angles on the sky, in radians (rd). For a perfect telescope and without turbulence,

[1] as opposed to multiple-aperture instruments such as imaging interferometers, see Section 9.

$\mathcal{P}$ is constant within the pupil and φ is zero; for a real telescope, the variations of ψ are due to the aberrations of the telescope itself and to those introduced by turbulence.

From now on, we shall write $P(u,v) = \mathcal{P}(\lambda u, \lambda v)$ and $\phi(u,v) = \varphi(\lambda u, \lambda v)$ in order to deal with dimensionless quantities.

In the following, we shall assume that P is simply the indicatrix of the pupil, i.e., that the intensity variations in the pupil are negligible. This assumption is generally valid in astronomical imaging and is called the near field approximation (Roddier, 1981). With this assumption, the PSF is completely described by the pupil phase ϕ.

Equation (1) indicates that the optical transfer function (OTF) $\tilde{h}^{\text{opt}}$ is the auto-correlation of $P\exp(j\phi)$:

$$\tilde{h}^{\text{opt}}(u,v) = \left(P\exp(j\phi) \otimes P\exp(j\phi)\right)(u,v), \tag{2}$$

where the correlation of two complex-valued functions f_1 and f_2 is defined by $f_1 \otimes f_2(x) \triangleq \int f_1^\star(t) f_2(t+x)\,\mathrm{d}t$. In the absence of aberrations (i.e., if $\phi = 0$), the OTF is thus the auto-correlation of P. Consequently, it has a cutoff spatial frequency of $D/\lambda\,\mathrm{rd}^{-1}$, where D is the pupil diameter, and is strictly zero beyond. In a real system for Space observation from the ground, turbulence-induced aberrations, if uncorrected, lead to a cutoff frequency much smaller than D/λ.

1.2.2 Origin of PSF degradations: intrinsic aberrations

Some aberrations are intrinsic to the instrument; they originate in imperfections in the design, fabrication and assembly, as well as in the environment of the instrument (thermo-mechanical stresses and vibrations for a space telescope for instance).

Optical telescopes for Space observation from the ground or for Earth observation from Space are usually instruments of very good optical quality, so that the overall amplitude of these aberrations is small (notably less than the imaging wavelength λ). Their spatial spectrum is related to their origins:

- some aberrations originate in the optical design; they are fixed and of low spatial frequencies;
- some aberrations are due to the fabrication process (polishing); they are also fixed but, conversely, of high spatial frequencies;
- some aberrations are due to misalignments, either because of an imperfect alignment during the integration or because of thermo-mechanical drifts during operation. Such aberrations are slowly varying and of low spatial frequencies;
- lastly, some aberrations may occur at the location where the optical components are supported. These too are slowly varying, but may be of variable spatial frequencies.

To sum it up, the salient features of intrinsic aberrations are that they are slowly varying and usually of small overall amplitude.

1.2.3 Origin of PSF degradations: atmospheric turbulence

Inhomogeneities of the temperature of atmospheric air induce inhomogeneities of the air refraction index, which perturb the propagation of light waves through the atmosphere. We shall here assume that these random index fluctuations follow the Kolmogorov law: their probability density function is Gaussian, with zero mean and a Power Spectral Density (PSD) proportional to $|\boldsymbol{\nu}|^{11/3}$, where $\boldsymbol{\nu}$ is the 3D spatial frequency (Roddier, 1981). This assumption is usually valid, at least for the spatial frequencies of today's single-aperture telescopes.

Astronomical observation from the ground In the case of astronomical observations from the ground, the light wave coming from a point-source is planar at the entrance of the atmosphere. By integration of the index fluctuations' PSD along the optical path and within the near field approximation, it is possible to derive the spatial statistics of the phase in the telescope's pupil. This phase is Gaussian as it results from the sum of all index perturbations from the high atmosphere down to the ground (Roddier, 1981). Its PSD depends on only one parameter denoted by r_0 and reads (Noll, 1976):

$$S_\phi(f) = 0.023\, r_0^{-5/3} f^{-11/3} \tag{3}$$

where f is the modulus of the 2D spatial frequency in the pupil and r_0, called Fried diameter (Fried, 1965), is the key parameter that quantifies the turbulence's strength. It is all the smaller as the turbulence is stronger; it is typically 10 cm in the visible in a relatively good site.

The typical evolution time τ of the turbulent phase in the pupil is given by the ratio of its characteristic scale r_0 over an average wind speed (which is, more precisely, the standard deviation of the wind speed modulus distribution (Roddier et al., 1982)):

$$\tau = r_0/\Delta v. \tag{4}$$

For $r_0 \approx 10\,\text{cm}$ and $\Delta v \approx 10\,\text{m.s}^{-1}$, one gets $\tau \approx 10^{-2}\,\text{sec}$. Images corresponding to an integration time that is notably longer than this time will be referred to as "long exposure". Those of shorter exposure time will be referred to as "short exposure". For a comprehensive exposition on the temporal statistics of the turbulent phase, see, e.g., (Conan et al., 1995).

The long exposure turbulent OTF (without AO correction) is the product of the telescope's OTF without atmosphere T by an atmosphere transfer function B of cutoff frequency r_0/λ (Roddier, 1981) :

$$\tilde{h}^{\text{opt}}(f) \triangleq \langle \tilde{h}_t^{\text{opt}}(f) \rangle = T(f)\, B(f) \;\text{ where } B(f) = \exp\left\{ -3.44 (\lambda f / r_0)^{\frac{5}{3}} \right\},$$

and $\langle\cdot\rangle$ denotes a temporal average on an arbitrarily long time. Because the cutoff frequency of T is D/λ, this equation shows that the phenomenon that limits the resolution depends on the ratio D/r_0: if $D < r_0$, the instrument is diffraction-limited (if its intrinsic aberrations are reasonable), i.e., its resolution is given by its diameter, whereas if $D \gg r_0$ the long-exposure resolution of the instrument is limited by turbulence and is not better than that of a telescope of diameter r_0.

As noted by Labeyrie (1970), when the exposure time is short enough to freeze the atmospheric turbulence (typically shorter than 10 ms , *cf* Eq. (4)), some high-frequency information is preserved in the images, in the form of speckles, whose typical size is λ/D and whose position is random. This is illustrated on Fig. 1. If a number of these (uncorrected) short-exposure images are processed jointly

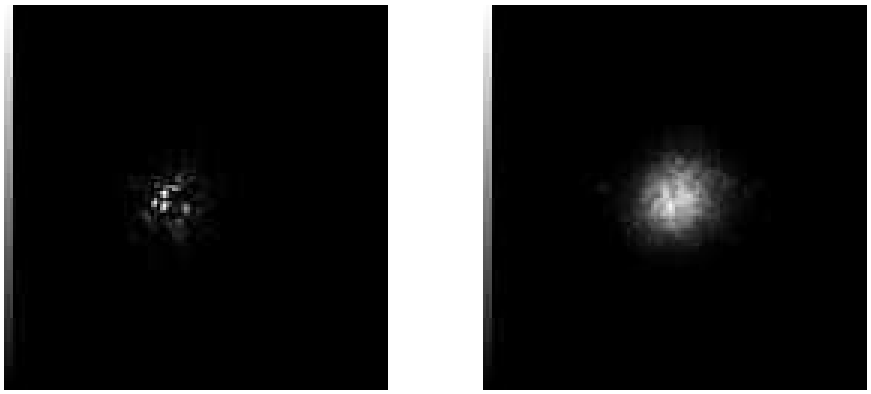


Figure 1: Simulated short-exposure (left) and long-exposure (right) images of a star through turbulence. The turbulence strength is $D/r_0 = 10$; the sampling rate respects the Shannon criterion.

in a more clever way than a simple average, it is thus possible to restore a high resolution image of the observed object.

Earth observation from Space In the case of Earth observation from Space, a light wave coming from a point-source on the ground is spherical and not planar as in astronomical observation. Because it is spherical, such a wave intersects less and thus interacts less with the lower layers of the atmosphere, which are the ones where the turbulence is strongest. It can be shown theoretically (Fried, 1966a) that the lower layers contribute less to the overall turbulence strength r_0, whose value is typically of a few tens of meters (see, e.g., (Blanc, 2002)). As a consequence, turbulence is not a limiting factor in the case of a space telescope observing the Earth.

1.2.4 Parameterization of the phase

The pupil phase ϕ can often be described parsimoniously when expanded on a modal basis. The Zernike polynomials (Noll, 1976) form an orthonormal basis on a disk and thus make a convenient basis for the expansion of the phase on the circular pupil of a telescope. Each of these polynomials is the product of a trigonometric function of the polar angle θ with a polynomial function of the radius r:

$$Z_i(\boldsymbol{r}) = R_n^m(r)\Theta_n^m(\theta) \tag{5}$$

where the trigonometric function reads:

$$\Theta_n^m(\theta) = \begin{cases} \sqrt{n+1} & \text{if} \quad m = 0 \\ \sqrt{2(n+1)}\cos(m\theta) & \text{if} \quad m \neq 0 \quad \text{and} \quad \text{i} \quad \text{even} \\ \sqrt{2(n+1)}\sin(m\theta) & \text{if} \quad m \neq 0 \quad \text{and} \quad \text{i} \quad \text{odd} \end{cases} \tag{6}$$

and the polynomial function reads:

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\,[(n+m)/2-s]!\,[(n-m)/2-s]!}\, r^{n-2s}. \tag{7}$$

Parameter n in Eqs. (6) and (7) is called the radial degree of the corresponding Zernike polynomial and parameter m is called the azimuthal degree; m and n have the same parity and are such that $0 \leq m \leq n$. The first Zernike polynomials are represented on Fig. 2. Several additional properties make this Zernike basis very commonly used:

- the first Zernike polynomials correspond to the well-known low-order optical aberrations: Z_4 is defocus, Z_5 and Z_6 are astigmatism, Z_7 and Z_8 are coma, and Z_{11} is spherical aberration;
- their ordering by increasing radial degree corresponds to an order of increasing spatial frequency.

The expansion of the phase ϕ on this basis reads:

$$\phi(\boldsymbol{r}) = \sum_{k=1}^{\infty} a_k\, Z_k(\boldsymbol{r}), \tag{8}$$

where $\boldsymbol{r}$ denotes the spatial coordinates in the pupil, normalized to a unit radius. For a multiple-aperture instrument, such an expansion can be used on each of the apertures. The first term is called piston (Z_1) and codes for the average optical path difference of a given aperture; the two following terms (Z_2 and Z_3) are tip/tilt and code for the position of the aperture's PSF. For a single-aperture telescope, the sum in Eq. (8) is usually started with $k = 4$, which corresponds to a centered PSF. Additionally, in practice, the sum is necessarily limited to a finite number k_{max} which depends on the problem at hand. It is of the order of 10 or a few tens when estimating the aberrations of a space telescope, and of the order of 100 or a few hundreds when observing Space from the ground.

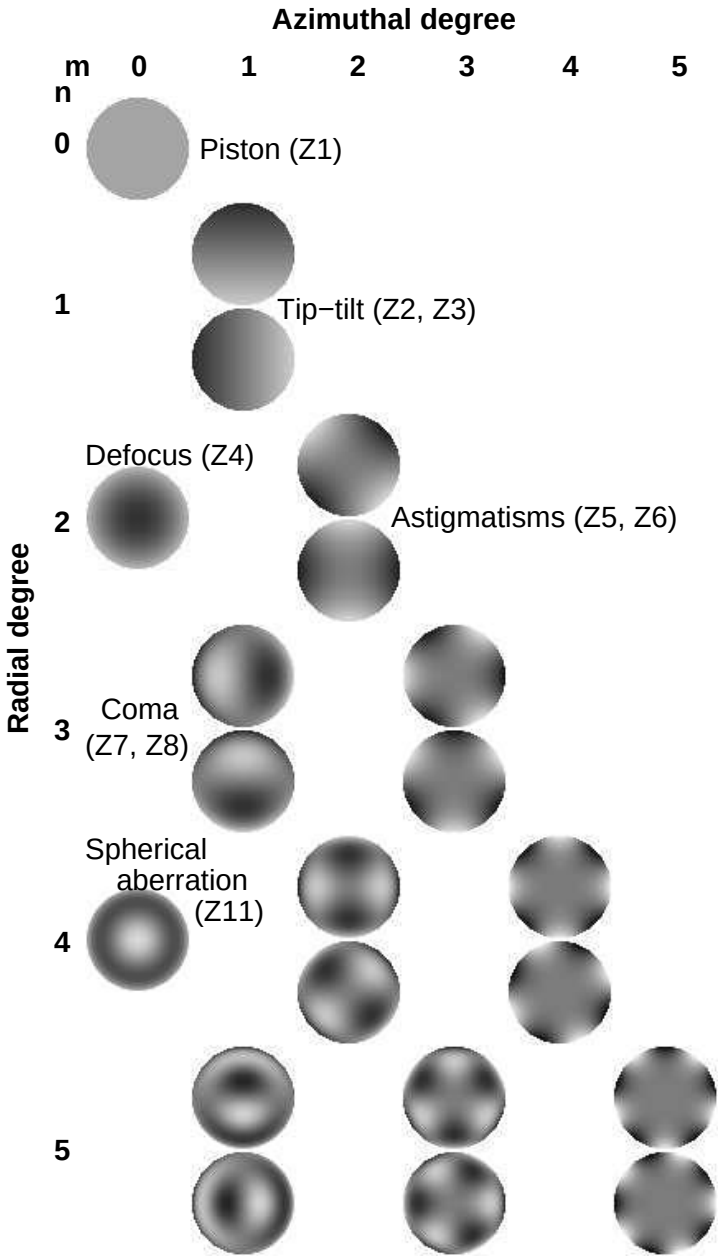


Figure 2: First Zernike polynomials.

1.2.5 Discrete image model

The image is recorded by a detector such as a CCD camera, which integrates the flux on a grid of pixels. This can be conveniently modeled as the convolution by a detector PSF h^{det} followed by a sampling operation. The global PSF of the instrument is thus:

$$h = h^{\text{det}} \star h^{\text{opt}}. \tag{9}$$

Due to the inevitable noise of the recording process (photon noise and detector noises), the recorded image reads:

$$\boldsymbol{i} = [h \star o]_{\text{Ш}} + \boldsymbol{n} \tag{10}$$

where $[\,\cdot\,]_{\text{Ш}}$ denotes the sampling operation.

This model is generally approximated by a discrete convolution with the sampled version of the (unknown) object o, and written in matrix form:

$$\boldsymbol{i} = \boldsymbol{h} \star \boldsymbol{o} + \boldsymbol{n} = H\,\boldsymbol{o} + \boldsymbol{n}, \tag{11}$$

where H is the matrix representing the discrete convolution by the sampled version $\boldsymbol{h}$ of h, and where $\boldsymbol{i}$ is the vector obtained by stacking together the columns of the corresponding image. Similarly, $\boldsymbol{o}$ is the vector obtained by stacking together the columns of the sampled object.

1.3 Basics of phase diversity

1.3.1 Uniqueness of the phase estimate

We mentioned in Section 1.1 that phase retrieval from a single image generally faces non-uniqueness of the solution, even if the object is known. This is due to the relationship between the OTF and the pupil phase (Eq. (2)), as shown below.

For any complex-valued function f, a simple change of variables in the integration shows that the function f' defined as $f'(t) \triangleq f^{\star}(-t)$ and f have identical auto-correlations: $f \otimes f = f' \otimes f'$.

Let $f(t) = P(t)e^{j\phi(t)}$, one obtains $f'(t) = P(-t)e^{-j\phi(-t)}$ as P is real-valued; thus, if P is centro-symmetrical (i.e., even), for any phase $\phi(t)$, the phase defined by $\phi'(t) = -\phi(-t)$ yields the same OTF, i.e., $\tilde{h}^{\text{opt}}(\phi') = \tilde{h}^{\text{opt}}(\phi)$ (as noted by Gonsalves (1976)), and thus the same image. This result can be cast into a somewhat more informative form: if the phase is decomposed (uniquely) into its even and odd components, i.e., $\phi(t) = \phi_{even}(t) + \phi_{odd}(t)$, then one gets $\phi'(t) = -\phi_{even}(t) + \phi_{odd}(t)$. In other words, there is an indetermination on the sign of the even part of the phase (Blanc, 2002).

Recording a second image of the same object with the instrument suffering from the same unknown phase plus a known additional *even* one removes this indetermination and adds enough information to retrieve both the phase and the possibly unknown object. More quantitative results on the uniqueness of the phase

estimate can be found in Idier et al. (2005). Let ϕ_d be this "diversity" phase (often defocus), the two images read:

$$\boldsymbol{i}_1 = \left[\boldsymbol{h}^{\text{det}} \star \boldsymbol{h}^{\text{opt}}(\phi)\right] \star \boldsymbol{o} + \boldsymbol{n_1} \tag{12}$$

$$\boldsymbol{i}_2 = \left[\boldsymbol{h}^{\text{det}} \star \boldsymbol{h}^{\text{opt}}(\phi + \phi_d)\right] \star \boldsymbol{o} + \boldsymbol{n_2} \tag{13}$$

1.3.2 Inverse problems at hand

The phase diversity technique can be used in two different contexts: one can be interested in imaging a remote object, for instance in solar astronomy or Space surveillance. Or one can be interested in measuring the aberrations of an imaging system, either to correct the latter in real-time, or to restore *a posteriori* the images it takes. These two problems are obviously very related but they are not identical.

In particular, when interested in imaging a remote object through unknown aberrations, one can live with the aforementioned sign ambiguity on the phase provided the object be recovered satisfactorily. And indeed, multi-frame "blind" deconvolution from short-exposure turbulent images has been successfully demonstrated (Schulz, 1993, Thiébaut and Conan, 1995), where blind here means deconvolution without a dedicated WFS but with the use of the strong constraints that each PSF is fully described by a pupil phase (*cf* Eq. (1)), and that the unknown object is identical in all images. Yet, it can be advantageous to record WFS data simultaneously with the images, in particular because blind deconvolution is usually impaired by the presence of local minima in the criterion to minimize. The WFS can be a focal-plane WFS consisting of a diverse image for each recorded image or a pupil-plane WFS such as a Shack-Hartmann (Fontanella, 1985, Primot et al., 1988, Mugnier et al., 2001).

When interested in estimating wave-fronts, the presence of (at least) two images is necessary to avoid a sign indetermination in the phase, as shown in subsection 1.3.1.

In both problematics, the basis of the inversion consists in estimating the phase[2] and the object that are consistent with the measurements, given the recorded images. The most "natural" solution, which is the one used by Gonsalves originally (Gonsalves, 1982), is based on the minimization of the following least-square criterion[3]:

$$J(\boldsymbol{o}, \boldsymbol{\phi}) = \|\boldsymbol{i}_1 - H(\boldsymbol{\phi})\boldsymbol{o}\|^2 + \|\boldsymbol{i}_2 - H(\boldsymbol{\phi} + \boldsymbol{\phi}_d)\boldsymbol{o}\|^2 \tag{14}$$

as a function of $(\boldsymbol{o}, \boldsymbol{\phi})$.

The remainder of this contribution is organized as follows: Section 2 reviews the domains of application of phase diversity. Then, Sections 3 and 4 review the wave-front estimation methods associated with this technique and their properties,

[2] the phases in the case of a sequence of image pairs.

[3] For simplicity, this criterion is stated for the case of two images per phase screen and one single phase screen; it is readily generalizable to more than two images and several phase screens.

while Section 5 examines the possible object estimation (i.e., image restoration) methods. Section 6 gives some background on the various minimization algorithms that have been used for phase diversity. Section 7 illustrates the use of phase diversity on experimental data for wave-front sensing. Finally, Sections 8 and 9 highlight two fields of phase diversity wave-front sensing that have witnessed noteworthy advances: Section 8 reviews the methods used to estimate the large-amplitude aberrations that one faces when imaging through turbulence, and proposes a novel approach for this difficult problem. And Section 9 reviews the developments of phase diversity for a recent application: the phasing (also called cophasing) of multi-aperture telescopes.

2 Applications of phase diversity

The concept of phase diversity has been first proposed by Gonsalves in 1982 as a WFS for adaptive optics. Since 1990, this method has been successfully used in several applications including astronomy, space observation and Earth observation. Phase diversity has the particularity of providing the estimation of the un-aberrated object as well as the aberrations responsible for the blurring. This method directly uses image data for the estimation of the aberrations. It is thus sensitive to all aberrations degrading the quality of the imaging telescope, contrarily to wave-front sensors such as the Shack-Hartman, which use a dedicated light path and thus suffer from non-common-path aberrations. Furthermore, the optical hardware of this technique is simple. These are at least some of the reasons why phase diversity is becoming a widespread method both to compensate quasi-static optical aberrations and to obtain diffraction-limited imaging through turbulence.

2.1 Quasi-static aberration correction of optical telescopes

Imperfections of an optical telescope can originate from design, fabrication of the optical system (e.g., polishing errors), misalignments (integration and launch) and thermo-mechanical stresses. These aberrations correspond to different ranges in term of spatial frequencies but all are slowly changing. Phase diversity using the image data from the science camera obviates the need for important auxiliary optics and thus is a strong candidate for the calibration of telescopes.

2.1.1 Monolithic-aperture telescope calibration

Space-based telescopes In the case of imaging space or Earth from space, the images are only perturbed by the imperfections of the optical system.

A first practical application of the phase diversity technique has been the determination of the Hubble Space Telescope aberrations (Roddier and Roddier, 1991, 1993, Fienup et al., 1993). In this case, the observed object was known (an unresolved star), so only the aberrations had to be estimated, resulting in a much easier problem, referred to *phase-diverse phase retrieval* (Ellerbroek et al., 1997).

Using the ability of phase diversity to also work with extended objects (including those extending beyond the field of view), studies have been made for the calibration of telescopes imaging the Earth (Blanc et al., 2003b). The implementation of the real-time correction of static optical aberrations has also successfully been demonstrated (Kendrick et al., 1998).

Ground-based telescopes Phase diversity has also been used to calibrate space imaging systems on Earth. The images obtained from Earth are mostly degraded by the deleterious effects of the atmospheric turbulence but also by the static aberrations of the system. Aberrations induced by the atmosphere are zero mean and quickly changing, unlike aberrations due to imperfections of the system.

The calibration of the whole optical system, from the entrance pupil of a telescope to the focal plane, can be done by averaging a large number of aberration estimates corresponding to a series of short-exposure pairs of images of an astronomical object (Acton et al., 1996, Baba and Mutoh, 2001). How these estimates are obtained is explained in Subsection 2.2.1. The calibration through the atmosphere has also been done, by Lee et al. (1997b), in the case where the optical instrument contains an AO system. In the latter reference, the diversity introduced in the images is unusual: no additional defocused image is required, and successive changes to the adaptive optics introduce the diversity.

If one wants to only calibrate the AO and the camera (and not the telescope itself), the most effective procedure is to install an internal point-source at the entrance of the AO system. The calibration of the non-common path aberrations of the VLT AO system called NAOS and its camera called CONICA has been recently done this way; see Blanc et al. (2003a), Hartung et al. (2003) and Section 7 for details. Phase diversity is also a practical tool for calibrating deformable mirrors (Löfdahl et al., 2000).

2.1.2 Cophasing of multi-aperture telescopes

The resolution of a telescope is ultimately limited by its aperture diameter. The latter is limited by current technology to about 10 m for ground-based telescopes and to a few meters for space-based telescopes because of volume and mass considerations.

Multi-aperture telescopes (a.k.a. interferometers) have the potential to remove these limitations. In order to reach the diffraction-limited resolution, all sub-apertures must be precisely phased with respect to one another. This so-called *cophasing* of the interferometer can be performed by use of a phase diversity sensor. Indeed, for a multi-aperture telescope as well as for a single-aperture one, the image is the result of interferences between all aperture points. Thus there is information in the image (whether focused or defocused) about the misalignments between sub-apertures, which are the specific aberrations of interferometry and can be described on each sub-aperture by the first three Zernike polynomials, called piston and tip-tilt (see Eq. (8) and Fig. 2). Section 9 is dedicated to this relatively recent application of phase diversity.

2.2 Diffraction-limited imaging through turbulence

Phase diversity can be used to correct the phase errors due to atmospheric turbulence in two ways: it can be used as an *a posteriori* correction technique (image restoration) or as a real-time WFS for adaptive optics. Note that the ability of phase diversity to recover both Wave-Front phase *and* amplitude has been demonstrated on simulated (Gonsalves, 1997) and experimental data (Jefferies et al., 2002).

2.2.1 *A posteriori* correction

For this application, the object is the parameter of interest. Image restoration by means of phase diversity can either correct all the aberrations degrading an imaging system without an AO or can be used after AO correction, as a second step, to correct for the residual aberrations.

A special processing approach has been proposed to use phase diversity for imaging through the atmosphere, called *Phase-diverse speckle* (a technique that blends the speckle imaging and the phase diversity concepts). Several short-exposure pairs of phase diversity data (in- and out-of-focus) are collected. This method has been applied for imaging through turbulence without AO, in particular for imaging satellites (Seldin et al., 1997, Thelen et al., 1999b) and for imaging the Sun.

Additionally, when the object being imaged through turbulence is very extended, the point spread function is no longer space-invariant in the field-of-view. The problem of correcting for turbulence-induced blur becomes thus more complicated. Phase diversity can accommodate for space-variant blur. Two methods have been investigated for solving this problem: correcting separately sub-fields which are smaller than the isoplanatic patch, which is the field of view in which the point spread function can be considered as space invariant (Löfdahl and Scharmer, 1994, Seldin and Paxman, 1994, Paxman et al., 1996) or using a tomographic phase reconstruction (Gonsalves, 1994, Acton et al., 1996, Thelen et al., 1999a, 2000, Paxman et al., 1994, 1998). With sub-fielding, a series of overlapping sub-frame reconstructions is combined to provide the entire corrected field-of-view. In the other, more sophisticated, approach, the volumic nature of turbulence is taken into account by reconstructing the phase in several screens located at different altitudes.

Post correction by means of phase diversity is also useful for AO corrected telescopes. Firstly because of non-common path aberrations, either unseen because outside the AO loop or corrected by the AO loop while not in the science path. Secondly because AO correction is always partial (Roggemann, 1991, Conan et al., 1994, Conan, 1994). Phase-diverse techniques have been successfully demonstrated for post-correction of binary stars in Seldin et al. (1996b), of satellites in Seldin et al. (1996a) and of the Sun in Löfdahl and Scharmer (2002).

2.2.2 Real-time wave-front correction

The correction of atmospheric turbulence can be done in real-time by using AO systems. Phase diversity is potentially a good candidate for use as a real-time AO WFS for a number of reasons: it is very simple optically; it is also easy to calibrate; and it directly relies on the image so it corrects all aberrations degrading the images (no non-common-path aberration). However, the computational time required on today's computers to obtain estimates of the wave-front with phase diversity is, for the moment, considerable compared to the evolution time of the turbulence (a few ms) so that current AO systems generally use other (pupil-plane) sensors.

Demonstrations of real-time correction have been obtained for very few corrected aberrations by Gates et al. (1994), Kendrick et al. (1994a, 1998). Efforts in making phase diversity estimation faster have thus been made: first by proposing better numerical algorithms in Vogel et al. (1998), Löfdahl et al. (1998a), then by modifying the error metric used to estimate the aberrations and object from the data (Kendrick et al., 1994b, Scharmer, 1999, Löfdahl and Scharmer, 2000). Phase diversity sensors depend on an imaging model (Equation (10)) involving convolutions which are usually implemented using Fast Fourier Transforms and thus computationally demanding. The idea of new metrics is to reduce the number of computed FFTs. The use of these metrics for real-time correction has been demonstrated only for few aberrations: even these new methods suffer from rapidly increasing computing time as the number of aberrations increases.

Another difficulty of the phase diversity WFS for this application is that it exhibits phase wrapping when the peak-to-valley phase variation is higher than 2π, which is often the case for turbulence-induced dynamic aberrations before closing the loop of the AO system. This is due to the fact that this sensor is sensitive only to the phase modulo 2π, as can be seen from Eq. (1). Recent works provide some methods to alleviate this problem—see Section 8 for details.

3 Phase estimation methods

The main problem to address in the phase diversity framework is to estimate the unknown quantities (the object $\boldsymbol{o}$ and/or the aberrated phase $\boldsymbol{\phi}$) from the data (focused and defocused images). The choice of a relevant estimator is thus essential. This section presents the conventional phase estimator found in the phase diversity literature. More precisely, it focuses on the estimation of the aberrated phase from a focal image $\boldsymbol{i}_1$ and an additional defocused one $\boldsymbol{i}_2$ obtained from a single-aperture telescope. The estimation methods presented here can be easily generalized to more than one phase screen i.e., to the phase-diverse speckle context. The estimation of the object will be discussed in Section 5 and the specificity of the estimation from segmented-aperture telescope in Section 9.

3.1 Joint Estimator

3.1.1 Joint criterion

The conventional processing scheme found in the literature is based on the joint estimation of the aberrations and of the observed object (Paxman et al., 1992). The Bayesian interpretation of such an approach is that it consists in computing the Joint Maximum A Posteriori (JMAP) estimator:

$$\begin{aligned}(\hat{\boldsymbol{o}}, \hat{\boldsymbol{\phi}})_{\text{MAP}} &= \arg\max_{\boldsymbol{o},\boldsymbol{\phi}} p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}) \\ &= \arg\max_{\boldsymbol{o},\boldsymbol{\phi}} p(\boldsymbol{i}_1|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}_n) p(\boldsymbol{i}_2|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta}_n) p(\boldsymbol{o}; \boldsymbol{\theta}_o) p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi). \end{aligned} \quad (15)$$

where $p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta})$ is the joint probability density function of the data ($\boldsymbol{i}_1$,$\boldsymbol{i}_2$), the object $\boldsymbol{o}$ and the aberrations $\boldsymbol{\phi}$. It may also depend on a set of hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_n, \boldsymbol{\theta}_o, \boldsymbol{\theta}_\phi)$. The likelihood of the data $\boldsymbol{i}_k$ is denoted by $p(\boldsymbol{i}_k|\boldsymbol{o}, \boldsymbol{\phi}; \boldsymbol{\theta})$, $p(\boldsymbol{o}; \boldsymbol{\theta}_o)$ and $p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi)$ are the *a priori* probability density functions of $\boldsymbol{o}$ and $\boldsymbol{\phi}$.

The majority of the estimation structures used in the phase diversity literature can be rewritten as Equation (15) even if they were not originally introduced in a Bayesian framework. Gonsalves (1982) proposed to use a joint least-square approach for the estimation of aberrations parameters. A maximum likelihood estimation of the unknowns was later presented in Paxman et al. (1992) under Gaussian and Poisson noise models. The Gonsalves least-square approach is equivalent to the Joint Maximum Likelihood (JML) approach presented in Paxman et al. (1992) under the Gaussian noise model. The JML, in turn, is obtained by setting $p(\boldsymbol{o}; \boldsymbol{\theta}_o) = p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi) = 1$ in the JMAP approach (Equation (15)). Bucci et al. (1999) introduces regularization under a deterministic approach. A first stochastic interpretation of joint estimation is presented in Vogel et al. (1998) and in Thelen et al. (1999b). By introducing regularization on the aberrations ($p(\boldsymbol{\phi}; \boldsymbol{\theta}_\phi) \neq 1$),

they propose a so-called Generalized Maximum Likelihood (GML) estimation[4]. The use of statistical information on both the object and the aberrations leads to the JMAP estimator (Vogel et al., 1998, Blanc et al., 2003b) of Equation (15).

Noise We shall assume here, for simplicity, that the noise is stationary white Gaussian with the same variance σ^2 for each image. The case of different variances has been presented in Löfdahl and Scharmer (1994). Hence, the likelihood $p(\boldsymbol{i}_k|\boldsymbol{o},\boldsymbol{\phi};\boldsymbol{\theta})$ reads

$$p(\boldsymbol{i}_k|\boldsymbol{o},\boldsymbol{\phi};\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{N^2/2}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{i}_k - H_k\boldsymbol{o})^t\,(\boldsymbol{i}_k - H_k\boldsymbol{o})\right),\ k=\{1,2\}. \tag{16}$$

where N^2 is the number of pixels in the image and the hyperparameter vector $\boldsymbol{\theta_n}$ reduces to σ^2. This stationary white Gaussian model is a reasonable approximation for bright and extended object (Earth or solar observations).

Object prior probability distribution Various methods have been proposed to introduce regularization on the object in the phase diversity literature. Some authors (Löfdahl and Scharmer (1994), Lee et al. (1997a)) use a low-pass filter. Terminating the iterations of the criterion before convergence is also a (somewhat *ad-hoc*) regularization strategy (Seldin and Paxman (1994), Thelen et al. (1999b)). A quadratic regularization model has been proposed (Vogel et al. (1998), Bucci et al. (1999)). We choose the latter method which is easily interpretable in a Bayesian framework, as a Gaussian prior probability distribution for the object. The general expression for such a prior is :

$$p(\boldsymbol{o};\boldsymbol{\theta}_o) = \frac{1}{(2\pi)^{N^2/2}\det(R_o)^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{o}-\boldsymbol{o}_m)^t R_o^{-1}(\boldsymbol{o}-\boldsymbol{o}_m)\right). \tag{17}$$

where $\boldsymbol{o}_m$ is the mean object and R_o its covariance matrix.

Phase prior probability distribution Concerning the aberrations, implicit regularization is achieved by expanding the phase on a finite linear combination of basis functions. Usually the aberrated phase is expanded on a finite set of Zernike polynomials (see Noll (1976) and subsection 1.2.4):

$$\phi(\boldsymbol{r}) = \sum_{k=4}^{k_{\max}} a_k Z_k(\boldsymbol{r}). \tag{18}$$

Note that coefficients a_{1-3} have not been introduced as mentioned in Section 1.2.4: the piston coefficient a_1 is the average phase and has no influence on the point spread function, and the tilt coefficients a_{2-3} introduce a shift in the image

[4]The denomination GML comes from the statistics literature and refers to a JML criterion that is penalized by a regularization term on some of the unknowns only.

that is of no importance for extended objects. In the following, we will note $\boldsymbol{a} = (a_4, ..., a_{k_{\max}})^t$, the $\{k_{\max} - 3\}$-dimensional vector gathering the aberration coefficients to be estimated. Additionally, in the case of imaging through turbulence, a statistical prior on the turbulent phase is available according to Kolmogorov model (Thelen et al., 1999b). It leads to a Gaussian prior probability distribution for the aberrations, with a zero mean and a covariance matrix R_a given by Noll (1976):

$$p(\boldsymbol{\phi}(\boldsymbol{a}); \boldsymbol{\theta}_\phi) = \frac{1}{(2\pi)^{(k_{\max}-3)/2} \det(R_a)^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{a}^t R_a^{-1} \boldsymbol{a}\right). \tag{19}$$

The *a priori* information on the aberrations $\boldsymbol{a}$ is the covariance matrix R_a, so that in this case, $\boldsymbol{\theta}_\phi = R_a$. Note that in the particular case where the aberrations are only intrinsic (see Subsection 1.2.2) and high-frequency (polishing) errors are negligible, a few Zernike coefficients are enough to describe all the aberrations, regularization due to the truncated expansion of the phase is sufficient and the *a priori* probability density function $p(\boldsymbol{\phi}(\boldsymbol{a}); \boldsymbol{\theta})$ can be omitted (leading to a GML estimator).

Criterion Under the above Gaussianity assumptions, we have :

$$\begin{aligned} p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{a}; \boldsymbol{\theta}) = & \frac{1}{(2\pi)^{N^2/2}\, \sigma^{N^2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{i}_1 - H_1\boldsymbol{o})^t(\boldsymbol{i}_1 - H_1\boldsymbol{o})\right] \\ & \times \frac{1}{(2\pi)^{N^2/2}\, \sigma^{N^2}} \exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{i}_2 - H_2\boldsymbol{o})^t(\boldsymbol{i}_2 - H_2\boldsymbol{o})\right] \\ & \times \frac{1}{(2\pi)^{N^2/2} \det(R_o)^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{o} - \boldsymbol{o}_m)^t R_o^{-1}(\boldsymbol{o} - \boldsymbol{o}_m)\right] \\ & \times \frac{1}{(2\pi)^{(k_{\max}-3)/2} \det(R_a)^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{a}^t R_a^{-1} \boldsymbol{a}\right). \end{aligned} \tag{20}$$

The JMAP approach amounts to maximizing $p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{a}; \boldsymbol{\theta})$, which is equivalent to minimizing the criterion:

$$\begin{aligned} L_{\text{JMAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) &= -\ln p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{a}; \boldsymbol{\theta}) \\ &= N^2 \ln \sigma^2 + \frac{1}{2}\ln\det(R_o) + \frac{1}{2}\ln\det(R_a) \\ &+ \frac{1}{2\sigma^2}(\boldsymbol{i}_1 - H_1\boldsymbol{o})^t(\boldsymbol{i}_1 - H_1\boldsymbol{o}) + \frac{1}{2\sigma^2}(\boldsymbol{i}_2 - H_2\boldsymbol{o})^t(\boldsymbol{i}_2 - H_2\boldsymbol{o}) \\ &+ \frac{1}{2}(\boldsymbol{o} - \boldsymbol{o}_m)^t R_o^{-1}(\boldsymbol{o} - \boldsymbol{o}_m) + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + A. \end{aligned} \tag{21}$$

where A is a constant.

Expression of $\hat{o}$ Canceling the derivative of L_{JMAP} with respect to the object gives (Gonsalves, 1982, Paxman et al., 1992) a closed-form expression for the object

$\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta})$ that minimizes the criterion for given $(\boldsymbol{a},\boldsymbol{\theta})$:

$$\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta}) = R\,(H_1^t \boldsymbol{i}_1 + H_2^t \boldsymbol{i}_2 + \sigma^2 R_o^{-1} \boldsymbol{o}_m). \tag{22}$$

where $R = (H_1^t H_1 + H_2^t H_2 + \sigma^2 R_o^{-1})^{-1}$. Substituting $\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta})$ into the criterion of Eq. (21) yields a "new" criterion that does not explicitly depend on the object:

$$\begin{aligned} L'_{\text{JMAP}}(\boldsymbol{a},\boldsymbol{\theta}) &= L_{\text{JMAP}}(\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta}),\boldsymbol{a},\boldsymbol{\theta}) \\ &= N^2 \ln \sigma^2 + \frac{1}{2}\ln\det(R_o) + \frac{1}{2}\ln\det(R_a) \\ &\quad + \frac{1}{2\sigma^2}(\boldsymbol{i}_1^t \boldsymbol{i}_1 + \boldsymbol{i}_2^t \boldsymbol{i}_2) \\ &\quad - \frac{1}{2\sigma^2}(\boldsymbol{i}_1^t H_1 + \boldsymbol{i}_2^t H_2 + \sigma^2 \boldsymbol{o}_m^t R_o^{-1}) R (H_1^t \boldsymbol{i}_1 + H_2^t \boldsymbol{i}_2 + \sigma^2 R_o^{-1} \boldsymbol{o}_m) \\ &\quad + \frac{1}{2}\boldsymbol{a}^t R_a^{-1} \boldsymbol{a} + A. \end{aligned} \tag{23}$$

The dimension of the parameter space over which the minimization of this new criterion is performed is dramatically reduced compared to the minimization of the criterion $L_{\text{JMAP}}(\boldsymbol{o},\boldsymbol{a},\boldsymbol{\theta})$ since the N^2 object parameters have been eliminated (Paxman et al., 1992). Note that there is no such closed form expression of $\hat{\boldsymbol{o}}$ with a Poisson noise model.

3.1.2 Circulant approximation and expression in Fourier domain

H_1 and H_2 correspond to convolution operators thus they are Toeplitz-block-Toeplitz (TBT) matrices. If $(o - o_m)$ can be assumed stationary, the covariance matrix R_o is also TBT. Such matrices can be approximated by circulant block circulant matrices, with the approximation corresponding to a periodization (Hunt, 1973). Under this assumption, the covariance matrix R_o and the convolution matrices H_1 and H_2 are diagonalized by the discrete Fourier transform (DFT). We can write

$$\begin{aligned} R_o &= F^{-1}\text{diag}[S_o]F, \\ H_1 &= F^{-1}\text{diag}[\tilde{h}_1]F, \\ H_2 &= F^{-1}\text{diag}[\tilde{h}_2]F. \end{aligned}$$

where F is the two-dimensional DFT matrix, $\text{diag}[x]$ denotes a diagonal matrix having x on its diagonal, tilde denotes the two-dimensional DFT, and S_o is the object power spectral density model. Thus the criterion L_{JMAP} and the closed-form expression $\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta})$ can be written in the discrete Fourier domain, leading to a faster computation:

$$L_{\text{JMAP}}(\boldsymbol{o},\boldsymbol{a},\boldsymbol{\theta}) = N^2 \ln \sigma^2 + \frac{1}{2}\sum_v \ln S_o(v) + \frac{1}{2}\ln\det(R_a) + \sum_v \frac{1}{2\sigma^2}|\tilde{\boldsymbol{\imath}}_1 - \tilde{\boldsymbol{h}}_1\tilde{\boldsymbol{o}}|^2 + \sum_v \frac{1}{2\sigma^2}|\tilde{\boldsymbol{\imath}}_2 - \tilde{\boldsymbol{h}}_2\tilde{\boldsymbol{o}}|^2 + \sum_v \frac{|\tilde{\boldsymbol{o}} - \tilde{\boldsymbol{o}}_m|^2}{2S_o(v)} + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + A. \tag{24}$$

$$\text{and} \quad \hat{\tilde{o}}(\boldsymbol{a},\boldsymbol{\theta},v) = \frac{\tilde{h}_1^*(\boldsymbol{a},v)\tilde{\imath}_1 + \tilde{h}_2^*(\boldsymbol{a},v)\tilde{\imath}_2 + \frac{\sigma^2\tilde{o}_m(v)}{S_o(v)}}{|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}}. \tag{25}$$

where v is the spatial frequency. The expression in the Fourier space of the criterion $L'_{\text{JMAP}}(\boldsymbol{a},\boldsymbol{\theta})$ is:

$$\begin{aligned} L'_{\text{JMAP}}(\boldsymbol{a},\boldsymbol{\theta}) &= L_{\text{JMAP}}(\hat{\boldsymbol{o}}(\boldsymbol{a},\boldsymbol{\theta}),\boldsymbol{a},\boldsymbol{\theta}) \\ &= N^2 \ln\sigma^2 + \frac{1}{2}\sum_v \ln S_o(v) \\ &\quad + \frac{1}{2}\sum_v \frac{\left|\tilde{\imath}_1(v)\tilde{h}_2(\boldsymbol{a},v) - \tilde{\imath}_2(v)\tilde{h}_1(\boldsymbol{a},v)\right|^2}{\sigma^2\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}\right)} \\ &\quad + \frac{1}{2}\sum_v \frac{\left|\tilde{h}_1(\boldsymbol{a},v)\tilde{o}_m(v) - \tilde{\imath}_1(v)\right|^2 + \left|\tilde{h}_2(\boldsymbol{a},v)\tilde{o}_m(v) - \tilde{\imath}_2(v)\right|^2}{S_o(v)\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}\right)} \\ &\quad + \frac{1}{2}\ln\det(R_a) + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + A. \end{aligned} \tag{26}$$

Note that the objective function first proposed by Gonsalves (1982) is only composed of the first and third terms of criterion 26. $L'_{\text{JMAP}}(\boldsymbol{a},\boldsymbol{\theta})$ must be minimized with respect to the aberrations $\boldsymbol{a}$. There is no closed-form expression for $\hat{\boldsymbol{a}}$, so the minimization is done using an iterative method (see Section 6 for a description of minimization methods). But before minimizing the criterion, the value of the regularization parameters must be chosen.

3.1.3 Tuning of the hyperparameters

Noise The noise model requires the tuning of the variance of the noise in the image $\theta_n = \{\sigma^2\}$. It can be estimated using the total flux in the image and the previously calibrated electronic noise level of the camera.

Aberrations In the case of atmospheric turbulence, the *a priori* information on the aberrations $\boldsymbol{a}$ is the covariance matrix R_a. Noll (1976) has shown that this matrix is completely defined by the ratio D/r_0 where D is the diameter of the telescope and r_0 is the Fried diameter (Fried, 1966b). The value of the latter can be obtained by a seeing-monitor or by SCIDAR measurement for example.

Object We choose the following model for S_o:

$$S_o(v) \triangleq E[|\tilde{o}(v) - \tilde{o}_m(v)|^2] = k/[v_o^p + v^p] - |\tilde{o}_m(v)|^2. \tag{27}$$

where E stands for the mathematical expectation. This heuristic model and similar ones have been quite widely used (Kattnig and Primot, 1997, Conan et al., 1998). This model introduces four hyperparameters $\boldsymbol{\theta}_o = (k, v_o, p, o_m)$. The tuning of the hyperparameters of the object $\boldsymbol{\theta}_o$ is not easy: their optimum values depend on the structure of the object. Thus, they must be estimated for each object. Unfortunately, in a joint estimation of the object and the aberrations, these hyperparameters $\boldsymbol{\theta}_o$ can not be jointly estimated with $\boldsymbol{o}$ and $\boldsymbol{a}$. Indeed, the criterion of Equation (21) degenerates when, for example, one seeks $\boldsymbol{\theta}_o$ together with $\boldsymbol{o}$ and $\boldsymbol{a}$. In particular, for the pair $\{\hat{\boldsymbol{\theta}}_o = (k = 0, v_o, p, o_m),\ \hat{o} = o_m\}$, which does not depend on the data, the criterion tends to minus infinity. So before minimizing L_{JMAP}, these hyperparameters must be chosen empirically by the user or estimated using a sounder statistical device.

3.2 Marginal estimator

In this method, the aberrations $\boldsymbol{a}$ and the hyperparameters ($\boldsymbol{\theta}_n$,$\boldsymbol{\theta}_o$) linked to the noise and the object are first estimated. Then if the parameter of interest is the object, it can be restored, in a second step, using the estimated aberrations and hyperparameters (see Section 5 for a detailed explanation).

The marginal estimator restores the sole aberrations by integrating the object out of the problem [5]. It is a Maximum A Posteriori (MAP) estimator for $\boldsymbol{a}$, obtained by integrating the joint probability density function:

$$\begin{aligned} \hat{\boldsymbol{a}}_{\text{MAP}} &= \arg\max_{\boldsymbol{a}} p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{a}; \boldsymbol{\theta}) = \arg\max_{\boldsymbol{a}} \int p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \boldsymbol{a}; \boldsymbol{\theta}) d\boldsymbol{o} \\ &= \arg\max_{\boldsymbol{a}} \int p(\boldsymbol{i}_1|\boldsymbol{a}, \boldsymbol{o}; \boldsymbol{\theta}) p(\boldsymbol{i}_2|\boldsymbol{a}, \boldsymbol{o}; \boldsymbol{\theta}) p(\boldsymbol{a}; \boldsymbol{\theta}) p(\boldsymbol{o}; \boldsymbol{\theta}) d\boldsymbol{o}. \end{aligned} \tag{28}$$

Let $\boldsymbol{I} = (\boldsymbol{i}_1\ \boldsymbol{i}_2)^t$ denote the vector that concatenates the data. As a linear combination of jointly Gaussian variables ($\boldsymbol{o}$ and $\boldsymbol{n}$), $\boldsymbol{I}$ is a Gaussian vector. Maximizing $p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{a}; \boldsymbol{\theta}) = p(\boldsymbol{I}, \boldsymbol{a}; \boldsymbol{\theta})$ is thus equivalent to minimizing the following

[5] In the vocabulary of probabilities, to integrate out (i.e. to marginalize) a quantity means to compute a marginal probability law by summing over all possible values of the quantity.

criterion:

$$L_{\text{MAP}}(\boldsymbol{a},\boldsymbol{\theta}) = \frac{1}{2}\ln\det(R_I) + \frac{1}{2}(\boldsymbol{I}-\boldsymbol{m}_I)^t R_I^{-1}(\boldsymbol{I}-\boldsymbol{m}_I) + \frac{1}{2}\ln\det(R_a) + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + B \tag{29}$$

where B is a constant, $\boldsymbol{m}_I = (H_1\boldsymbol{o}_m \; H_2\boldsymbol{o}_m)^t$ and $R_I \triangleq E[\boldsymbol{I}\boldsymbol{I}^t] - E[\boldsymbol{I}]E[\boldsymbol{I}]^t$ is the covariance matrix of $\boldsymbol{I}$.

3.2.1 Expression of R_I^{-1}

The expression of R_I^{-1} is obtained by the block matrix inversion lemma (Gantmacher, 1966):

$$R_I^{-1} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \tag{30}$$

with

$$\begin{aligned} Q_{11} &= \left[(H_1R_oH_1^t + \sigma^2 I_d) - H_1R_oH_2^t(H_2R_oH_2^t + \sigma^2 I_d)^{-1}H_2R_oH_1^t\right]^{-1} \\ Q_{12} &= -Q_{11}(H_1R_oH_2^t)(H_2R_oH_2^t + \sigma^2 I_d)^{-1} \\ Q_{21} &= -(H_2R_oH_2^t + \sigma^2 I_d)^{-1}(H_2R_oH_1^t)Q_{11} \\ Q_{22} &= \left[(H_2R_oH_2^t + \sigma^2 I_d) - H_2R_oH_1^t(H_1R_oH_1^t + \sigma^2 I_d)^{-1}H_1R_oH_2^t\right]^{-1}. \end{aligned} \tag{31}$$

3.2.2 Determinant of R_I

Let Δ be a matrix that reads

$$\Delta = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

in a block form. Its determinant is given by $\det(\Delta) = \det(A)\det(D - CA^{-1}B)$. Using this formula, it is easy to calculate the determinant of R_I:

$$\begin{aligned} \det(R_I) = {} & \det(H_1R_oH_1^t + \sigma^2 I_d) \\ & \times \det\left[H_2R_oH_2^t + \sigma^2 I_d - H_2R_oH_1^t(H_1R_oH_1^t + \sigma^2 I_d)^{-1}H_1R_oH_2^t\right]. \end{aligned} \tag{32}$$

3.2.3 Marginal criterion

Substituting the definition of $\boldsymbol{m}_I$ and the expression of R_I^{-1} of Equation (30) into $(\boldsymbol{I}-\boldsymbol{m}_I)^t R_I^{-1}(\boldsymbol{I}-\boldsymbol{m}_I)$ yields the following expression:

$$\begin{aligned} &(\boldsymbol{I}-\boldsymbol{m}_I)^t R_I^{-1}(\boldsymbol{I}-\boldsymbol{m}_I) \\ &\quad = (\boldsymbol{i}_1 - H_1\boldsymbol{o}_m)^t Q_{11}(\boldsymbol{i}_1 - H_1\boldsymbol{o}_m) + (\boldsymbol{i}_1 - H_1\boldsymbol{o}_m)^t Q_{12}(\boldsymbol{i}_2 - H_2\boldsymbol{o}_m) \\ &\qquad + (\boldsymbol{i}_2 - H_2\boldsymbol{o}_m)^t Q_{21}(\boldsymbol{i}_1 - H_1\boldsymbol{o}_m) + (\boldsymbol{i}_2 - H_2\boldsymbol{o}_m)^t Q_{22}(\boldsymbol{i}_2 - H_2\boldsymbol{o}_m). \end{aligned}$$

Basic algebraic manipulations yield the following expression for the marginal criterion:

$$\begin{aligned} L_{\text{MAP}}(\boldsymbol{a},\boldsymbol{\theta}) = & \frac{1}{2}\ln\det(R_I) + \frac{1}{2}\ln\det(R_a) \\ & + \frac{1}{2\sigma^2}(\boldsymbol{i}_1^t\boldsymbol{i}_1 + \boldsymbol{i}_2^t\boldsymbol{i}_2) \\ & - \frac{1}{2\sigma^2}(\boldsymbol{i}_1^t H_1 + \boldsymbol{i}_2^t H_2 + \sigma^2\boldsymbol{o}_m^t R_o^{-1})R(H_1^t\boldsymbol{i}_1 + H_2^t\boldsymbol{i}_2 + \sigma^2 R_o^{-1}\boldsymbol{o}_m) \\ & + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + B \end{aligned} \tag{33}$$

where $R = (H_1^t H_1 + H_2^t H_2 + \sigma^2 R_o^{-1})^{-1}$.

3.2.4 Relationship between the joint and the marginal criteria

The comparison of the expression of the criterion L_{MAP} (Equation (33)) and of the criterion L_{JMAP} (Equation (23)) shows that the two criteria are related by the following relationship:

$$\begin{aligned} L_{\text{MAP}}(\boldsymbol{a},\boldsymbol{\theta}) = & \frac{1}{2}\ln\det(R_I) - N^2\ln\sigma^2 \\ & - \frac{1}{2}\ln\det R_o + L'_{\text{JMAP}}(\boldsymbol{a},\boldsymbol{\theta}) + C \end{aligned} \tag{34}$$

where C is a constant. If we focus only on the terms depending on the phase (i.e., suppose that the hyperparameters are known), relationship 34 can be summarized by (Goussard et al., 1990):

$$L_{\text{MAP}}(\boldsymbol{a}) = \frac{1}{2}\ln\det(R_I) + L_{\text{JMAP}}(\boldsymbol{a}) + C' \tag{35}$$

where C' is a constant. Thus, the difference between the marginal and the joint estimator consists of a single additional term dependent on the phase, which is $\ln\det(R_I)$. Although the two estimators differ only by one term, we shall see in Section 4 that their properties differ considerably.

3.2.5 Expression in the Fourier domain

In practice, the marginal estimator is computed in the Fourier domain. Using the circulant approximations (see Subsection 3.1.2) for H_1, H_2 and R_o and noting that $\ln\det(R_o) = \sum_v \ln S_o(v)$, the term $\ln\det(R_I)$ (Equation (32)) can be expressed as follows:

$$\begin{aligned} \ln\det(R_I) = & \sum_v \ln S_o(v) + N^2\ln\sigma^2 \\ & + \sum_v \ln\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}\right). \end{aligned} \tag{36}$$

Combining Equations 26, 34 and 36 gives the marginal estimator in the Fourier domain:

$$
\begin{aligned}
L_{\text{MAP}}(\boldsymbol{a},\boldsymbol{\theta}) = & \sum_v \ln S_o(v) + N^2 \ln \sigma^2 \\
& + \sum_v \ln\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)} \right) \\
& + \frac{1}{2}\sum_v \frac{\left|\tilde{i}_1(v)\tilde{h}_2(\boldsymbol{a},v) - \tilde{i}_2(v)\tilde{h}_1(\boldsymbol{a},v)\right|^2}{\sigma^2\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}\right)} \\
& + \frac{1}{2}\sum_v \frac{\left|\tilde{h}_1(\boldsymbol{a},v)\tilde{o}_m(v) - \tilde{i}_1(v)\right|^2 + \left|\tilde{h}_2(\boldsymbol{a},v)\tilde{o}_m(v) - \tilde{i}_2(v)\right|^2}{S_o(v)\left(|\tilde{h}_1(\boldsymbol{a},v)|^2 + |\tilde{h}_2(\boldsymbol{a},v)|^2 + \frac{\sigma^2}{S_o(v)}\right)} \\
& + \frac{1}{2}\ln\det(R_a) + \frac{1}{2}\boldsymbol{a}^t R_a^{-1}\boldsymbol{a} + B. \qquad (37)
\end{aligned}
$$

Let us see, now, how the hyperparameters can be estimated.

3.2.6 Unsupervised estimation of the hyperparameters

For the marginal estimator, the estimation of $\boldsymbol{\theta}_o = (k, v_o, p, o_m)$ and $\boldsymbol{\theta}_n = \sigma^2$ can be jointly tackled with the aberrations, according to:

$$
(\hat{\boldsymbol{a}}, \hat{\boldsymbol{\theta}}_o, \hat{\boldsymbol{\theta}}_n) = \underset{\boldsymbol{a},\boldsymbol{\theta}_o,\boldsymbol{\theta}_n}{\arg\max}\, p(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{a}; \boldsymbol{\theta}). \qquad (38)
$$

The criterion $L_{\text{MAP}}(\boldsymbol{a})$ of Equation (29) becomes $L_{\text{MAP}}(\boldsymbol{a}, \sigma^2, k, v_o, p, o_m, \boldsymbol{\theta}_a)$. It must be minimized with respect to the aberrations $\boldsymbol{a}$ and the five hyperparameters (σ^2, k, v_o, p, o_m). If we adopt the change of variable $\mu = \sigma^2/k$, the cancellation of the derivative of the criterion with respect to k gives a closed-form expression $\hat{k}(\boldsymbol{a}, \mu, v_o, p, o_m, \boldsymbol{\theta}_a)$ which minimizes the criterion for given values of the other parameters. Injecting $\hat{k}$ into L_{MAP} yields $L_{\text{MAP}}(\boldsymbol{a}, \mu, v_o, p, o_m, \boldsymbol{\theta}_a)$. There is no closed-form expression for $\hat{\mu}$, $\hat{v}_o$, $\hat{p}$ and $\hat{o}_m$ but it is easy to calculate the analytical expression of the gradients of the criterion with respect to these hyperparameters and then to use numerical methods for the minimization of the criterion.

3.3 Extended objects

In order to process extended scenes (Earth or solar observations), the problem of edge effects must be addressed. The latter is due to the fact that the joint and the marginal criteria are expressed in the Fourier domain thanks to an approximation: the convolutions are made using FFTs. This introduces a periodization which produces severe wrap-around effects on extended objects. To solve this problem, two solutions have been proposed in the phase diversity literature: the apodization technique (Löfdahl and Scharmer, 1994) and the guard-band technique (Seldin and Paxman, 1994).

3.3.1 Apodization

To reduce the computing time required to minimize the joint and the marginal criteria, they are computed in Fourier space and thus do not take into account the effects of boundaries of the images. To reduce the edges effects while still computing $\hat{o}(\boldsymbol{a})$ with FFTs, the images can be apodized. The entire apodization of the data by a Hanning window has been first used by Paxman and Crippen (1990) but lead to poor results. Löfdahl and Scharmer (1994) suggested to only apodize the edges of the images by a modified Hanning window (an example of a 1D Hanning window and of a modified Hanning window are shown in Figure 3 for comparison). In this technique, the summation in the joint criterion expression (Equation (24),

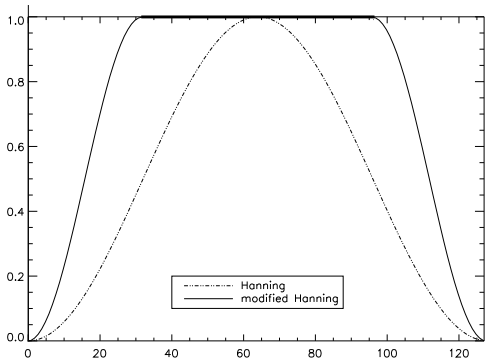


Figure 3: A 1D Hanning window (dotted line) and a 1D modified Hanning window (solid line) for comparison.

line 2) is computed in the image space instead of the Fourier space (according to the Parseval's theorem). It allows one to only keep in the summation data that have not been apodized (i.e., the ones for which the apodization function is unity). Note that this method can be easily adapted to the marginal criterion. This type of apodization has been already used in speckle techniques (Von der Lühe, 1993) and works well with phase diversity data. The advantage of this technique is to provide fast computation of the criterion. Its disadvantage, apart from the fact that it is approximate, is that a part of the data is apodized and is not used in the estimation.

3.3.2 Guard band

Another way of tackling the edge effects has been proposed for the joint estimator by Seldin and Paxman (1994), and is given below for both estimators.

Joint estimator The technique consists firstly in acknowledging the fact that object pixels beyond the field of view of the image do influence the data due to the convolution operator involved in the image formation (see Equation (10)), and estimating the object value on the guard-band pixels (i.e., pixels beyond the effective

field of view) as well. The criterion is minimized numerically with respect to the aberrations *and* the object (no fast solution for computing $\hat{o}$). The guard band width depends on the severity of the aberrations (i.e., on the effective PSF support width). Secondly, in practice, the object and the PSF 2-D arrays are immersed in arrays of size larger than the sum of their support in order to be able to compute $\boldsymbol{h} \star \boldsymbol{o}$ exactly by means of FFTs.

Marginal estimator To apply the guard band technique to the marginal estimator $L_{\text{MAP}}(\boldsymbol{a}, \boldsymbol{\theta})$, a new algorithm is used, called the "alternating" marginal estimator and noted $L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta})$. The relationship between the joint estimator and the marginal one (see Equation (34)) can be summarized by $L_{\text{MAP}}(\boldsymbol{a}, \boldsymbol{\theta}) = L'_{\text{JMAP}}(\boldsymbol{a}, \boldsymbol{\theta}) + \varepsilon(\boldsymbol{a}, \boldsymbol{\theta})$. The alternating marginal criterion is then defined by $L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) = L_{\text{JMAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) + \varepsilon(\boldsymbol{a}, \boldsymbol{\theta})$. And:

$$\begin{aligned}
\arg\min_{\boldsymbol{o},\boldsymbol{a},\boldsymbol{\theta}} L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) &= \arg\min_{\boldsymbol{a},\boldsymbol{\theta}} \left\{ \arg\min_{\boldsymbol{o}} L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) \right\} \\
&= \arg\min_{\boldsymbol{a},\boldsymbol{\theta}} \left\{ \arg\min_{\boldsymbol{o}} \left[L_{\text{JMAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta}) \right] + \varepsilon(\boldsymbol{a}, \boldsymbol{\theta}) \right\} \\
&= \arg\min_{\boldsymbol{a},\boldsymbol{\theta}} \left[L'_{\text{JMAP}}(\boldsymbol{a}, \boldsymbol{\theta}) + \varepsilon(\boldsymbol{a}, \boldsymbol{\theta}) \right] \\
&= \arg\min_{\boldsymbol{a},\boldsymbol{\theta}} L_{\text{MAP}}(\boldsymbol{a}, \boldsymbol{\theta}). \qquad (39)
\end{aligned}$$

The minimization of $L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta})$ with respect to $\boldsymbol{o}$, $\boldsymbol{a}$ and $\boldsymbol{\theta}$ is therefore equivalent to the minimization of $L_{\text{MAP}}(\boldsymbol{a}, \boldsymbol{\theta})$ with respect to the sole $\boldsymbol{a}$ and $\boldsymbol{\theta}$. The guard-band can then be applied to the criterion $L^{\text{alt}}_{\text{MAP}}(\boldsymbol{o}, \boldsymbol{a}, \boldsymbol{\theta})$.

In the guard band technique, the measured data are unperturbed but the disadvantage of this method is the extensive computation time (due to the iterative estimation of the object).

4 Properties of the phase estimation methods

In this section we study the properties of the two phase estimation methods presented in the previous section, by means of simulations. We compare their asymptotic properties, the influence of the hyperparameters on the quality of the estimated phase and finally their performance.

4.1 Image simulation

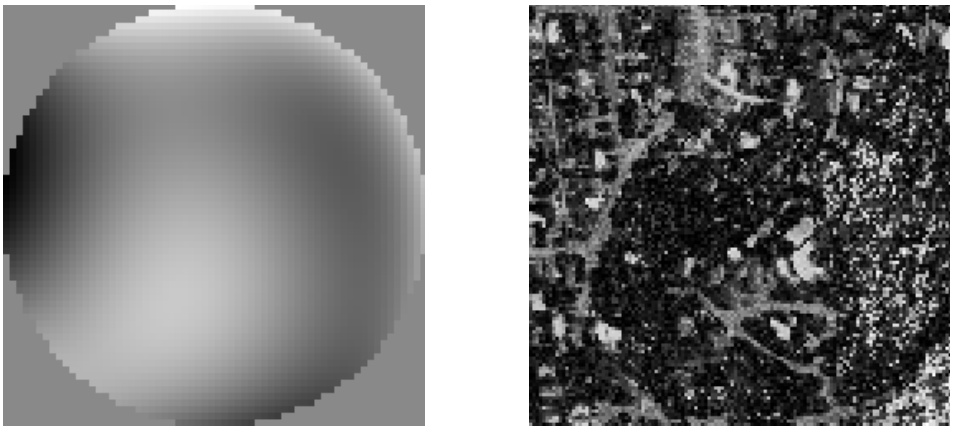
The simulations have been obtained in the following way: our object is an Earth view. The aberrations are due to the imperfections of the optical system. The phase is a linear combination of the first 21 Zernike polynomials, with coefficients listed in Table 1; the estimated phase will be expanded on the same polynomials. The defocus amplitude for the second observation plane is 2π radians, peak-to-valley. The simulated images are monochromatic and are sampled at the Shannon rate. They have been obtained by convolution between the point spread function and the object, computed in the Fourier domain using FFTs. The result is corrupted by a stationary white Gaussian noise (see Figure 4). The images generated in this way are periodic. This is an artificial situation, under which R_o is truly circulant block circulant. The fact that the images are periodic allows us to estimate the phase without the additional computing cost of the guard-band technique.

Table 1: Values of the coefficients used for simulations.

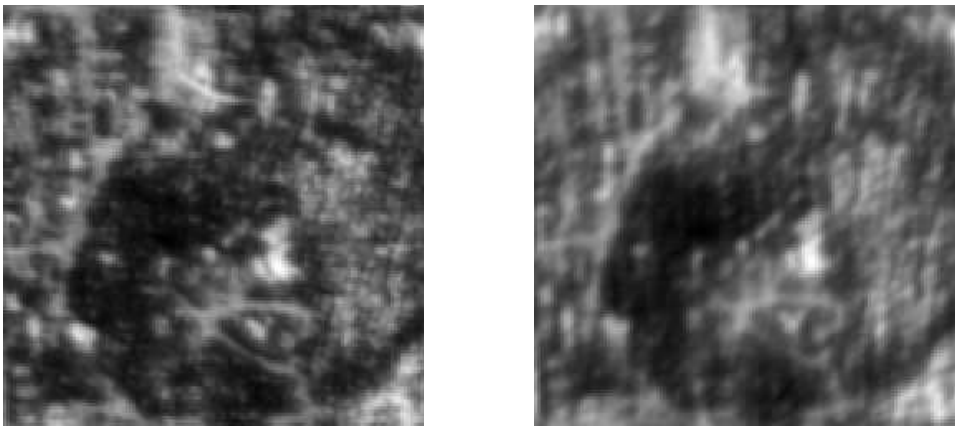
Coefficient	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
Value (rd)	-0.2	0.3	-0.45	0.4	0.3	-0.25	0.35	0.2	0.1
Coefficient	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}
Value (rd)	0.05	-0.05	0.05	0.02	0.01	-0.01	-0.02	0.01	0.01

4.2 Asymptotic properties of the two estimators for known hyperparameters

For the time being, we consider that the hyperparameters are the "true" ones, i.e., we fit the power spectral density of the object using the true object, and we assume that σ^2 is known (note that the mean object is set to zero). Additionally, as in all the following, no regularization on the aberrations is introduced, save the fact that only the Zernike coefficients a_4 to a_{21} are estimated. Consequently, the marginal estimation which was based on a MAP approach now corresponds to a Maximum Likelihood (ML) estimation. Similarly, the JMAP approach corresponds to a GML estimation because the phase is not regularized (see Subsection 3.1.1). The minimized criteria will then be denoted L_{ML} and L_{GML}, respectively. Figure 5 shows the bias, standard deviation and root mean square error (RMSE) on the phase provided



(a) (b)



(c) (d)

Figure 4: Aberrated phase (a) of RMS value $\lambda/7$ and true object (b) used for the simulation. Simulated focused (c) and defocused (d) images.

by the joint method (left) and the marginal method (right) as a function of the noise level for three image sizes (128×128, 64×64 and 32×32 pixels). These three quantities are defined as:

- the empirical bias, $b = \sum_{k=4}^{21} \left[\langle \hat{a}_k \rangle - a_k^{true} \right]$
- the empirical standard deviation, $\sigma = \left[\sum_{k=4}^{21} \langle (\hat{a}_k - \langle \hat{a}_k \rangle)^2 \rangle \right]^{1/2}$
- the empirical RMSE, $e = (b^2 + \sigma^2)^{1/2}$

The empirical average is done on 50 different noise realizations. Furthermore, for the image sizes of 32×32 and 64×64 pixels, the quantities are averaged on all the sub-images of 32×32 pixels (respectively 64×64) contained in the image of 128×128 pixels. For joint estimation, the bias increases with the noise level. Furthermore, processing a larger number of data is not favorable in term of bias. On the contrary, the standard deviation of the phase estimate is a decreasing function of the image size. Finally the RMSE, which is dominated by the bias term, does not decrease as the number of data increases. This pathological behavior meets several statistical studies (Champagnat and Idier, 1995, Little and Rubin, 1983): the estimate does not converge towards the true value as the size of the data set tends to infinity. An intuitive explanation of this phenomenon is that, if a larger image is used to estimate the aberrations, the size of the object, which is jointly reconstructed, increases also, so that the ratio of the number of unknowns to the number of data does not tend towards zero. On the contrary, for marginal estimation, the ratio of unknowns to data tends towards zero because the number of unknowns stays the same whatever the size of the data set. In this case, the bias, standard deviation and RMSE of the phase decreases when the number of data increases (see Figure 5 right). Indeed, under broad conditions, the marginal estimator is expected to converge, since it is a true ML estimator (Lehmann, 1983, Carvalho and Slock, 1997).

The curve of the joint estimation standard deviation presents an irregularity for the noise level of 14% and an image size of 128×128 pixels. This surprising result can be interpreted by looking at the different phase estimates obtained in this condition, which are shown Figure 6. The minimization of the joint criterion leads to two different sets of aberration coefficients. This explains why the standard deviation reaches a much larger value for this simulation condition. It also shows that the joint criterion presents local minima. Furthermore, we have empirically checked that it presents local minima whatever the size of the data set, whereas we do not witness such local minima with the marginal criterion (see Figure 5, right). Indeed, the marginal criterion tends to be asymptotically more and more regular. The latter observation is in agreement with the asymptotic Gaussianity of the likelihood, which is expected under suitable statistical conditions.

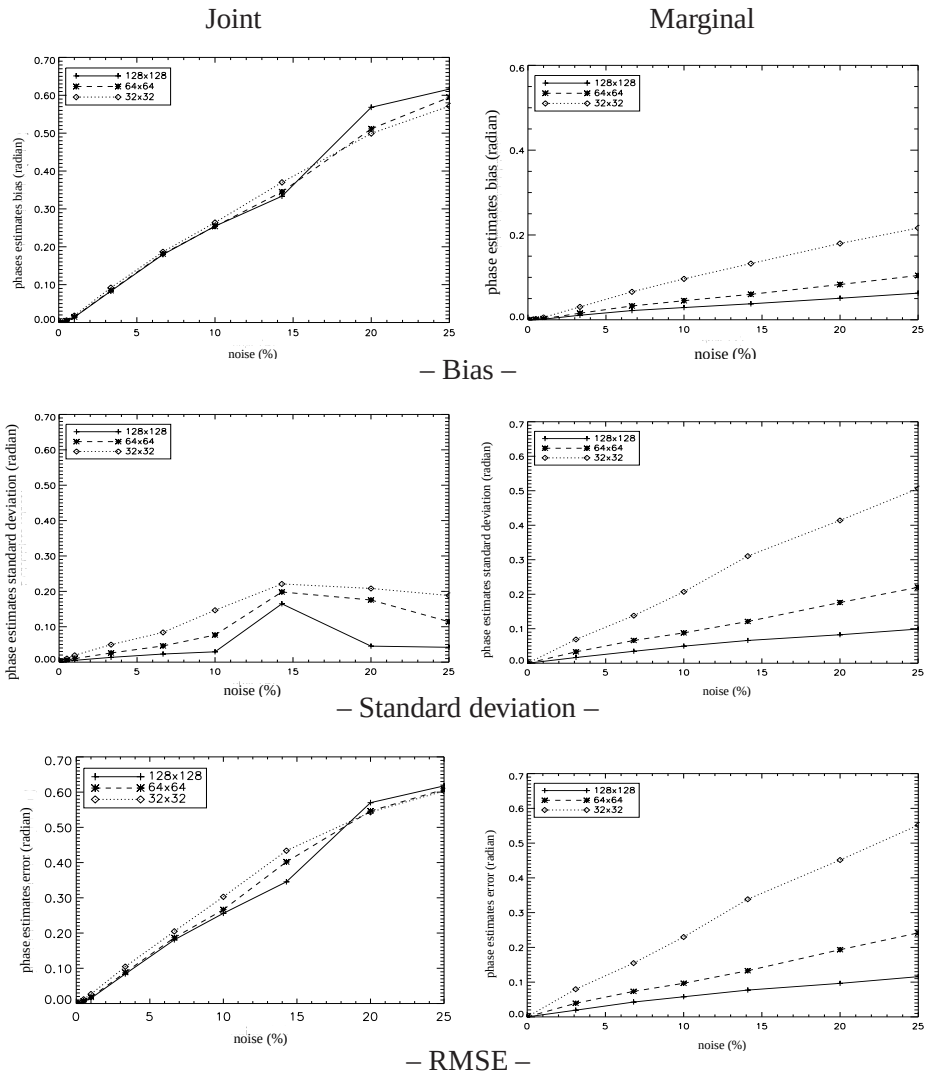


Figure 5: Bias, standard deviation and RMSE of phase estimates as a function of noise level given in percent (it is the ratio between the noise standard deviation and the mean flux per pixel). Left figures are for the joint estimator, right figures for the marginal one. The solid, dashed and dotted lines respectively correspond to images of dimensions 128×128, 64×64 and 32×32 pixels. All these estimates have been obtained as empirical averages on 50 independent noise realizations.

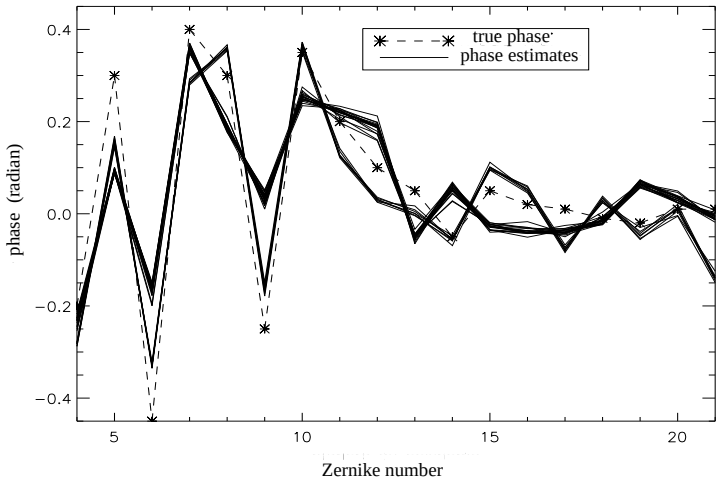


Figure 6: The different aberration estimates obtained with the joint method for the image size of 128×128 pixels and a noise level of 14 %. Dashed line: true aberrations; solid line: aberration estimates.

4.3 Joint estimation: influence of the hyperparameters

An important problem for the estimation of the aberrations (and the object) is the tuning of the hyperparameters. For the joint estimator, we have pointed out that they must be adjusted by hand. Particularly important is the global hyperparameter, which we shall denote by μ and is the one that quantifies the trade-off between goodness of fit to the data and to the prior[6]. Let us study its influence on the joint method. Figure 7 shows the RMSE on the phase estimates and on the object estimate as a function of the value of this hyperparameter (its true value is $\mu = 1$). The RMSE on the object is defined as $\left[\sum_{\boldsymbol{r}} \langle (\hat{o}(\boldsymbol{r}) - o^{true}(\boldsymbol{r}))^2 \rangle\right]^{1/2} / \left[\sum_{\boldsymbol{r}} \hat{o}(\boldsymbol{r})^2\right]^{1/2}$. We see that the best value of this hyperparameter (i.e., the one that gives the lower error on the estimate) is not the same for the object and for the phase. It means that the object and the phase can not be jointly optimally restored. Note that the optimal hyperparameter value for the object coincides with the true value $\mu = 1$. If the parameter of interest is the phase, the object must be under-regularized to have a better estimation of the aberrations. The behavior of RMSE on the phase strongly depends on the noise level: for a high noise level (14% here), there is an optimal basin but for lower noise levels (4% and below), any value under 1, including a near null regularization ("near null regularization" means that the parameter μ is not equal to zero but to a small arbitrary constant (10^{-16} in our case) in order to avoid numerical problems due to computer precision), is almost optimal with respect to estimation of the aberrations even though the jointly estimated object

[6]For the Gaussian prior used in this work, tuning μ is equivalent to tuning a scale factor in the object power spectral density S_o.

becomes of the poorest quality. This observation sheds some light on the fact that, when the parameters of interest are the aberrations and when the noise level is low, estimation without object regularization can be successfully used as the literature testifies it (Thelen et al., 1999b, Meynadier et al., 1999, Seldin and Paxman, 2000, Carrara et al., 2000). This empirical observation has also led us to study the asymptotic behavior of the joint estimator with near null regularization. Figure 8 shows the results. In this case, when the number of data increases, the RMSE on the aberrations estimates decreases. Although the number of data samples over the number of unknowns is the same as that of the estimation with the true hyperparameters, the estimator behaves as if the object were not being estimated. This surprising behavior of joint aberration estimates when the regularization parameter μ vanishes has been recently explained in Idier et al. (2005); this study has shown that the GML is a consistent phase estimator i.e., it converges towards the true value as the number of data increases.

4.4 Marginal estimation: unsupervised estimation

For the marginal estimator, we have to show that the unsupervised estimation of the hyperparameters (i.e., when the hyperparameters are estimated jointly with the aberrations) gives good aberration estimates. To this end, we compare the quality of the aberrations reconstruction obtained either by minimizing $L_{\text{ML}}(\boldsymbol{a})$ with the true hyperparameters or by minimizing $L_{\text{ML}}(\boldsymbol{a}, \mu, v_o, p)$, for several image sizes. From Figure 9, we see that for low noise levels, the unsupervised restoration is very good (the maximum difference is less than 5%). For 128×128 pixels, it is quite good (the maximum difference is less than 15%) for any noise level. Only for 32×32 pixels and high noise levels is the reconstruction seriously degraded because of the lack of information contained in the noisy data.

4.5 Performance comparison

We present the performance comparison of the joint and the marginal estimators for phase estimation. In order to compare these estimators in a realistic way, we use the joint estimator with a near null regularization (which gives good results for the estimation of the aberrations as seen in Subsection 4.3) and the unsupervised marginal estimator, described in Subsection 3.2.6. We compare the RMSE of the phase estimates as a function of noise level for two image sizes (32×32 pixels and 128×128 pixels). The results for the two estimators are plotted in Figure 10. We can see two different domains: when the signal to noise ratio (SNR) is high (noise level$< 5\%$), the two estimators approximately give the same results. At lower SNR (5% <noise level$< 20\%$), marginal estimation is significantly better. Note that this result has been checked on experimental data. The performance comparison depends on the studied object. A comprehensive comparison of the two methods has been done in Blanc (2002). This study has shown that the marginal method leads to better phase estimates for high noise levels. For low noise levels,

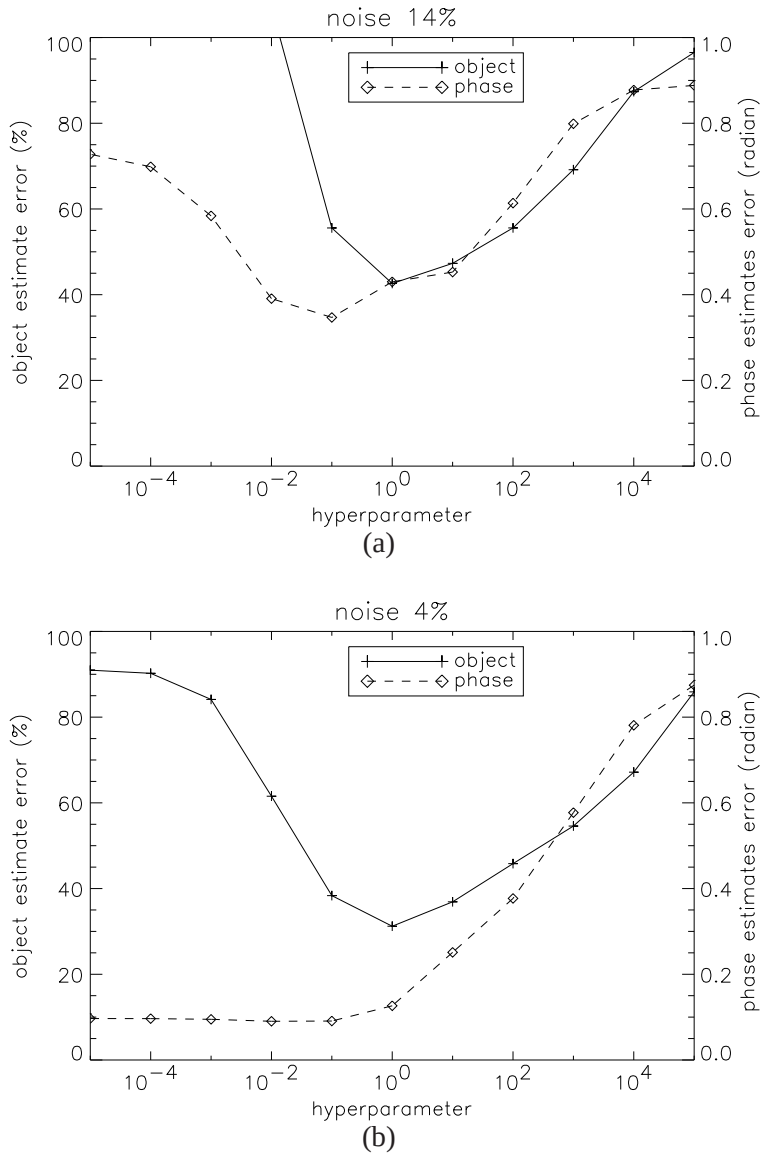


Figure 7: Plots of RMSE for joint phase estimates (dashed line—see right vertical axis) and joint object estimate (solid line - see left vertical axis) as a function of the value of the hyperparameter μ for an image size of 32×32 pixels. Figure (a) is for a noise level of 14%, Figure (b) is for 4%.

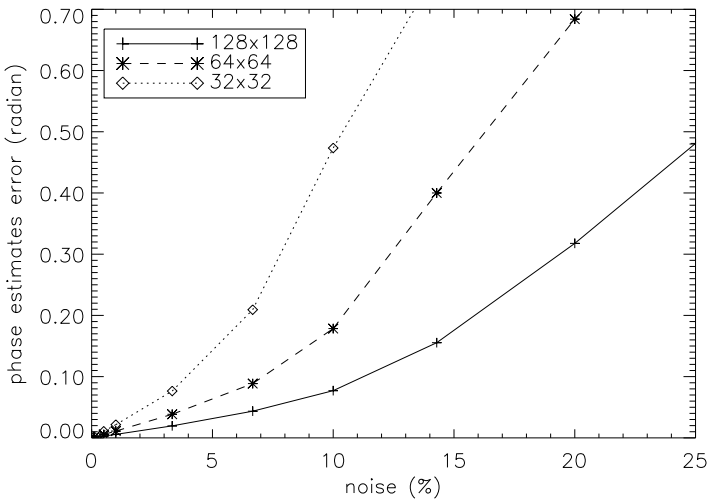


Figure 8: Performance of the joint estimator: RMSE of the phase estimates as a function of noise level for a near null regularization, for three image sizes.

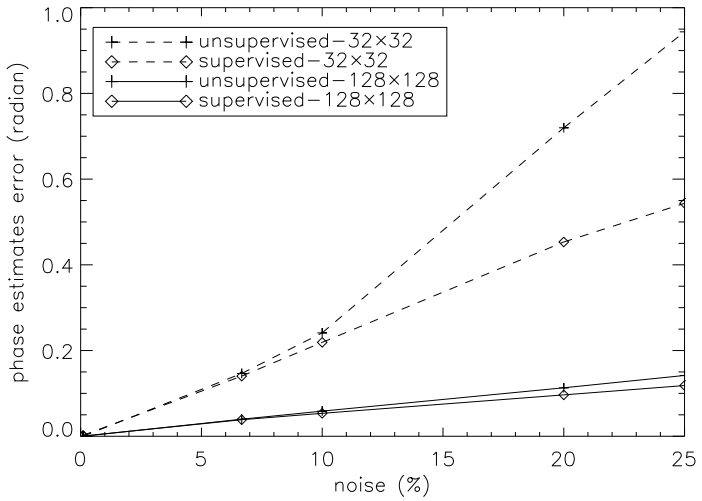


Figure 9: Performance of the marginal estimator with the true hyperparamaters (plus signs) and with an unsupervised estimation (diamonds), as measured by the RMSE of the phase estimates as a function of the noise level.

the joint estimator performs sometimes slightly better than the marginal one, the significance of the difference between the two phase estimates depending on the observed object. These performance differences between several observed objects are probably due to the Gaussian hypothesis on the object used in the marginal approach—see Subsection 3.2.

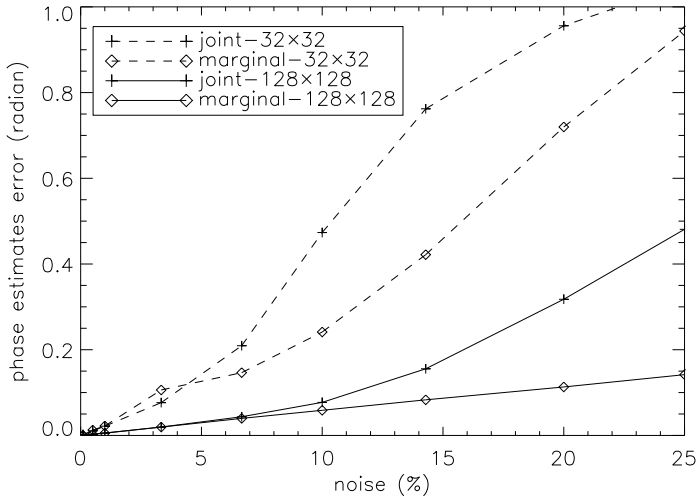


Figure 10: RMSE of the aberrations estimates for the unsupervised marginal estimator (diamonds), and for the joint one with nearly no regularization (plus signs), as a function of the noise level.

4.6 Conclusion

In these two latter sections, we have presented two methods for estimating the aberrations. The conventional estimator found in the literature which is interpretable as a joint Maximum A Posteriori approach, is based on a joint estimation of the aberrated phase and the observed object. It has been shown, by means of simulations, that it has bad asymptotic properties unless it is under-regularized and that it does not allow an optimal joint estimation of the object and the aberrated phase. The joint estimator without an object regularization can nevertheless be successfully used to estimate the aberrations, thanks to the fact that the Generalized Maximum Likelihood is a consistent phase estimator. The marginal estimator which estimates the sole phase by Maximum A Posteriori, is obtained by integrating the observed object out of the problem. This reduces drastically the number of unknowns, allows the unsupervised estimation of the regularization parameters and provides better asymptotic properties than the joint approach. Finally, the comparison of the quality of the phase restoration has shown that the marginal method leads to better phase estimates for high noise levels and that the two estimators give quite

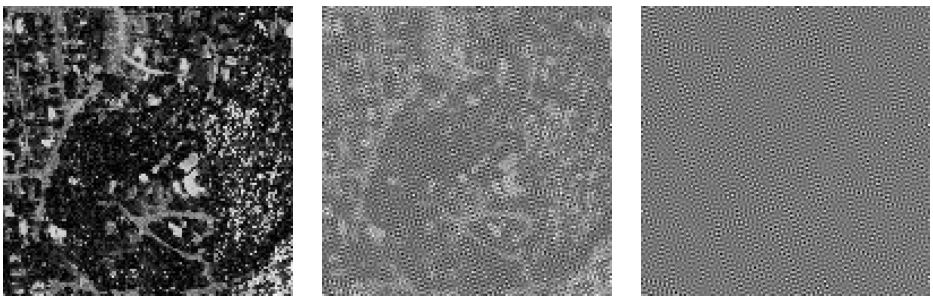
similar performance (sometimes slightly better for the joint one depending on the observed object) for low noise levels.

5 Restoration of the object

When phase diversity is used as a wave-front sensor, the estimation of the aberrations is the unique goal. On the contrary, in the case of image restoration, the object is the parameter of interest. Let us see, in this section, how the object can be estimated by the joint and by the marginal approaches.

5.1 With the joint method

In the case of the joint method, the object is estimated jointly with the aberrations by JMAP. Section 3.1.3 has shown that the hyperparameters of the object must be adjusted empirically by the user which is not easy especially with extended object. Furthermore, we have seen Subsection 3.1.3 that the joint method does not allow an optimally joint estimation of the object and of the aberrations. In particular, a near-null regularization is favorable for a good estimation of the aberrations. Let us now see the object obtained with no regularization. Figure 11 shows the results of the restoration of the object for two noise levels 1% and 14 % and an image size of 64×64 pixels (the first image is the true object). The image simulations are obtained in the same conditions as in Section 4.1. As we can see, when there is no regularization on the object, at the minimum of the joint criterion, the object estimate is completely buried in noise. In order to obtain a good object estimate, one needs to regularize the object estimation, which has been done either by incorporating an explicit regularization term into the criterion (see Section 3.1.1) or by interrupting the iterative minimization before the convergence (Strand, 1974, Thelen et al., 1999b). Yet, as mentioned above, this is prejudicial to the phase estimation and thus does not lead to the best object estimate. Additionally, it leads to the problem of hyperparameter tuning.



(a) (b) (c)

Figure 11: Joint estimation: (a) True object. (b) and (c) object restored by the joint method (with near null regularization used). The noise level is equal to 1% for (b). In this case, the RMSE is equal to 115%. Restoration (c) is for a noise level of 14%, the RMSE is 1500%.

5.2 With the marginal method

5.2.1 Principle

The marginal estimator has been obtained by integrating the observed object out of the problem. It is based on a Maximum A Posteriori approach and restores the sole aberrations. The previous section has shown that this estimator has good asymptotic properties and allows the unsupervised estimation of the noise variance and of the regularization parameters of the object. Furthermore, the marginal method provides a simple way to restore the object. The idea is to calculate $\hat{\boldsymbol{o}}$ once the aberrations $\hat{\boldsymbol{a}}_{\text{marg}}$ and the hyperparameters $\hat{\boldsymbol{\theta}}_{\text{marg}}$ are estimated by the marginal method. In particular, this estimation can be done by MAP :

$$\begin{aligned}\hat{\boldsymbol{o}}_{\text{MAP}}(\hat{\boldsymbol{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}}) &= \arg\max_{\boldsymbol{o}} f(\boldsymbol{o}|\boldsymbol{i}_1, \boldsymbol{i}_2, \hat{\boldsymbol{a}}_{\text{marg}}; \hat{\boldsymbol{\theta}}_{\text{marg}}) \\ &= \arg\max_{\boldsymbol{o}} f(\boldsymbol{i}_1, \boldsymbol{i}_2, \hat{\boldsymbol{a}}_{\text{marg}}|\boldsymbol{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) f(\boldsymbol{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) \end{aligned} \tag{40}$$

Maximizing $f(\boldsymbol{i}_1, \boldsymbol{i}_2, \hat{\boldsymbol{a}}_{\text{marg}}|\boldsymbol{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) f(\boldsymbol{o}; \hat{\boldsymbol{\theta}}_{\text{marg}})$ is thus equivalent to minimizing the following criterion:

$$L_{\text{MAP}}(\boldsymbol{o}) \propto \frac{||\boldsymbol{i}_1 - H_1(\hat{\boldsymbol{a}}_{\text{marg}})\boldsymbol{o}||^2}{\sigma^2} + \frac{||\boldsymbol{i}_2 - H_2(\hat{\boldsymbol{a}}_{\text{marg}})\boldsymbol{o}||^2}{\sigma^2} + (\boldsymbol{o} - \boldsymbol{o}_m)^t R_o^{-1} (\boldsymbol{o} - \boldsymbol{o}_m) \tag{41}$$

The cancellation of the derivative of this criterion with respect to $\boldsymbol{o}$ gives the following closed-form expression:

$$\hat{\boldsymbol{o}}_{\text{MAP}}(\hat{\boldsymbol{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}}) = R^{-1}(H_1^t \boldsymbol{i}_1 + H_2^t \boldsymbol{i}_2 + \sigma^2 R_o^{-1} \boldsymbol{o}_m), \tag{42}$$

where $R = (H_1^t H_1 + H_2^t H_2 + \sigma^2 R_o^{-1})$. The object $\hat{\boldsymbol{o}}_{\text{MAP}}$ is the bi-frame Wiener filter associated with the aberration and hyperparameter estimates $\hat{\boldsymbol{a}}_{\text{marg}}$ and $\hat{\boldsymbol{\theta}}_{\text{marg}}$. Note that the restoration is given by a Wiener filter as with the joint estimator. But the two methods differ in their tuning of the regularization parameters and in their estimation of the aberrations.

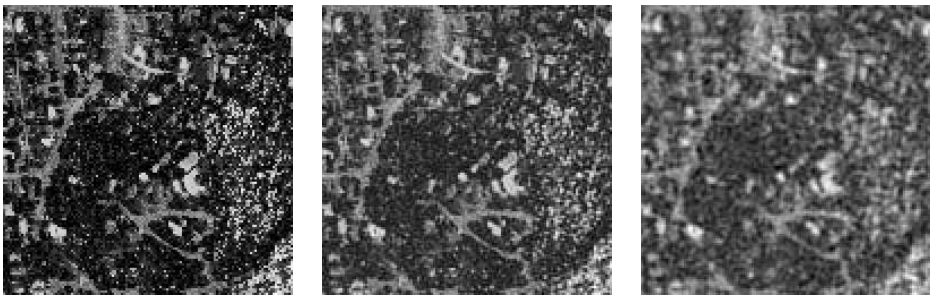
In the case of a compact object, the criterion $L_{\text{MAP}}(\boldsymbol{o})$ and the closed-form expression of $\hat{\boldsymbol{o}}_{\text{MAP}}$ can be written in the Fourier domain. For objects more extended than the image, the computation of the criterion requires the introduction of a guard band and the minimization is done iteratively (see Section 3.3.2 for details).

To summarize, the object is restored in two steps: first the phase and the hyperparameters linked to the noise and the object ($\boldsymbol{\theta}_n$,$\boldsymbol{\theta}_o$) are estimated. Then the object is restored by a MAP approach. This allows one to benefit from the good properties of this estimator: $\hat{\boldsymbol{o}}_{\text{MAP}}(\hat{\boldsymbol{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}})$ converges towards $\hat{\boldsymbol{o}}_{\text{MAP}}(\hat{\boldsymbol{a}}_{\text{true}}, \hat{\boldsymbol{\theta}}_{\text{true}})$ as the size of the data increases.

Additionally, although the restoration of the object is formally done in a second step, in practice, it is in fact performed at each computation of the marginal criterion (see Equations 34 and 25), so it is available at convergence.

5.2.2 Results

Figure 12 presents the restoration of the object obtained by this approach (same conditions as for the joint restoration). The RMSE on the object is equal to 22% for the noise level of 1% and 38% for a noise level of 14%. The quality of the object estimates provided by this method is satisfactory. Thus this method provides a robust and easy (although indirect) MAP estimation of the object.



(a) (b) (c)

Figure 12: Marginal estimation: (a) True object. (b) and (c) object restored with the marginal method. The noise level is equal to 1% for (b). In this case, the RMSE is equal to 22%. Restoration (c) is for a noise level of 14%, the RMSE is 38%.

5.2.3 Influence of the hyperparameters

Section 4.3 has shown that the JMAP estimation of the object and the phase does not allow one to optimally reconstruct the two quantities. We now study the influence of the global hyperparameter on the RMSE of marginal estimates. By contrast with the joint estimator, Figure 13 indicates that here there is a unique optimal hyperparameter for both the object and the aberrations. We have drawn the curves only for the 4% noise level because the general behavior of this estimator is the same for all noise levels. The marginal method is thus able to optimally restore the aberrations *and* the object.

5.3 With a "hybrid" method

5.3.1 Principle

For low noise levels, the joint method with no regularization on the object provides a very degraded object estimate but a good restoration of the aberrations (sometimes even better than the marginal method—see Section 4.5). Additionally, the estimation of the aberrations by the joint method requires less computation time than the restoration by the marginal method because the latter requires hyperparameter estimation. We thus propose to use the joint method to estimate the aberrations, then to estimate the sole hyperparameters for these aberrations by the

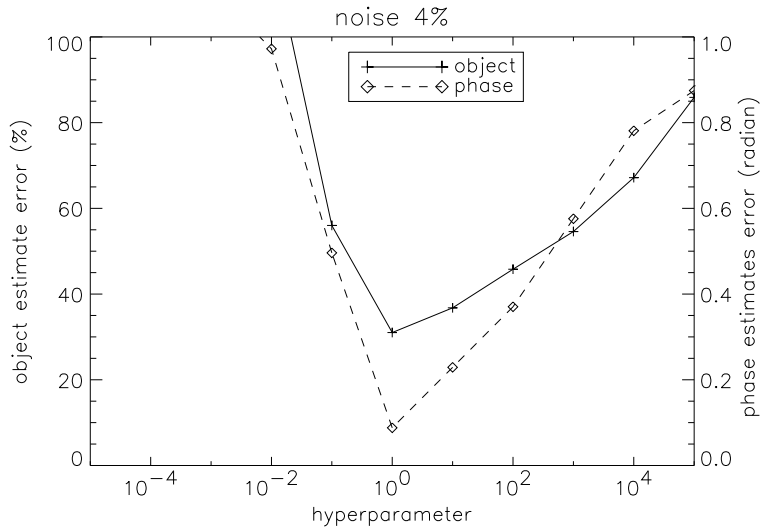


Figure 13: Plots of the RMSE for the marginal phase estimate (dashed line, see right vertical axis) and for the marginal object estimate (solid line, see left vertical axis), as a function of the value of the hyperparameter μ, for an image size of 32×32 pixels and a noise level of 4%.

marginal method. This step is fast since only few parameters remain to be estimated. At this point, we are in the same conditions as in the case of the marginal estimation: the aberrations and the hyperparameters are known. Then we estimate the object by a MAP approach. This method is called the "hybrid" method and consists in the three steps detailed below.

5.3.2 The three steps

- Estimation of the aberrations
  The estimate of the aberrations is obtained jointly with the object by a Generalized Maximum Likelihood (GML) approach, i.e., without a regularization of the object (see Section 3.1.1 and 4.2):

$$(\hat{\boldsymbol{o}}, \hat{\boldsymbol{a}})_{\text{GML}} = \arg\max_{\boldsymbol{o},\boldsymbol{a}} f(\boldsymbol{i}_1, \boldsymbol{i}_2 | \boldsymbol{a}; \boldsymbol{o}, \theta) f(\boldsymbol{a}; \theta) \tag{43}$$

  From this estimation, the sole aberrations $\hat{\boldsymbol{a}}_{\text{GML}}$ are kept.

- Estimation of the hyperparameters
  The estimation of the hyperparameters relative to the noise and to the object is done by a ML approach. The joint probability density function of the object $\boldsymbol{o}$, the images $\boldsymbol{i}_1$ and $\boldsymbol{i}_2$ and the aberrations $\boldsymbol{a}$ is marginalized with respect to the object $\boldsymbol{o}$, as done for the marginal estimator (except that here

the aberrations $\boldsymbol{a}$ are fixed to their values estimated by GML):

$$\hat{\theta} = \arg\max_{\theta} \int f(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{o}, \hat{\boldsymbol{a}}_{\text{GML}}; \theta) \mathrm{d}\mathbf{o} = \arg\max_{\theta} f(\boldsymbol{i}_1, \boldsymbol{i}_2, \hat{\boldsymbol{a}}_{\text{GML}}; \theta) \quad (44)$$

The expression of the associated criterion as well as its implementation and computation are identical to the ones used in Section 3.2.6 for the marginal estimation of the hyperparameters and the aberrations. This step is very fast due to the fact that the minimization is done with fixed aberrations and only the (three) hyperparameters are estimated.

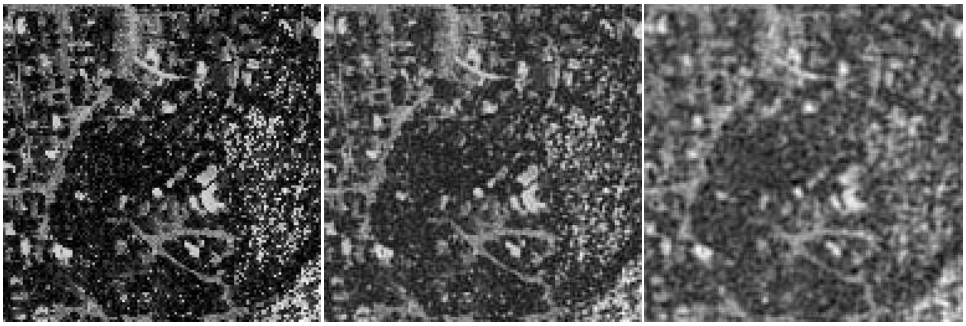
- Estimation of the MAP object
  From the aberrations $\hat{\boldsymbol{a}}_{\text{GML}}$ and the hyperparameters $\hat{\theta}$ estimated in the two previous steps, the object is restored by a MAP approach:

$$\hat{\mathbf{o}}_{\text{MAP}}(\hat{\boldsymbol{a}}_{\text{GML}}, \hat{\theta}) = \arg\max_{\boldsymbol{o}} f(\boldsymbol{i}_1, \boldsymbol{i}_2, \hat{\boldsymbol{a}}_{\text{GML}} | \boldsymbol{o}; \hat{\theta}) f(\boldsymbol{o}; \hat{\theta}) \quad (45)$$

  This step is identical to the restoration of the object from the marginal estimator (see the previous subsection) except that here this step must be explicitly computed.

5.3.3 Results

We present the results of the object estimation obtained by the hybrid method in Figure 14 (the simulation conditions are always the same). The RMSE on the



(a) (b) (c)

Figure 14: Hybrid estimation: (a) True object. (b) and (c) restored object with the marginal method. The noise level is equal to 1% for (b). In this case, the RMSE is equal to 22%. Restoration (c) is for a noise level of 14%, the RMSE is 41%.

object is 22% for a noise of 1% and 41% for a noise level of 14 %. The object estimates obtained by the joint method with no regularization (Section 5.1) does not bear comparison with these estimates: here the quality of the restoration is close to the one provided by the marginal method. For the noise level of 1%, the hybrid method and the marginal one give nearly identical object estimates.

For the lower signal-to-noise ratio, the marginal method performs slightly better. This is not surprising: for high noise level, the quality of the aberration estimates obtained by the marginal method is better than the one given by the joint method (see Section 4.5).

5.4 Conclusion

We have shown that the joint method does not allow one to obtain good object estimates. For high noise levels, the marginal method provides a simple and robust way to estimate the object. For lower noise level, we propose to use the good phase estimate obtained by the joint method, in a novel hybrid method, leading to a fast and satisfactory estimation of the object.

6 Optimization methods

6.1 Projection-based methods

As mentioned in Section 1, phase diversity techniques originate from the problem of phase retrieval, where the observed object $\boldsymbol{o}$ is a point source, and only one focused image $\boldsymbol{i}$ is available. In such conditions, it is clear from Equations (1) and (11) that the observed image reads $\boldsymbol{i} = \left|\mathrm{FT}^{-1}\left(P\, e^{j\phi}\right)\right|^2 + \boldsymbol{n}$. Neglecting the noise component, and assuming that P is known (as the indicatrix of the pupil), the issue of recovering the pupil phase ϕ can be restated in a more basic way: how to recover the phase ϕ of a function of known modulus P, given the modulus $\rho = \boldsymbol{i}^{1/2}$ of its Fourier transform?

The most common approach for solving this problem is to alternately enforce the known moduli in the two domains, which leads to the Gerchberg-Saxton algorithm (Gerchberg and Saxton (1972); see also Gerchberg (1974), Papoulis (1975) and Fienup (1982)):

1. given a current value of ϕ, compute $\boldsymbol{g}_0 = \mathrm{FT}^{-1}\left(P\, e^{j\phi}\right)$
2. replace the modulus of $\boldsymbol{g}_0$ by ρ, i.e., compute $\boldsymbol{g} = \rho\, \boldsymbol{g}_0 / \left|\boldsymbol{g}_0\right|$
3. compute $\boldsymbol{G} = \mathrm{FT}\left(\boldsymbol{g}\right)$
4. take the phase of $\boldsymbol{G}$ as the new current value of ϕ, and go back to Step 1.

The enforcement of the constraints can be mathematically interpreted as projections onto two subsets S_1 and S_2 of the set of phase vectors ϕ. At convergence, the aim is to find a phase vector that belongs to $S_1 \cap S_2$.

Variants of the same alternating procedure are also encountered, for instance to tackle the problem where P is an unknown non-negative function of known support (Fienup, 1982, Bauschke et al., 2002).

In the simplest phase diversity problem, a couple of images *of a point source* is available according to Equations (12) and (13). Adapted versions of the Gerchberg-Saxton algorithm have been proposed to cope with this situation (Misell, 1973, Baba and Mutoh, 2001).

In the general phase diversity problem, the object is unknown, and the Gerchberg-Saxton algorithm does not easily generalize (see Baba and Mutoh (1994) for an attempt). Instead, optimization procedures are more commonly found in the literature of phase diversity, either to minimize the plain least-square criterion of Equation (14), or a penalized version of it, e.g., Equation (21). In the authors' mind, least-square minimization techniques should not be considered as stopgap solutions when the Gerchberg-Saxton algorithm is not readily usable. On the contrary, for several reasons, adopting the optimization framework should rather be seen as highly recommendable to solve phase diversity (or even phase retrieval) problems:

- The Gerchberg-Saxton algorithm will not converge if $S_1 \cap S_2$ is an empty set, i.e., if the data-set is not *feasible*. This difficulty is shared by nearly all projection techniques. Obviously, the non-feasible case is far from academic. It is rather expected to correspond to practical situations, as far as noisy measurements and approximated models must be handled. In (Censor et al., 1983) and in (Byrne, 1997), modified projection techniques are introduced to cope with non-feasible, linear problems. Convergence is then shown towards a least-square solution.

- In (Fienup, 1982), the Gerchberg-Saxton algorithm is tested against a least-square approach based on a gradient-search method, on a phase retrieval problem. The former approach is reported to converge very slowly compared to the latter.

Finally, it is the authors' view that much clarity is gained when an objective function (such as a least-square criterion) is explicitly defined, first and foremost. Only then, an appropriate algorithm is to be selected among different families of optimization schemes, one of which being based on successive projections. In the remaining part of the present Section, two other families are introduced: line-search methods, and trust-region schemes. For sake of simplicity, we shall mainly restrict our attention to the minimization of $L = L'_{\text{JMAP}}$ defined by Equation (23) (or indifferently to $L = L_{\text{MAP}}$ defined by Equation (33)), as a function of aberration parameters $\boldsymbol{a}$.

6.2 Line-search methods

Line-search minimization is by far the most commonly adopted framework in the context of phase diversity, see, e.g., Gonsalves (1982), Fienup (1982), Paxman et al. (1992), Blanc et al. (2003b).

Each iteration of a line-search method is twofold: firstly, a search direction $\boldsymbol{p}_\ell$ is computed. Then, a stepsize α_ℓ is determined, which corresponds to how far to move along $\boldsymbol{p}_\ell$. The resulting iteration reads:

$$\boldsymbol{a}_{\ell+1} = \boldsymbol{a}_\ell + \alpha_\ell \boldsymbol{p}_\ell$$

Important particular cases are obtained according to specific choices of the search direction.

6.2.1 Strategies of search direction

A *steepest descent method* is obtained when $\boldsymbol{p}_\ell$ is chosen opposite to the criterion gradient $\boldsymbol{g}_\ell = \nabla L(\alpha_\ell)$. Such a method is computationally simple but rather slow to converge.

Conjugate gradient methods form a very useful extension of steepest descent, since they usually converge much faster, for almost the same low computational requirement. They were originally designed to minimize convex quadratic functions,

but they also provide an efficient and popular approach to general optimization problems. According to conjugate gradient methods, each direction $\boldsymbol{p}_\ell$ is built as a linear combination of $-\boldsymbol{g}_\ell$ and the previous direction $\boldsymbol{p}_{\ell-1}$:

$$\boldsymbol{p}_\ell = -\boldsymbol{g}_\ell + \beta_\ell \boldsymbol{p}_{\ell-1}$$

where β_ℓ is chosen to generate mutually conjugated search directions when the criterion is convex quadratic. For instance, the Fletcher-Reeves choice for the conjugate directions corresponds to

$$\beta_\ell = \frac{\boldsymbol{g}_\ell^t \boldsymbol{g}_\ell}{\boldsymbol{g}_{\ell-1}^t \boldsymbol{g}_{\ell-1}}$$

In the phase diversity context, conjugate gradient methods are proposed in Gonsalves (1982), Fienup (1982), Paxman et al. (1992), Blanc et al. (2003b).

$\boldsymbol{p}_\ell = -H_\ell^{-1}\boldsymbol{g}_\ell$ yields *Newton's method* when H_ℓ is the Hessian $\nabla^2 L(\alpha_\ell)$. The asymptotic convergence of Newton's method is fast, but its behavior far from the solution is hardly predictable, especially when L is not a convex criterion. Moreover, it is more computationally demanding than the previous methods, since each search direction $\boldsymbol{p}_\ell$ is the solution of a linear system.

Quasi-Newton methods correspond to

$$\boldsymbol{p}_\ell = -B_\ell^{-1}\boldsymbol{g}_\ell \tag{46}$$

where B_ℓ is "well-chosen". Typically, matrix B_ℓ results from a trade-off: (i) it is close to H_ℓ for large values of ℓ, (ii) it is symmetric and structurally positive definite, even for the first iterations (contrarily to the Hessian H_ℓ), (iii) it is more easily invertible than H_ℓ. The most popular quasi-Newton algorithm is the BFGS method (for Broyden-Fletcher-Goldfarb-Shanno), where B_ℓ^{-1} is maintained and updated, rather than matrix B_ℓ itself. For large scale problems, it is still too demanding to handle matrices B_ℓ^{-1}. Limited-memory BFGS directly approximates $B_\ell^{-1}\boldsymbol{g}_\ell$, using information gathered from the m earlier iterations, where m is a usually a small number. For more details on BFGS and on limited-memory BFGS, see Nocedal and Wright (1999). Vogel et al. (1998) reports that the performance of BFGS applied to the minimization of $L = L'_{\text{JMAP}}$ is disappointing; Vogel (2000) resorts to limited-memory BFGS.

6.2.2 Step size rules

Choosing a clever strategy of search direction is not a sufficient condition to obtain an efficient minimization algorithm. Modern analyses of convergence combine both the search direction and the choice of the step size α_ℓ. The presumably "ideal" step size rule $\alpha_\ell = \arg\min_\alpha L(\boldsymbol{a}_\ell + \alpha_\ell \boldsymbol{p}_\ell)$ usually has no closed-form expression. We are thus led to employ so-called *inexact* step size strategies (i.e., inexact line search strategies (Moré and Thuente, 1994). To find even a local minimizer of $f(\alpha) = L(\boldsymbol{a}_\ell + \alpha_\ell \boldsymbol{p}_\ell)$ needs a certain amount of iterations of a uni-variate

minimization method, depending on the required precision. In this respect, a key result is brought by *sufficient decrease conditions* such as Wolfe's (Nocedal and Wright, 1999): inexact step size rules that satisfy such conditions, in conjunction with appropriate search directions, form the ingredients of algorithms that are both implementable and converging.

6.3 Trust-region methods

Trust-region methods are based on a model function M_ℓ whose behavior near the current point $\boldsymbol{a}_\ell$ is similar to that of the actual criterion L. The model M_ℓ is usually chosen convex quadratic. Then, the trust-region approach correspond to the following strategy:

1. Solve
$$\boldsymbol{p}^r_{\ell+1} = \arg\min_{\boldsymbol{p}} M_\ell(\boldsymbol{a}_\ell + \boldsymbol{p}) \tag{47}$$
where $\boldsymbol{a}_\ell + \boldsymbol{p}$ is restricted to belong to a *trust region*. The latter is typically chosen as a ball of radius, say r.

2. If the candidate solution $\boldsymbol{p}^r_{\ell+1}$ does not produce a sufficient decrease $L(\boldsymbol{a}_\ell) - L(\boldsymbol{a}_\ell + \boldsymbol{p}^r_{\ell+1})$, then the trust region is considered as too large: the radius r is decreased, and the minimization step is re-solved, as many times as necessary.

In practice, it may be computationally expensive to solve Equation (47). Actually, it suffices to compute inexact solutions, provided that a condition of sufficient decrease still holds.

The Levenberg-Marquardt method is often presented as a method of choice to solve non-linear least-square problems (Press et al., 1992). It is based on updates of the form of Equation (46), with $B_\ell = J^t_\ell J_\ell + \lambda_\ell I$, where

- $J^t_\ell J_\ell$ is a positive semi-definite approximation of the Hessian,
- I is the identity matrix,
- λ_ℓ is a parameter that must be adjusted along the iterations to ensure the sufficient decrease of L.

The Levenberg-Marquardt can be viewed as a pioneering trust region method, where the trust region is a ball of radius λ_ℓ, and $M_\ell(\boldsymbol{p}) = \|J_\ell \boldsymbol{p} + \boldsymbol{g}_\ell\|^2$ (Nocedal and Wright, 1999).

In the context of phase diversity, Löfdahl and Scharmer (1994) proposes a simplified form of the Levenberg-Marquardt method, where $\lambda_\ell = 0$. Such a choice does not ensure the theoretical convergence of $\{\boldsymbol{a}_\ell\}$ for arbitrary initial points. Vogel et al. (1998) also resorts to a variant of the Levenberg-Marquardt method, and Luke et al. (2000) applies limited memory BFGS with trust regions.

7 Application of phase diversity to an operational system: calibration of NAOS-CONICA

7.1 Practical implementation of phase diversity

7.1.1 Choice of the defocus distance

An appropriate choice of the defocus distance d is essential to optimize the performance of the phase diversity sensor. If it is too small, the images are almost identical (not enough diversity), which brings back the problems associated with phase retrieval. If it is too large, the defocused image contains virtually no information.

The RMS defocus coefficient a_4^d (such that $\phi_d(\boldsymbol{r}) = a_4^d Z_4(\boldsymbol{r})$, see Equation (13)) depends on d, on the wavelength λ, the telescope diameter D, and the focal length F through:

$$a_4^d = \frac{\pi d}{8\sqrt{3}\lambda(F/D)^2} (\text{in radian}) \tag{48}$$

$\phi_d(\boldsymbol{r})$ is quadratic, minimum for $\boldsymbol{r} = 0$ and maximum for $\boldsymbol{r} = 1$. The corresponding peak-to-valley optical path difference Δ is equal to

$$\Delta = \frac{\lambda}{2\pi} \times 2\sqrt{3}a_4^d = \frac{\sqrt{3}\lambda a_4^d}{\pi} = \frac{d}{8(F/D)^2} \tag{49}$$

Some studies (Lee et al., 1999, Meynadier et al., 1999) have shown that choosing d such that Δ is approximately equal to λ provides accurate results. In fact the "optimal" defocus distance depends on the object structure, on the amplitude of the aberrations and on the SNR of the images. In practice a large domain around this value (typically, $\Delta \approx \lambda \pm \lambda/2$) still provides good results. This result has been first obtained experimentally (Meynadier et al., 1999). It has been later confirmed in a simplified theoretical framework, using the expression of the Cramer-Rao lower bound for the variance of unbiased estimators (Lee et al., 1999, Prasad, 2004).

7.1.2 Image centering

The estimation of aberrations from experimental data requires the determination of two parameters in addition to those describing the aberrated phase (Löfdahl and Scharmer, 1994). These parameters correspond to the relative alignment in x and y between the focused and the defocused images. They can be described by tip-tilt coefficients a_2 and a_3 in the Zernike basis and can be estimated in the same way as those corresponding to the wave-front, except that they relate only to the defocused image. This estimation of the alignment parameters eliminates the need for sub-pixel alignment of the images. But the images still need to be first re-centered to an accuracy of about one or two pixels, for example by cross-correlation, because phases larger than $\pm\pi$ are difficult to estimate by phase diversity and often induce local minima (see Section 8).

7.1.3 Spectral bandwidth

The phase diversity concept, as described so far, is a monochromatic WFS. Nevertheless, Meynadier et al. (1999) has shown that the use of broadband filters does not significantly degrade the accuracy as long as $\Delta\lambda/\lambda$ is reasonable (typically $\Delta\lambda/\lambda \leq 0.15$). Note that the concept can be extended to polychromatic images through the appropriate change in the data model (Seldin and Paxman, 2000).

7.2 Calibration of NAOS and CONICA static aberrations

7.2.1 The instrument

The European Very Large Telescope (VLT) instrument NAOS-CONICA is composed of the adaptive optics (AO) (Roddier, 1999) system NAOS (Rousset et al., 1998) and of the high resolution camera CONICA (Lenzen et al., 1998). It is aimed at providing very high quality images on one (UT4-Yepun) of the 8-m telescopes of the European Cerro Paranal observatory. Figure 15 presents the simplified outline of such an AO instrument. It consists of two different optical paths separated by a beam-splitter: the imaging path and the AO path (having in common the so-called "common path"). The imaging path is composed of the different filters and objectives of CONICA. The AO path consists of the WFS which estimates the optical phase delays and of the real-time computer, which applies the appropriate corrections via the deformable mirror (DM). In order to reach the ultimate performance

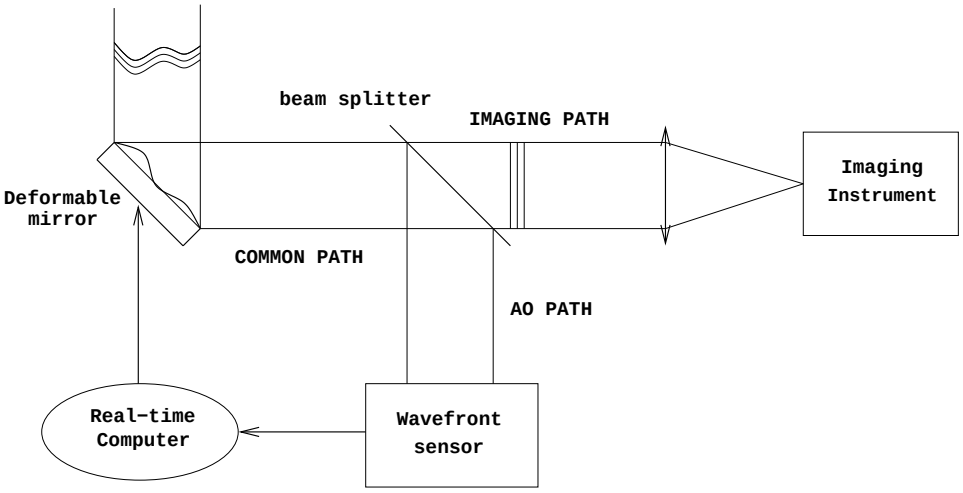


Figure 15: Simplified outline of the VLT instrument NAOS-CONICA.

of the instrument, it has been necessary to calibrate the remaining aberrations induced by its optical components. Defects of the wave-front originating from any component within the AO loop (common and AO path) are seen by the AO WFS and thus corrected. This is not the case for a degradation of image quality induced

by components outside the AO loop, in the imaging path (i.e., the beam splitters and filters and objectives of the camera). Phase diversity has been used to calibrate these unseen aberrations. More details about the calibration procedure can be found in Blanc et al. (2003a) and Hartung et al. (2003). Because of the huge number of observation modes of NAOS-CONICA, the calibration has been split into several parts: NAOS' dichroics, i.e., the beam splitters, and CONICA, i.e., the different filters and objectives. The goal is to be able to assign the aberration contributions to the various optical components. Two different procedures have been used to obtain the phase diversity data, i.e., the focused and defocused images. The first procedure has provided the calibration of the optical path of CONICA, the second one, the aberrations of NAOS' dichroics.

7.2.2 Calibration of CONICA stand-alone

Phase diversity setup The estimation of CONICA stand-alone aberrations (objectives and filters) is obtained through the use of pinholes located at differently defocused positions on a wheel in the camera entrance focal plane. Figure 16 depicts the setup for the corresponding measurements. CONICA has as many as 40 possible filters and 7 different camera objectives. A calibration point source is slid in front of CONICA. The telescope pupil is simulated by a cold pupil placed inside CONICA. After rotating the wheel that holds the pinholes, one records the focused and defocused images and then performs the phase diversity estimation.

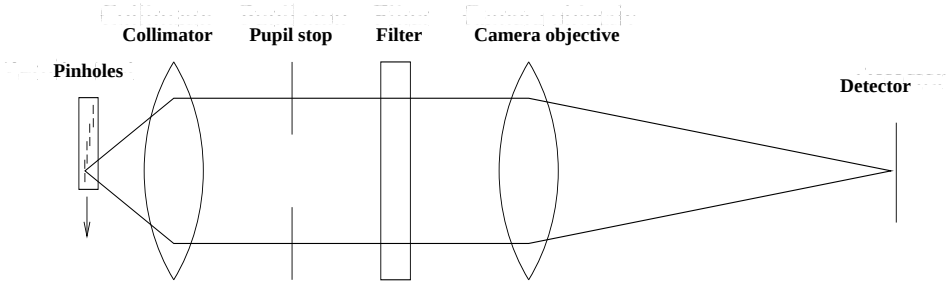


Figure 16: Calibration of CONICA stand-alone: use of pinholes in the entrance focal plane.

Example of CONICA aberration estimation The true value of the aberrations is of course not available in the real world. A practical way to be confident in the correctness of the estimated aberrations is to compare the recorded images to the ones reconstructed with these estimated aberrations.

Figure 17 shows an example of the comparison between experimental images and the reconstructed ones. Note that, here, the observed object is close to a Dirac

function. The phase estimation is obtained using the joint method with no regularization on the object. The phase estimate is expanded on the first 15 Zernike polynomials. The Strehl ratio (SR[7]) computed from the aberrations' estimate is equal to 87%. It compares nicely to the SR directly computed on the focal plane image which is equal to 85 %.

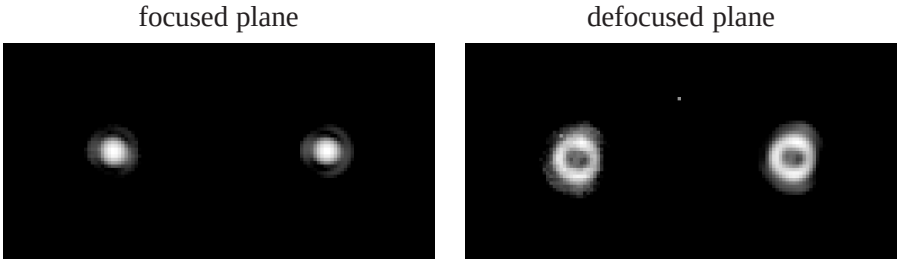


Figure 17: Comparison between measured images and PSF's reconstructed from the estimated aberrations. Left: focused plane; right: defocused plane (a log scale is adopted for each image).

Calibration of the filters and objectives The aberrations estimated from CONICA phase diversity data correspond to the contributions of a filter and a camera objective. Thanks to the fact that the camera objectives are achromatic and that there are many different filters with very different aberrations, one can separate the contributions of the filters and of the objectives. One first measures the total aberrations for a given camera and a given filter i ($1 \leq i \leq n$). This operation is repeated for all the n filters. The camera objective contribution $\boldsymbol{a}^{\text{Cam}}$ is then estimated by taking the median value of these total aberrations $\boldsymbol{a}^{\text{Ctot,filt}i}$ for all different filters:

$$\boldsymbol{a}^{\text{Cam}} = \text{median}\left(\boldsymbol{a}^{\text{Ctot,filt1}}, \boldsymbol{a}^{\text{Ctot,filt2}}, ...\boldsymbol{a}^{\text{Ctot,filt}n}\right) \tag{50}$$

Lastly, filter aberrations are obtained as the difference between the measured aberrations and the estimated camera aberrations:

$$\boldsymbol{a}^{\text{filt}i} = \boldsymbol{a}^{\text{Ctot,filt}i} - \boldsymbol{a}^{\text{Cam}} \tag{51}$$

Figure 18 shows the calibration results in the J-H band, for a specific camera objective (C50S) and several filters. The solid line corresponds to the median representing the camera objective aberrations.

[7]The SR is a common way to describe the quality of the point spread function. It is given by the ratio of the measured to the theoretical, diffraction-limited, peak intensity in the image of a point-source.

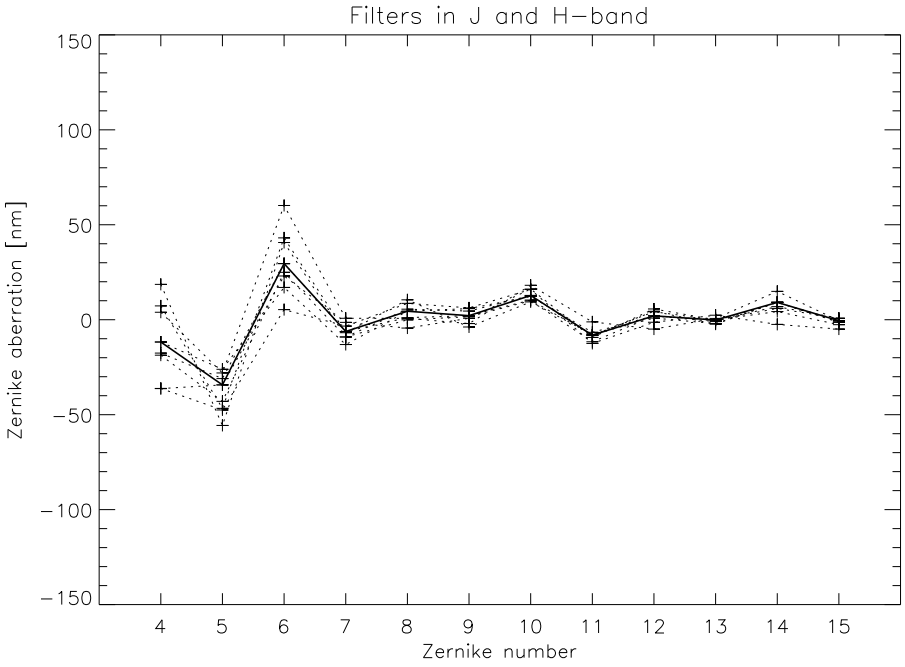


Figure 18: CONICA's aberrations measured by several filters in the J and H bands with camera objective C50S. The solid lines indicate the median value representing the aberrations of the camera objective.

7.2.3 Calibration of the NAOS dichroics

Phase diversity setup The estimation of the NAOS dichroic aberrations is obtained through the use of the AO system. Note that NAOS has five different dichroics. A focused image of a fiber source, located in the entrance focal plane of NAOS, is recorded in closed-loop in order to avoid the common-path aberrations from the optical train between the source and the dichroic. Then a known defocus is introduced on the DM with the AO loop still closed to record the defocused image. This approach gives the sum of $\boldsymbol{a}^{\mathrm{NCtot}}$, the NAOS dichroic aberrations and the CONICA aberrations.

Calibration of NAOS dichroic aberrations The contribution of the NAOS dichroics $\boldsymbol{a}^{\mathrm{dichro}}$ can be determined by subtracting the total CONICA instrument aberrations $\boldsymbol{a}^{\mathrm{Ctot}}$ (provided by the previous calibration of CONICA stand-alone) from the overall NAOS-CONICA instrument aberrations $\boldsymbol{a}^{\mathrm{NCtot}}$:

$$\boldsymbol{a}^{\mathrm{dichro}} = \boldsymbol{a}^{\mathrm{NCtot}} - \boldsymbol{a}^{\mathrm{Ctot}} \tag{52}$$

7.2.4 Closed loop compensation

Thanks to the individual calibration of the different optical components, the correction coefficients are known for any possible configuration of the instrument. The pre-compensation of these static aberrations can be done by using NAOS, which is able to introduce known static aberrations on the DM in closed loop. In this manner the DM takes the shape needed for compensation of the static optical aberrations. To demonstrate the final gain in optical quality, we compare the originally acquired images without correction for static aberrations with the images obtained after closed loop compensation. Figure 19 shows two examples of closed loop compensations in J-band and in K-band. We achieve a striking correction in J-band, visible with the naked eye between the images before and after correction. In K-band the non-corrected image is already very close to the diffraction limit and the improvement is hardly visible. However, the computation of the SR (see Figure 19) shows that even in K-band the performed correction is still significant.

7.3 Conclusion

These calibrations have shown that phase diversity is a simple and powerful approach to improve the overall optical performance of an AO system. This is of a great interest for future very high SR systems, in which an accurate estimation and correction of aberrations is essential to achieve the ultimate performance and to reach the scientific goals (for instance, regarding exo-planet detection).

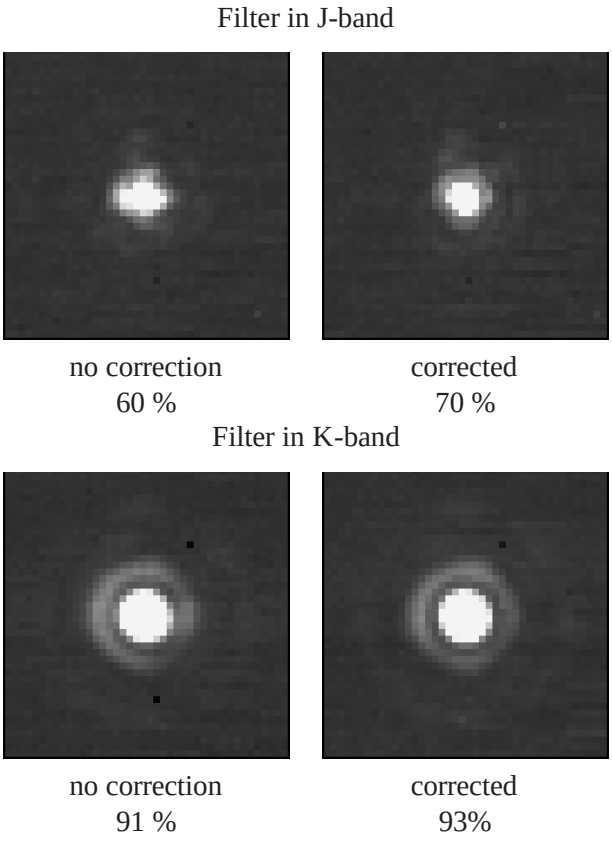


Figure 19: Comparison of the PSF's before and after closed loop compensation. The SR of each PSF is indicated in percent.

8 Emerging methods: measurement of large aberrations

8.1 Problem statement

Large-amplitude aberrations refer to the fact that variations higher than 2π can be encountered. In such a situation, the data used in phase diversity (i.e.,, focused and defocused images) do not contain the full information on the aberrated pupil phase. In fact Equation (1), recalled below, shows that the PSF is related to the pupil phase ϕ by a non-linear relationship such that the phase appears in a complex exponential:

$$h^{\text{opt}}(x,y) = \left|\text{FT}^{-1}\left(\mathcal{P}(\lambda u,\lambda v)\, e^{j\varphi(\lambda u,\lambda v)}\right)\right|^2 (x,y).$$

Any 2π-variation of any phase point in the pupil leaves the PSF unchanged and thus the image unchanged. Hence the phase diversity data only contain information on the wrapped phase $\phi[2\pi]$. The sole data (whatever their number) do not allow one to discriminate between the true continuous phase and all the equivalent phases. Figure 20 shows a 1D example of two phases which yield the same image. In the joint criterion as in the marginal one, the term corresponding to the

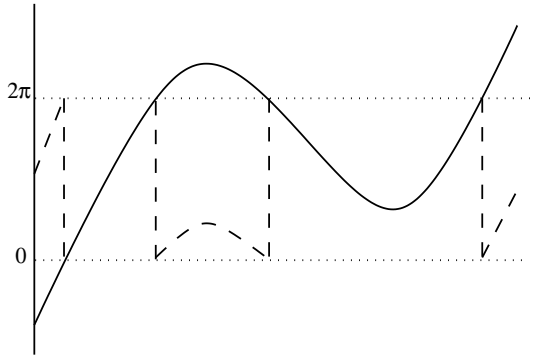


Figure 20: The pupil phase in solid line and the one in dashed line yield the same PSF.

likelihood presents an infinite number of equivalent local minima. Figure 21 illustrates this in a 1D representation: the unwrapped true phase corresponds to one of these minima but if only using the data, it is not differentiated from the others. Note that, unfortunately, the defocused image which removes the indetermination of the sign of the even part of the pupil phase (see Subsection 1.3.1), has no effect on this indetermination. In practice, this indetermination appears for any strongly degraded system (i.e., for phase amplitudes greater than 2π). It is the case, for example, for astronomical observations from the grounds where images are strongly degraded by the atmospheric turbulence. This indetermination does not affect *a posteriori* correction. Indeed, for this application, the object is the parameter of

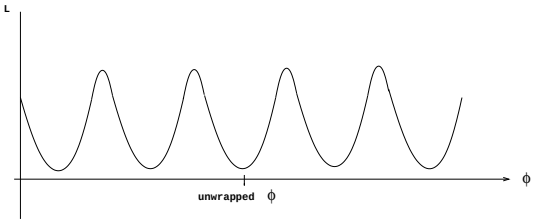


Figure 21: Symbolic 1D representation of the likelihood term

interest thus it does not bear the indetermination provided that the wrapped phase $\phi[2\pi]$ is correctly estimated. By contrast, real-time wave-front correction requires the estimation of the unwrapped phase [8] and thus the 2π-indetermination must be removed. Regularity criteria are usually invoked to perform *phase unwrapping*. However, the latter is a difficult task, both from the informational viewpoint (only poor information is available), and from the computational viewpoint (the criterion to optimize is highly multimodal).

8.2 Large aberration estimation methods

In the phase diversity literature, two different ways have been proposed to estimate large aberrations: the first one is based on a single step estimation of the unwrapped phase, the second one is composed by two steps: first, the wrapped phase is estimated then, if necessary, the phase is unwrapped.

8.2.1 Estimation of the unwrapped phase

This method consists in simultaneously estimating and unwrapping the phase. It requires introducing information in the estimation process in addition to those contained in the data, in order to remove the 2π-indetermination and to construct a criterion which presents (if possible) a unique global minimum corresponding to the unwrapped phase. This information can be brought by an appropriate choice of the basis function for the expansion of the phase and/or by the use of an explicit regularization term on the phase parameters. In the phase diversity literature, two different basis have been used to expand the phase.

Zernike polynomials In all previous sections, and as often done in the phase diversity literature, the phase has been expanded on the Zernike modal basis (see Subsection 1.2.4). A parameterization of the phase on an infinite number of Zernike polynomials leaves the likelihood criterion unchanged (Figure 21). But in practice, the phase is expanded on a finite number of polynomials. In this case,

[8] Especially if the scientific imaging is not performed at the same wavelength as the wave-front sensing, or if it is polychromatic.

the minima of the criterion are no more equivalent and a unique global minimum corresponds to the unwrapped phase. Even if this decomposition seems to remove the ambiguity, in practice all the minima are very close to each other (see Figure 22) and it is almost impossible to reach the global minimum (corresponding to the unwrapped phase). In order to better differentiate between the global minimum

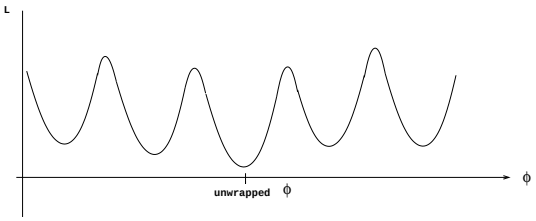


Figure 22: 1D representation of the likelihood term for the pupil phase expanded on a truncated Zernike decomposition.

and these others, it is essential to add more information on the phase by introducing a regularization term in the criterion. For imaging through turbulence, one can use the statistical knowledge on the atmospheric turbulence and introduce the *a priori* probability distribution presented in Subsection 3.1.1, Equation (19). This simultaneous estimation and unwrapping of the phase has been chosen by Thelen et al. (1999b).

Delta functions Another possible basis for the phase expansion is the delta functions: the phase is estimated point-by-point in the pupil (i.e., at each sampled point in the pupil plane). By contrast with the truncated Zernike basis, this expansion of the phase leaves the indetermination unchanged, i.e., the likelihood term keeps the shape of Figure 21. Thus, in order to reach the unwrapped phase solution, one must introduce *a priori* information on the phase to "distort" the criterion and remove the equivalence between all 2π-variation phase estimates. The solution proposed by Jefferies et al. (2002) is to parameterize the phase by a function convolved with a smoothing kernel. This is a mean to impose smoothness on the estimated phase and it has the same unwanted effect as using a truncated Zernike basis: it distorts the criterion into something like Figure 22.

To summarize, for the Zernike modal regularization approach as well as for the point-by-point one, the regularized criterion presents many local minima. The global minimum corresponding to the searched unwrapped phase is unique but very close to the other local solutions. The success of the minimization of such criteria by local minimization method (see Section 6) is very uncertain and strongly depends on the starting point of the iterative minimization (which should be very close to the unwrapped solution). Thus, the simultaneous estimation and unwrapping of the phase is very difficult. This is due to the fact that each process (estimation and unwrapping) is, by itself, a difficult problem.

8.2.2 Estimation of the wrapped phase (then unwrapping)

An alternative approach is to keep the problem separated in two steps: first, estimate the wrapped phase $\phi[2\pi]$ then, if necessary, unwrap the phase to obtain ϕ. This approach is guided by, on one hand, the above mentioned difficulties to solve the two problems simultaneously, and, on the other hand, by noting that for some applications, the estimation of the sole wrapped phase $\phi[2\pi]$ is enough (as for post-correction). In the phase diversity literature, Löfdahl et al. (1998b), Baba and Mutoh (2001) have chosen this two-step method. Their application was the estimation of static aberrations of a telescope by observing an unresolved point source (a star) through turbulence. We have also developed a phase estimation method based on this two-step approach. It is briefly presented in the following. More details can be found in Blanc (2002).

Estimation of $\phi[2\pi]$: construction of a regularized criterion In this approach, we first acknowledge the fact that it is almost impossible to find the global minimum of a criterion that has many almost equivalent minima. Thus, to estimate $\phi[2\pi]$ without the minimization problems associated with local minima, all minima corresponding to a 2π-variation of the phase must remain strictly equivalent: the criterion has to keep the form of Figure 21. Concerning the likelihood term, it requires the choice of an appropriate basis for the phase expansion. As we indicate Subsection 8.2.1, the delta functions keep the equivalence. Hence, the phase is expanded on this basis. The implementation of the point-by-point phase estimation leads to a very large parameter space (compared to a truncated Zernike basis) and thus requires the use of an explicit regularization term on the phase. In order to keep the equivalence between all local minima, this term must also remain the indetermination.

- Choice of the phase regularization function

  The goal of the *a priori* information on the phase is to ensure a good smoothing of the small gradients (i.e., corresponding to noise) but to be insensitive to the large gradients (corresponding to 2π-variations). We choose the following expression for the phase regularization function:

$$J_{regul}(\phi) = \sum_{(l,m)\in S} \left[\left| e^{j(\phi_{l-1,m}-\phi_{l,m})} - e^{j(\phi_{l,m}-\phi_{l+1,m})} \right|^2 + \left| e^{j(\phi_{l,m-1}-\phi_{l,m})} - e^{j(\phi_{l,m}-\phi_{l,m+1})} \right|^2 \right] \quad (53)$$

  The summation is done on all the pixels within the pupil (S is the pupil support). Furthermore, we impose a strict support constraint i.e., all terms $|...|^2$ that contain, at least, a pixel out of the pupil support, are suppressed.

  To ensure that the regularization function is insensitive to any 2π-variation of the phase, the *a priori* information has been imposed on the phasors i.e.,

$e^{j\phi}$. Note that, even if the regularization term involves the phasors, the estimated parameters are still the phase values on pixels. In order to keep all the indeterminations of the global criterion, the regularization function has been constructed in such a way that it is insensitive, as the data is, to a global piston and to tip-tilt (see Section 1.2.4 and 3.1.1).

A first-order expansion of Equation (53) provides insight on the effect of the regularization term on the phase. For small phase differences, Equation (53) is approximately given by:

$$J_{regul}(\phi) \simeq \sum_{(l,m)\in S} \left[|\phi_{l-1,m} - 2\phi_{l,m} + \phi_{l+1,m}|^2 + |\phi_{l,m-1} - 2\phi_{l,m} + \phi_{l,m+1}|^2 \right]$$

It corresponds to a quadratic penalization on the second order differences of the phase values.

- MAP estimation

  We propose a MAP approach for the point-by-point phase estimation:

$$\hat{\phi}_{\text{MAP}} = \arg\min_{\phi} J_{\text{MAP}}(\phi) \tag{54}$$

The expression of $\hat{\phi}_{\text{MAP}}$ is obtained by using Equation (37) where the regularization term on the Zernike coefficients $\boldsymbol{a}^t R_a^{-1} \boldsymbol{a}$ is replaced by Equation (53). Equation (37) becomes:

$$\begin{aligned} L_{\text{MAP}}(\phi, \boldsymbol{\theta}) = & \sum_v \ln S_o(v) + N^2 \ln \sigma^2 \\ & + \sum_v \ln \left(|\tilde{h}_1(\phi, v)|^2 + |\tilde{h}_2(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right) \\ & + \frac{1}{2} \sum_v \frac{\left| \tilde{i}_1(v)\tilde{h}_2(\phi, v) - \tilde{i}_2(v)\tilde{h}_1(\phi, v) \right|^2}{\sigma^2 \left(|\tilde{h}_1(\phi, v)|^2 + |\tilde{h}_2(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right)} \\ & + \frac{1}{2} \sum_v \frac{\left| \tilde{h}_1(\phi, v)\tilde{o}_m(v) - \tilde{i}_1(v) \right|^2 + \left| \tilde{h}_2(\phi, v)\tilde{o}_m(v) - \tilde{i}_2(v) \right|^2}{S_o(v) \left(|\tilde{h}_1(\phi, v)|^2 + |\tilde{h}_2(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right)} \\ & + \gamma \sum_{(l,m)\in S} \left[\left| e^{j(\phi_{l-1,m} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l+1,m})} \right|^2 \right. \\ & \left. + \left| e^{j(\phi_{l,m-1} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l,m+1})} \right|^2 \right] \end{aligned} \tag{55}$$

where γ is the hyperparameter that quantifies the trade-off between goodness of fit to the data and to the prior. There is no closed-form expression for $\hat{\phi}$

that minimizes the criterion $L_{\text{MAP}}(\phi, \boldsymbol{\theta})$, so the minimization has to be done using an iterative method.

- Tuning of the hyperparameters
  - Noise and object: As in Subsection 3.2.6, the hyperparameters linked to the noise and the object are estimated jointly with the phase parameters. We recall that:

$$(\hat{\phi}, \hat{\boldsymbol{\theta}}_o, \hat{\boldsymbol{\theta}}_b) = \underset{\phi, \boldsymbol{\theta}_o, \boldsymbol{\theta}_b}{\arg\min}\, J_{\text{MAP}}(\phi, \hat{\boldsymbol{\theta}}_o, \hat{\boldsymbol{\theta}}_b, \hat{\boldsymbol{\theta}}_\phi) \tag{56}$$

  - Phase: The value of γ has to be adjusted by hand. In fact, γ can not be jointly estimated with ϕ (the solution $\gamma = 0$ corresponding to a null regularization minimizes the criterion taken as a function of γ for any given value of the phase).

Phase-unwrapping The preceding estimation procedure produces an estimate of the wrapped phase. For applications requiring the phase itself, it is necessary, in a second step, to calculate the phase from the wrapped phase. In this second step, a phase-unwrapping method is used. The phase-unwrapping problem is found in a variety of applications and several phase-unwrapping methods have been proposed in the literature (see for example Ghiglia and Pritt (1998)). We will not develop here these methods. In the following, we will focus on the wrapped phase estimation, which is specific to phase diversity contrarily to the phase unwrapping problem.

8.3 Simulation results

We will show simulation results obtained with the MAP estimator on turbulence-induced aberrations.

8.3.1 Choice of an error metric

Because we focus on the estimation of the wrapped phase $\phi[2\pi]$, we have developed a metric which quantifies the quality of the phase estimate within the interval $[-\pi, \pi]$.

$$\epsilon = \sqrt{\frac{1}{\mathcal{A}} \sum_{(l,m)\in S} c(l,m)^2} \tag{57}$$

$$\text{with} \quad c(l,m) = \sqrt{\left| e^{j\left(\phi^{\text{true}}_{l,m} - \phi^{\text{estimated}}_{l,m}\right)} - \left\langle e^{j\left(\phi^{\text{true}}_{l,m} - \phi^{\text{estimated}}_{l,m}\right)} \right\rangle \right|^2} \tag{58}$$

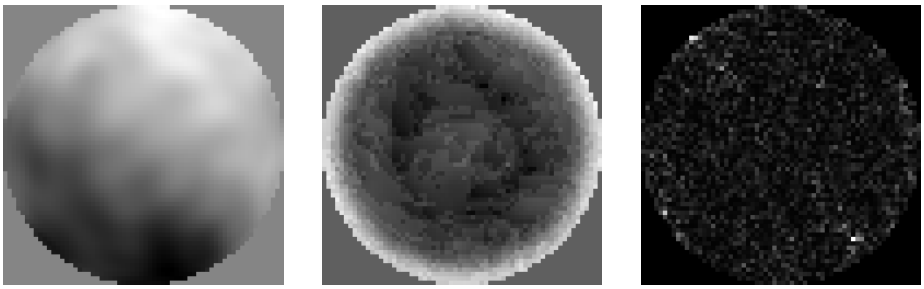
Where $\mathcal{A}$ is the number of pixels in the pupil. c is the 2D error map that indicates the difference between the true phase and its estimate at each pupil pixel. Note

that this error metric is insensitive, as the data is, to any 2π-variation of the phase and to piston terms. A first-order expansion of Equation (57) shows that, for small errors on the phase, ϵ corresponds to the standard deviation of the residual phase.

8.3.2 Results

Data generation Twenty turbulent wave-fronts are obtained using a modal method (Roddier, 1990): the phase is expanded on the Zernike polynomial basis, and given Kolmogorov statistics. The strength of the turbulence is fixed by the ratio D/r_o of the telescope diameter to the Fried parameter (see Subsection 1.2.3). For each of these turbulent phases, we compute a Shannon-sampled image. The image noise is a uniform Gaussian noise of 1%. The defocus amplitude between the two images is set to 2π radians, peak-to-valley.

Point source We first use a point source object. The strength of the turbulence is set to $D/r_o = 30$ (this corresponds to very strong turbulence conditions). The theoretical spatial standard deviation of the phase is then equal to 2.7λ (Noll, 1976), which corresponds to a peak-to-valley amplitude of 30π radians. Despite this high value, this case is a favorable situation because the object is known and the noise level is low. Accurate results are then obtained, even without regularization on the phase, i.e., the point-by-point phase is estimated using the maximum likelihood approach. The average error ϵ on the twenty turbulent wave-fronts is equal to $\lambda/30$. An example of a phase estimate is shown Figure 23 (middle), to be compared to the true phase (left). The comparison of these two phases by visual inspection does



true phase | phase estimate | 2D error map c.

Figure 23: Example of a turbulent phase estimate ($D/r_o = 30$) from a point source object. The 2D error map representation has been multiplied by 10 compared to the others. The error term ϵ is equal to $\lambda/28$.

not give a correct idea of the quality of the estimation. This is due to the fact that the phase estimate presents several 2π-jumps whereas the true phase is unwrapped. The good quality of the estimate can be seen through the 2D error map c (Figure 23 right) or on Figure 24, which shows the true phase and the estimated one both

wrapped within $[-\pi, \pi]$.

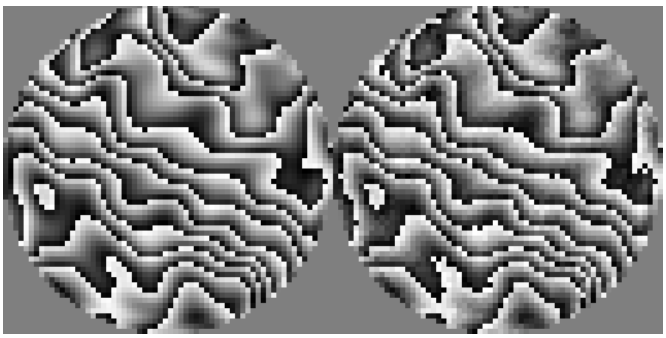


Figure 24: Example of a turbulent phase estimate ($D/r_o = 30$) from a point source object: left, the true phase within $[-\pi, \pi]$ and right, the phase estimate within $[-\pi, \pi]$.

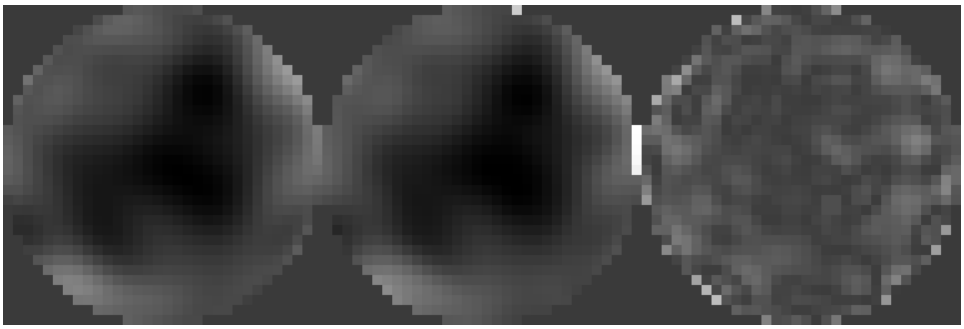
Extended object The discrete extended object used here is a spiral galaxy (Figure 25). Three different strengths of turbulence are studied: low ($D/r_o = 4$),

Figure 25: Extended object used for simulations.

medium ($D/r_o = 6$) and strong turbulence levels ($D/r_o = 8$). Phase estimates are obtained using the MAP approach of Equation (55).

- **Low turbulence level** The strength of turbulence is set to $D/r_o = 4$, corresponding to a peak-to-valley amplitude of 1.2π radians. The point-by-point phase estimates have been obtained using a focused image and a defocused one and the starting phase estimate of the iterative criterion minimization is

zero. The average error ϵ on the twenty phase estimates is equal to $\lambda/60$. Figure 26 (middle) shows an example of a phase estimate for $D/r_o = 4$, compared to the true phase (left). The right side of this Figure depicts the (ten times magnified) 2D error map c.



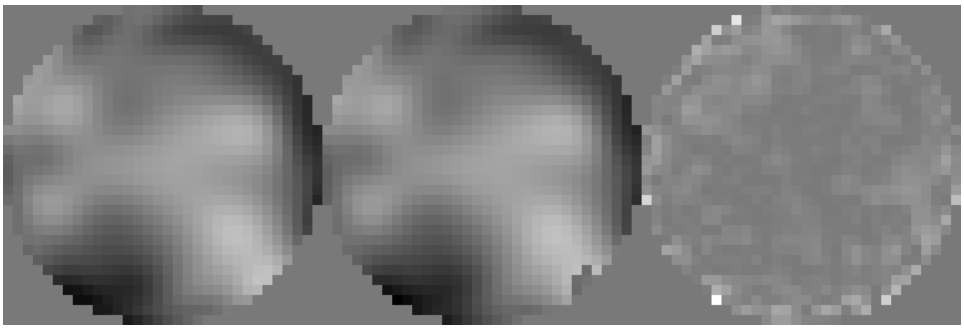
true phase phase estimate 2D error map c.

Figure 26: Example of a turbulent phase estimate ($D/r_o = 4$) from the galaxy object. The 2D error map representation has been multiplied by 10 compared to the others. The error term is equal to $\lambda/60$.

- **Medium turbulence level** The strength of turbulence is now set to $D/r_o = 6$, corresponding to a peak-to-valley amplitude of 1.5π radians. The quality of the point-by-point phase estimates obtained with a focused and a defocused image is poor ($\epsilon = \lambda/15$). This is probably due to the presence of local minima in the MAP criterion. Obviously, although we have constructed a criterion such that all minima corresponding to a 2π-variation of the phase are equivalent, there are other reasons for the presence of local minima such as the non-linear relationship between the phase and the image (see Equation (1)). To improve the estimation, we have used an additional defocused image (with a defocus from the focused plane equal to 4π). The reconstruction clearly shows an improvement when the three images are used: the mean error ϵ becomes $\lambda/40$.

- **Strong turbulence level** Finally, we consider stronger conditions of turbulence with $D/r_o = 8$, corresponding to a peak-to-valley amplitude of 2π radians. The use of three images leads to a poor quality of estimation, i.e., the mean error ϵ is equal to $\lambda/17$. We propose, as suggested by Jefferies et al. (2002), to improve the estimation by using a better starting phase estimate point for the criterion minimization. An initial estimate of low order Zernike polynomial coefficients (that exhibit greater power) is performed: the phase is expanded on the first Zernike coefficients a_4 to a_6. After estimation of these few Zernike coefficients, the estimation of the point-by-point phase parameters is started using the corresponding phase as the starting estimate for the minimization of the MAP criterion. The use of this starting point

for the minimization makes the estimation more precise: the average error ϵ obtained using three images becomes $\lambda/45$. This starting point is closer to a global minimum than the null phase initially used. Figure 27 (middle) shows an example of a phase estimate for $D/r_o = 8$, compared to the true phase (left). On the right side of this Figure, the 2D error map c is depicted.



true phase phase estimate 2D error map c.

Figure 27: Example of a turbulent phase estimate ($D/r_o = 8$) from the galaxy object. The 2D error map representation has been multiplied by 10 compared to the others. Three images have been used and the initial phase estimate is obtained using a first estimation of the Zernike coefficients a_4 to a_6. The error term is equal to $\lambda/45$.

To summarize, we have shown that the estimation of large aberrations from a known object and for a low noise level does not require the use of an explicit regularization term on the phase. For extended objects, the use of the two conventional images of the phase diversity perform well for low turbulence levels. For medium turbulence levels (here $D/r_o = 6$), the quality of the phase estimate is enhanced through the use of an additional defocused image. Finally, for higher turbulence cases, the accuracy is maintained if the point-by-point phase estimation is done with three images and is preceded by a first modal estimation of the low spatial frequencies of the aberrated phase. The measurement of large aberrations by use of phase diversity is a problem that has only been recently addressed. The results of the works referenced in these pages together with the ones presented here are quite promising and open the way to making phase diversity a practical wave-front sensor for adaptive optics.

9 Emerging applications: cophasing of multi-aperture telescopes

9.1 Background

The resolution of a telescope is ultimately limited by its aperture diameter. The latter is limited by current technology to about 10 m for ground-based telescopes and to a few meters for space-based telescopes because of volume and mass considerations. Interferometry allows one to go beyond this limit; it consists in making an array of sub-apertures interfere; the resulting instrument is called an interferometer or a multi-aperture telescope. So far, this technique has been used solely on ground-based instruments. The sub-apertures can either be telescopes *per se*, as in astronomy (e.g., VLTI, NPOI), or segments of a common primary mirror. If these segments are adjacent, such as in the Keck telescope, one refers to the instrument as a segmented telescope rather than an interferometer, even though it is conceptually one. Regarding high-resolution space-borne missions, interferometers are forecast in astronomy (with a segmented aperture for the JWST and a diluted aperture for the Darwin[9] mission for instance) and can also be considered for Earth observation (Mugnier et al., 2004).

For a correct performance, the aperture of such an instrument must be phased to within a small fraction of the wavelength. A critical sub-system of interferometers is thus the cophasing sensor (CS), whose goal is to measure the relative positioning (differential piston and tip/tilt) of the sub-apertures, which are the main sources of wave-front degradations, and possibly the higher-order aberrations on each sub-aperture.

Differential piston and tip-tilt measurement has been studied extensively and demonstrated for distant ground-based telescopes. Most of the proposed devices are based on a pupil-plane combination of the light coming from a given pair of sub-apertures. Because of their pupil-plane combination, the contrast of interference fringes decreases quickly as the object extension increases, which makes these devices useless on very extended scenes such as the Earth viewed from Space. Because of the pair-wise light combination, these devices become impractical for instruments made of many sub-apertures.

The fact that phase diversity can be used as a CS on a segmented aperture telescope was recognized very early (Paxman and Fienup, 1988). Additionally, in contrast with the above-mentioned devices, phase diversity enjoys two appealing characteristics: firstly, it is appropriate for an instrument with a large number of sub-apertures, because the complexity of the hardware does not scale with the number of sub-apertures and remains essentially independent of it. Secondly, it can be used on very extended objects.

This first property and the absence of non-common-path aberrations (see, e.g.,

[9] Darwin is a forecast ESA mission whose aim is to find and characterize Earth-like planets. The instrument will be a so-called nulling interferometer, which cancels the light coming from the star in order to detect the planet—see, e.g., http://www.esa.int/science/darwin for more information.

subsection 7.2) are two strong motivations for the choice of phase diversity as a CS, even when looking at an unresolved source. Phase diversity experiments as a CS on a point-source have been performed on the ground, notably to cophase the segments of the Keck telescope (Löfdahl et al., 1998b). Regarding space instruments, phase diversity has long been foreseen as a candidate of choice for the JWST segmented-aperture telescope (see, e.g., Redding et al. (1998), Carrara et al. (2000), Lee et al. (2003)), and has recently been selected as the CS for the Darwin interferometer (Cassaing et al., 2003, Mocoeur et al., 2005).

In 1994 (Kendrick et al., 1994a,b), phase diversity was experimentally validated as a CS on an extended source for correcting static aberrations in real-time with a segmented mirror. Quite recently, the loop was also closed with a phase diversity CS to correct static aberrations in the difficult framework of a six-telescope imaging interferometer using broadband light (Zarifis et al., 1999, Seldin and Paxman, 2000).

9.2 Experimental results on an extended scene

As mentioned in the previous subsection, phase diversity can be used with extended scenes. Actually, for applications such as Earth observation from Space, the scenes extend beyond the field of view recorded by the image and phase diversity is one of the very few possible CS's (Mugnier et al., 2004). We have designed, built and validated a prototype phase diversity CS for this application. After a short presentation of this prototype and its testbed, we present its latest results[10]. A more comprehensive presentation of the testbed along with earlier results can be found in (Sorrente et al., 2004, Baron, 2005).

A schematic view of the testbed, called BRISE for Banc Reconfigurable d'Imagerie sur Scènes Etendues, is shown on Fig. 28. BRISE is mainly composed of four modules (source, perturbation, detection and control), described below.

The source module delivers two objects: an extended scene, which is an Earth scene on a high-resolution photographic plate illuminated by an arc-lamp, and a reference point source, which is the output of a monomode fiber fed with a He-Ne laser.

The perturbation module has three functions: it images the source on the detector, defines the aperture configuration and introduces calibrated aberrations; its main component is the deformable mirror (DM), which performs the latter function. In order to introduce only piston and tip/tilt, we have chosen to manufacture a specific segmented DM consisting of three planar mirrors mounted on piezo-actuated platforms by Physik Instrument, which have exactly these three degrees of freedom.

The detection module is a water-cooled CCD camera that simultaneously records a focal-plane image and a defocused image of each of the two objects

[10]These experimental results are courtesy of I. Mocœur, after preliminary results by F. Baron. F. Cassaing is gratefully acknowledged for being the master architect of this prototype and testbed, and B. Sorrente for overseeing their realization.

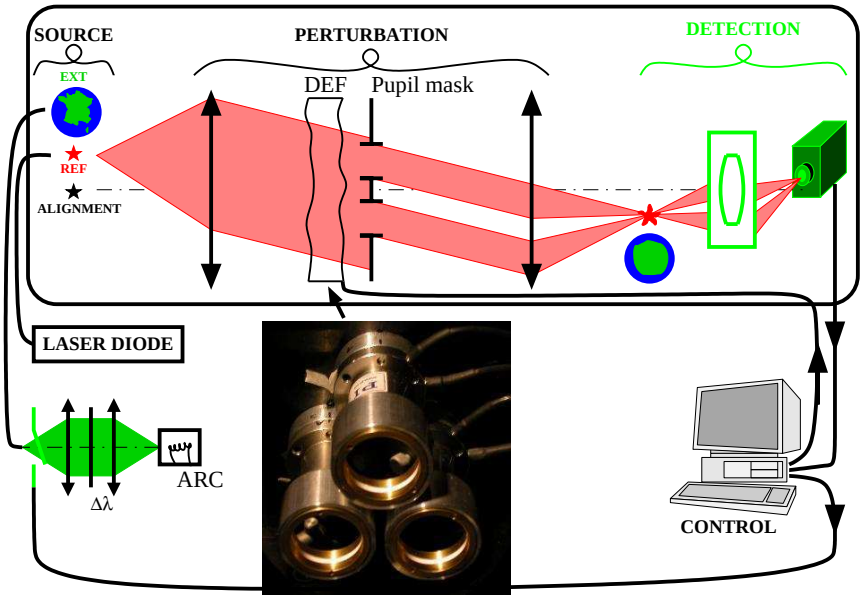


Figure 28: Schematic view of the BRISE testbed and photograph of the deformable mirror (DEF), courtesy F. Cassaing. EXT: extended object; REF: reference point-source; ARC: arc lamp used for the illumination of EXT.

to implement a phase diversity CS. Figure 29 shows an experimental example of such an image. The control module drives the experiment.

Special care has been given to the control of errors that could limit CS performance or the evaluation of the CS performance on extended objects. In particular, the two objects are observed simultaneously through very close paths, to minimize the differential effects of field aberrations, vibrations or air turbulence. A very accurate aberration calibration can thus be achieved thanks to the high SNR of the measurement obtained on the reference point source.

Figure 30 presents the piston measured at high photon level on a given subaperture as a function of the piston effectively introduced by the DM, for the reference point source at $\lambda_r = 633\,\text{nm}$ and for the extended scene, illuminated with white light and a spectral filter of width $40\,\text{nm}$ centered around $\lambda_e = 650\,\text{nm}$. For each introduced piston, three measurements are performed and reported on this figure. The point-source measurements exhibit an excellent linearity between roughly $-\lambda_r/2$ and $+\lambda_r/2$, at which points the expected modulo 2π wrapping occurs. With the extended object, the curve is linear on a slightly smaller piston range. Some features are different on this curve with respect to the one obtained with the reference point: the slope is not exactly unity, although this would not be a major problem in closed loop, and the sort of smooth wraparound that occurs around $+\lambda_e/2$ is somewhat surprising and currently interpreted as a consequence of the spectral bandwidth. Figure 31 shows the repeatability obtained on the piston measurement

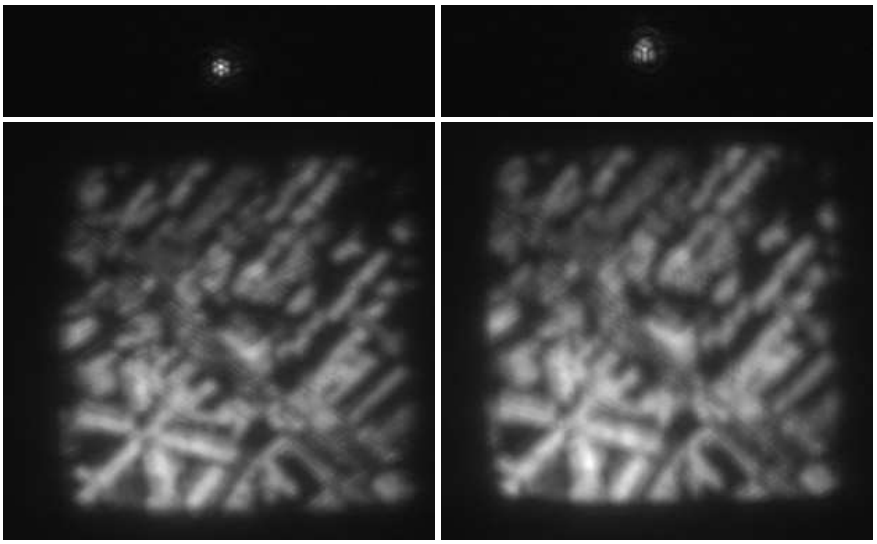


Figure 29: Focused (left) and defocused (right) experimental images of the extended scene (bottom) and reference point source (top) objects. These images are recorded simultaneously on different parts of the same detector and used for phase diversity.

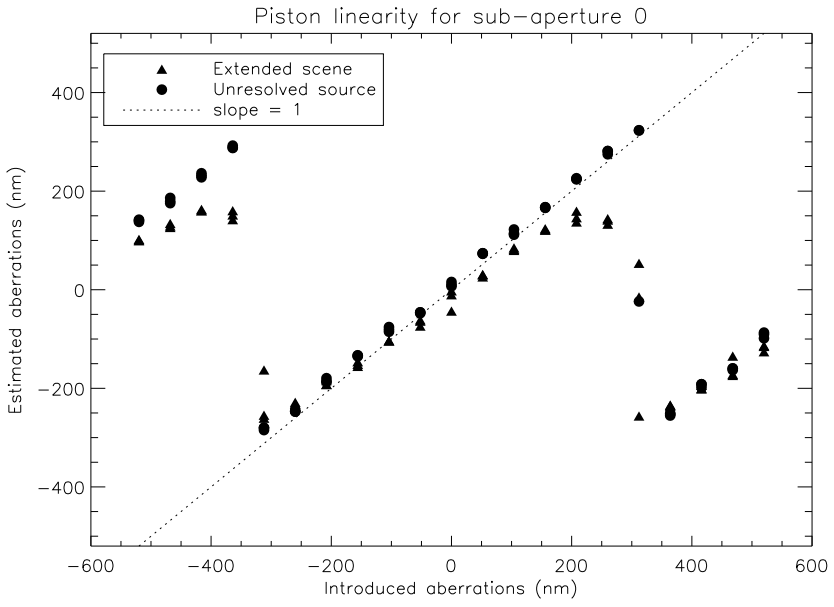


Figure 30: Piston measured at high photon level on the first sub-aperture, as a function of the piston effectively introduced by the DM.

with the extended object. The standard deviation of the estimated piston is, as expected, dominated by detector noise for low fluxes, and then inversely proportional to the square root of the number of photons per pixel (photon-noise regime). It is for instance below 1 nm as soon as the average flux is above 1000 photo-electrons per pixel.

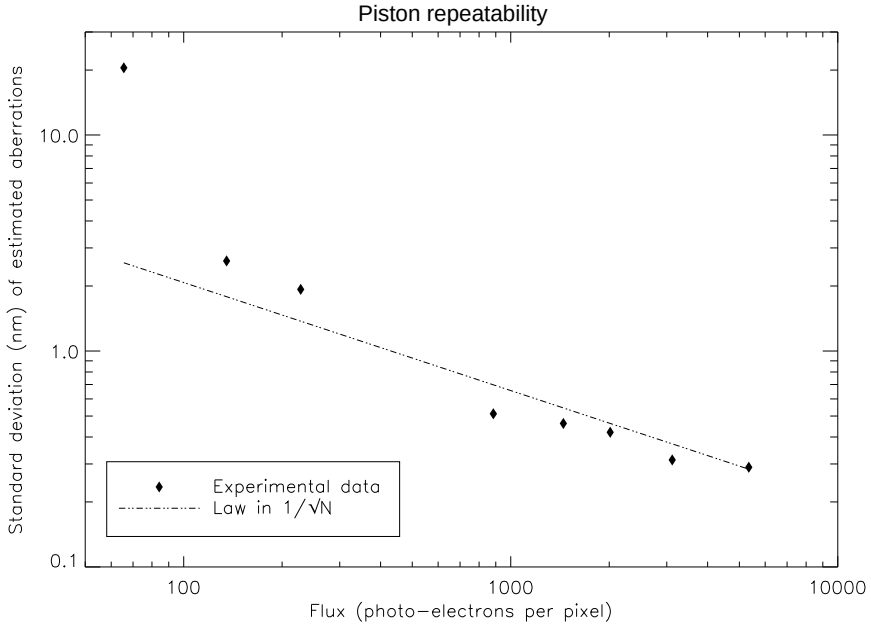


Figure 31: Repeatability obtained on the measurement of the piston on the first sub-aperture with the extended object, as a function of the average photon level per pixel.

9.3 Conclusion

Both the literature cited in this section and the quantitative results presented in it testify that phase diversity can be successfully used as a cophasing sensor on segmented-aperture telescopes and on interferometers, for point-sources as well as for extended objects. Some challenges remain to be met before this use is widespread. They are essentially the same as for single-aperture telescope wavefront sensing: the ability to sense large-amplitude aberrations, and the reduction of the computing cost for use in real-time applications, be it the compensation of atmospheric turbulence for ground-based instruments or the compensation of environmental perturbations for space-based instruments.

Bibliography

D. S Acton, D. Soltau, and W. Schmidt. Full-field wavefront measurements with phase diversity. *Astron. Astrophys.*, 309:661–672, 1996.

N. Baba and K. Mutoh. Iterative reconstruction method in phase-diversity imaging. *Appl. Opt.*, 33(20):4428–4433, 1994.

N. Baba and K. Mutoh. Measurement of telescope aberrations through atmospheric turbulence by use of phase diversity. *Appl. Opt.*, 40(4):544–552, 2001.

F. Baron. *Définition et test d'un capteur de cophasage sur télescope multipupilles: application à la détection d'exoplanètes et à l'observation de la Terre*. PhD thesis, Ecole Doctorale d'Astronomie et d'Astrophysique d'Ile de France, 2005.

H. H. Bauschke, P. L. Combettes, and D. R. Luke. Phase retrieval, Gerchberg-Saxton algorithm, and Fienup variants: A view from convex optimization. *J. Opt. Soc. Am. A*, 19(7):1334–1345, 2002.

A. Blanc. *Identification de réponse impulsionnelle et restauration d'images : apports de la diversité de phase*. PhD thesis, Université Paris XI Orsay, July 2002.

A. Blanc, T. Fusco, M. Hartung, L. M. Mugnier, and G. Rousset. Calibration of NAOS and CONICA static aberrations. Application of the phase diversity technique. *Astron. Astrophys.*, 399:373–383, 2003a.

A. Blanc, L. M. Mugnier, and J. Idier. Marginal estimation of aberrations and image restoration by use of phase diversity. *J. Opt. Soc. Am. A*, 20(6):1035–1045, 2003b.

M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, Sixth (corrected) edition, 1993.

O. M. Bucci, A. Capozzoli, and G. D'Elia. Regularizing strategy for image restoration and wave-front sensing by phase diversity. *J. Opt. Soc. Am. A*, 16(7):1759–1768, 1999.

C. L. Byrne. Convergent block-iterative algorithms for image reconstruction from inconsistent data. *IEEE Trans. Image Processing*, 6(9):1296–1304, 1997.

D. A. Carrara, B. J. Thelen, and R. G. Paxman. Aberration correction of segmented-aperture telescopes by using phase diversity. In M. A. Fiddy and R. P. Millane, editors, *Image reconstruction from incomplete data*, volume 4123, pages 56–63. Soc. Photo-Opt. Instrum. Eng., 2000.

E. De Carvalho and D. Slock. Maximum-likelihood blind FIR multi-channel estimation with Gaussian prior for the symbols. In *ICASSP*, pages 3593–3596, April 1997.

F. Cassaing, F. Baron, E. Schmidt, S. Hofer, L. M. Mugnier, M. Barillot, G. Rousset, T. Stuffler, and Y. Salvadé. DARWIN Fringe Sensor (DWARF): Concept study. In *Towards Other Earths*, volume SP-539, pages 389–392. ESA, 2003. Conference date: April 2003.

Y. Censor, Eggermont P. P. B., and D. Gordon. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numerische Mathematik*, 41:83–92, 1983.

F. Champagnat and J. Idier. An alternative to standard maximum likelihood for Gaussian mixtures. In *ICASSP*, pages 2020–2023, May 1995.

J.-M. Conan. *Étude de la correction partielle en optique adaptative*. PhD thesis, Université Paris XI Orsay, October 1994.

J. M. Conan, P. Y. Madec, and G. Rousset. Image formation in adaptive optics partial correction. In F. Merkle, editor, *Active and Adaptive Optics*, volume 48 of *ESO Conference and Workshop Proceedings*, pages 181–186, Garching bei München, Germany, 1994. ESO/ICO.

J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset. Myopic deconvolution of adaptive optics images using object and point spread function power spectra. *Appl. Opt.*, 37(21):4614–4622, July 1998.

J.-M. Conan, G. Rousset, and P.-Y. Madec. Wave-front temporal spectra in high-resolution imaging through turbulence. *J. Opt. Soc. Am. A*, 12(12):1559–1570, July 1995.

B. L. Ellerbroek, B. J. Thelen, D. J. Lee, and R. G. Paxman. Comparison of Shack-Hartmann wavefront sensing and phase-diverse phase retrieval. In Robert K. Tyson and Robert Q. Fugate, editors, *Adaptive Optics and Applications*, volume 3126, pages 307–320. Soc. Photo-Opt. Instrum. Eng., 1997.

J. R. Fienup. Phase retrieval algorithms: a comparison. *Appl. Opt.*, 21(15):2758–2769, August 1982.

J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin. Hubble space telescope characterized by using phase-retrieval algorithms. *Appl. Opt.*, 32(10): 1747–1767, 1993.

J.-C. Fontanella. Analyse de surface d'onde, déconvolution et optique active. *J. of Optics (Paris)*, 16(6):257–268, 1985.

D. L. Fried. Statistics of a geometric representation of wavefront distortion. *J. Opt. Soc. Am.*, 55(11):1427–1435, 1965.

D. L. Fried. Limiting resolution looking down through the atmosphere. *J. Opt. Soc. Am.*, 56(10):1380–1384, 1966a.

D. L. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *J. Opt. Soc. Am.*, 56:1372–1379, 1966b.

F. R. Gantmacher. L'algorithme de Gauss et quelques-unes de ses applications. In *Théorie des Matrices, Tome I*, chapter II, pages 42–50. Dunod, 1966.

E. L. Gates, S. R. Restaino, R. A. Carreras, R. C. Dymale, and G. C. Loos. Phase diversity as an on-line wavefront sensor: experimental results. In T. J. Schulz and D. L. Snyder, editors, *Image Reconstruction and Restoration*, volume 2302, pages 330–339. Soc. Photo-Opt. Instrum. Eng., 1994.

R. W. Gerchberg. Super-resolution through error energy reduction. *Opt. Acta*, 21: 709–720, 1974.

R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–246, 1972.

D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley-Interscience, 1998.

R. A. Gonsalves. Phase retrieval from modulus data. *J. Opt. Soc. Am.*, 66(9): 961–964, 1976.

R. A. Gonsalves. Phase retrieval and diversity in adaptive optics. *Opt. Eng.*, 21(5): 829–832, 1982.

R. A. Gonsalves. Nonisoplanatic imaging by phase diversity. *Opt. Lett.*, 19(7): 493–495, April 1994.

R. A. Gonsalves. Compensation of scintillation with a phase-only adaptive optic. *Opt. Lett.*, 22:588–590, 1997.

J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1968.

Y. Goussard, G. Demoment, and J. Idier. A new algorithm for iterative deconvolution of sparse spike. In *ICASSP*, pages 1547–1550, April 1990.

M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset, and R. Lenzen. Calibration of NAOS and CONICA static aberrations. Experimental results. *Astron. Astrophys.*, 399:385–394, 2003.

B. R. Hunt. The application of constrained least squares estimation to image restoration by digital computer. *IEEE Trans. Comp.*, C-22(9):805–812, 1973.

J. Idier, L. Mugnier, and A. Blanc. Statistical behavior of joint least square estimation in the phase diversity context. *IEEE Trans. Image Processing*, 14(12): 2107–2116, December 2005.

S. M. Jefferies, M. Lloyd-Hart, E. Keith Hege, and J. Georges. Sensing wave-front amplitude and phase with phase diversity. *Appl. Opt.*, 41(11):2095–2102, 2002.

A. P. Kattnig and J. Primot. Model of the second-order statistic of the radiance field of natural scenes, adapted to system conceiving. In S. K. Park and R. D. Juday, editors, *Visual information processing VI*, volume 3074, pages 132–141. Soc. Photo-Opt. Instrum. Eng., 1997.

R. L. Kendrick, D. S. Acton, and A. L. Duncan. Experimental results from the Lockheed phase diversity test facility. In T. J. Schulz and D. L. Snyder, editors, *Image reconstruction and restoration*, volume 2302, pages 312–322. Soc. Photo-Opt. Instrum. Eng., 1994a.

R. L. Kendrick, D. S. Acton, and A. L. Duncan. Phase-diversity wave-front sensor for imaging systems. *Appl. Opt.*, 33(27):6533–6546, September 1994b.

R. L. Kendrick, R. Bell, and A. L. Duncan. Closed loop wave front correction using phase diversity. In P. Y. Bely and J. B. Breckinridge, editors, *Space telescopes and instruments V*, volume 3356, pages 844–853. Soc. Photo-Opt. Instrum. Eng., 1998.

A. Labeyrie. Attainment of diffraction-limited resolution in large telescopes by Fourier analysing speckle patterns. *Astron. Astrophys.*, 6:85–87, 1970.

D. J. Lee, M. C. Roggemann, and B. M. Welsh. Cramer-Rao analysis of phase-diverse wave-front sensing. *J. Opt. Soc. Am. A*, 16(5):1005–1015, 1999.

D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby. Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing. *Appl. Opt.*, 36:9186–9197, December 1997a.

D. J. Lee, B. M. Welsh, and M. C. Roggemann. Diagnosing unknown aberrations in an adaptive optics system by use of phase diversity. *Opt. Lett.*, 22(13):952–954, 1997b.

L. H. Lee, G. Vasudevan, and E. H. Smith. Point-by-point approach to phase-diverse phase retrieval. In J. C. Mather, editor, *IR space telescopes and Instruments*, volume 4850, pages 441–452. Soc. Photo-Opt. Instrum. Eng., 2003.

E. Lehmann. *Theory of point estimation*. Wiley, 1983.

R. Lenzen, R. Hofmann, P. Bizenberger, and A. Tusche. CONICA: the high-resolution near-infrared camera for the ESO VLT. In A. M. Fowler, editor, *Infrared Astronomical Instrumentation*, volume 3354, pages 606–614. Soc. Photo-Opt. Instrum. Eng., 1998.

R. J. A. Little and D. B. Rubin. On jointly estimating parameters and missing data by maximizing the complete-data likelihood. *The American Statistician*, 37(3): 218–220, August 1983.

D. R. Luke, J. V. Burke, and R. G. Lyon. Fast algorithms for phase diversity and phase retrieval. In R. G. Lyon, editor, *Proceedings of the Workshop on Computational Optics and Imaging for Space Applications*, pages 130–150. NASA/GSFC, May 2000.

M. G. Löfdahl, A. L. Duncan, and G. B. Scharmer. Fast phase diversity wavefront sensor for mirror control. In D. Bonaccini and R. K. Tyson, editors, *Adaptive Optical System Technologies*, volume 3353, pages 952–963. Soc. Photo-Opt. Instrum. Eng., March 1998a.

M. G. Löfdahl, R. L. Kendrick, A. Harwit, K. E. Mitchell, A. L. Duncan, J. H. Seldin, R. G. Paxman, and D. S. Acton. Phase diversity experiment to measure piston misalignement on the segmented primary mirror of the Keck II telescope. In P. Y. Bely and J. B. Breckinridge, editors, *Space Telescopes and Instruments V*, volume 3356, pages 1190–1201. Soc. Photo-Opt. Instrum. Eng., 1998b.

M. G. Löfdahl and G. B. Scharmer. Wavefront sensing and image restoration from focused and defocused solar images. *Astron. Astrophys.*, 107:243–264, 1994.

M. G. Löfdahl and G. B. Scharmer. A predictor approach to closed-loop phase-diversity wavefront sensing. In J. B. Breckinridge and P. Jakobsen, editors, *UV, optical and IR space telescopes and instruments*, volume 4013, pages 737–748. Soc. Photo-Opt. Instrum. Eng., 2000.

M. G. Löfdahl and G. B. Scharmer. Phase diverse speckle inversion applied to data from the Swedish 1-meter solar telescope. In Keil and Avakyan, editors, *Innovative telescopes and instrumentation for solar astrophysics*, volume 4853. Soc. Photo-Opt. Instrum. Eng., 2002.

M. G. Löfdahl, G. B. Scharmer, and W. Wei. Calibration of a deformable mirror and Strehl ratio measurements by use of phase diversity. *Appl. Opt.*, 39(1):94–103, January 2000.

L. Meynadier, V. Michau, M.-T. Velluet, J.-M. Conan, L. M. Mugnier, and G. Rousset. Noise propagation in wave-front sensing with phase diversity. *Appl. Opt.*, 38(23):4967–4979, August 1999.

D. L. Misell. An examination of an iterative method for the solution of the phase problem in optics and electron optics: I. Test calculations. *J. Phys. D: Appl. Phys.*, 6:2200–2216, 1973.

I. Mocoeur, F. Cassaing, F. Baron, L. M. Mugnier, G. Rousset, B. Sorrente, and A. Blanc. Multi-telescope interferometer cophasing for astronomy. In *Semaine de l'astrophysique française*. EDP Sciences, 2005.

J. J. Moré and D. J. Thuente. Line search algorithms with guaranteed sufficient decrease. *ACM Transactions on Mathematical Software*, 20:286–307, 1994.

L. Mugnier, F. Cassaing, B. Sorrente, F. Baron, M.-T. Velluet, V. Michau, and G. Rousset. Multiple-aperture optical telescopes: some key issues for Earth observation from a GEO orbit. In *5th International Conference On Space Optics*, volume SP-554, pages 181–187, Toulouse, France, 2004. CNES/ESA, ESA.

L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem. Myopic deconvolution from wavefront sensing. *J. Opt. Soc. Am. A*, 18:862–872, April 2001.

J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Texts in Operations Research. Springer-Verlag, New York, NY, USA, 1999.

R. J. Noll. Zernike polynomials and atmospheric turbulence. *J. Opt. Soc. Am.*, 66 (3):207–211, 1976.

A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circuits Syst.*, CAS-22(9):735–742, September 1975.

R. G. Paxman and S. L. Crippen. Aberration correction for phased-array telescopes using phase diversity. In A. F. Gmitro, P. S. Idell, and I. J. LaHaie, editors, *Digital Image Synthesis and Inverse Optics*, volume 1351, pages 787–797. Soc. Photo-Opt. Instrum. Eng., 1990.

R. G. Paxman and J. R. Fienup. Optical misalignment sensing and image reconstruction using phase diversity. *J. Opt. Soc. Am. A*, 5(6):914–923, 1988.

R. G. Paxman, T. J. Schulz, and J. R. Fienup. Joint estimation of object and aberrations by using phase diversity. *J. Opt. Soc. Am. A*, 9(7):1072–1085, 1992.

R. G. Paxman, J. H. Seldin, M. G. Löfdahl, G. B. Scharmer, and C. U. Keller. Evaluation of phase-diversity techniques for solar-image restoration. *Astrophys. J.*, 1996.

R. G. Paxman, B. J. Thelen, D. A. Carrara, J. H Seldin, and K. W. Gleichman. Myopic deblurring of space-variant blur by using phase-diverse speckle. *IEEE Trans. Image Processing*, 1998.

R. G. Paxman, B. J. Thelen, and J. H. Seldin. Phase-diversity correction of turbulence-induced space-variant blur. *Opt. Lett.*, 19(16):1231–1233, August 1994.

S. Prasad. Information-optimized phase diversity speckle imaging. *Opt. Lett.*, 29 (6):563–565, 2004.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C, the art of scientific computing*. Cambridge University Press, New York, 2nd edition, 1992.

J. Primot, G. Rousset, and J.-C. Fontanella. Image deconvolution from wavefront sensing: atmospheric turbulence simulation cell results. In M.-H. Ulrich, editor, *Very Large Telescopes and their Instrumentation, Volume II*, volume 30 of *ESO Conference and Workshop Proceedings*, pages 683–692, Garching bei München, Germany, 1988. ESO.

D. Redding et al. Wavefront sensing and control for a Next Generation Space Telescope. In P. Y. Bely and J. B. Breckinridge, editors, *Space Telescopes and Instruments V*, volume 3356 (2), pages 758–772. Soc. Photo-Opt. Instrum. Eng., 1998.

C. Roddier and F. Roddier. Reconstruction of the Hubble Space Telescope mirror figure from out-of-focus stellar images. In P. Y. Bely and J. B. Breckinridge, editors, *Space Astronomical Telescopes and Instruments*, volume 1494, pages 78–84. Soc. Photo-Opt. Instrum. Eng., 1991.

C. Roddier and F. Roddier. Combined approach to the Hubble space telescope wave-front distortion analysis. *Appl. Opt.*, 32(16):2992–3008, 1993.

F. Roddier. The effects of atmospherical turbulence in optical astronomy. In E. Wolf, editor, *Progress in Optics*, volume XIX, pages 281–376. North Holland, Amsterdam, 1981.

F. Roddier. Curvature sensing and compensation: a new concept in adaptive optics. *Appl. Opt.*, 27(7):1223–1225, April 1988.

F. Roddier, editor. *Adaptive Optics in Astronomy*. Cambridge University Press, Cambridge, 1999.

F. Roddier, J. M. Gilli, and G. Lund. On the origin of speckle boiling and its effects in stellar speckle interferometry. *J. of Optics (Paris)*, 13(5):263–271, 1982.

N. Roddier. Atmospheric wavefront simulation using Zernike polynomials. *Opt. Eng.*, 29(10):1174–1180, 1990.

M. C. Roggemann. Limited degree-of-freedom adaptive optics and image reconstruction. *Appl. Opt.*, 30(29):4227–4233, 1991.

G. Rousset. Wave-front sensors. In Roddier (1999), chapter 5, pages 91–130.

G. Rousset, F. Lacombe, P. Puget, N. Hubin, E. Gendron, J.-M. Conan, P. Kern, P.-Y. Madec, D. Rabaud, D. Mouillet, A.-M. Lagrange, and F. Rigaut. Design of the Nasmyth Adaptive Optics System (NAOS) of the VLT. In D. Bonaccini and R. K. Tyson, editors, *Astronomical Telescopes & Instrumentation*, volume 3353. Soc. Photo-Opt. Instrum. Eng., March 1998.

G. B. Scharmer. Object-independent fast phase-diversity. In T. R. Rimmele, K. S. Balasubramaniam, and R. R. Radick, editors, *High resolution solar physics:*

Theory, observations and techniques, volume 183 of *Astron. Soc. Pacific Conf. Series*, pages 330–341, 1999.

T. J. Schulz. Multiframe blind deconvolution of astronomical images. *J. Opt. Soc. Am. A*, 10(5):1064–1073, 1993.

J. H. Seldin and R. G. Paxman. Phase-diverse speckle reconstruction of solar data. In T. J. Schulz and D. L. Snyder, editors, *Image Reconstruction and Restoration*, volume 2302, pages 268–280. Soc. Photo-Opt. Instrum. Eng., 1994.

J. H. Seldin and R. G. Paxman. Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity. In J. B. Breckinridge, R. A. Carreras, S. R. Czyzak, M. J. Eckart, R. D. Fiete, and P. S. Idell, editors, *Imaging Technology and Telescopes*, volume 4091, pages 48–63. Soc. Photo-Opt. Instrum. Eng., July 2000.

J. H. Seldin, R. G. Paxman, and B. L. Ellerbroek. Post-detection correction of compensated imagery using phase-diverse speckle. In M. Cullum, editor, *Proceedings of the ESO/OSA Topical meeting on adaptive optics*, number 54, pages 471–476. ESO Conference and workshop proceedings, ESO, 1996a.

J. H. Seldin, R. G. Paxman, B. L. Ellerbroek, and D. C. Johnston. Phase-diverse speckle restorations of artificial staellites imaged with adaptive-optics compensation. In *Adaptive optics*, number 13. OSA, 1996b.

J. H. Seldin, M. F. Reiley, R. G. Paxman, B. E. Stribling, B. L. Ellerbroeck, and D. C. Johnston. Space-object identification using phase-diverse speckle. In T. J. Schulz, editor, *Imaging reconstruction and restoration II*, volume 3170, pages 2–15. Soc. Photo-Opt. Instrum. Eng., 1997.

R. B. Shack and B. C. Plack. Production and use of a lenticular Hartmann screen (abstract). *J. Opt. Soc. Am.*, 61:656, 1971.

B. Sorrente, F. Cassaing, F. Baron, C. Coudrain, B. Fleury, F. Mendez, V. Michau, L. Mugnier, G. Rousset, L. Rousset-Rouvière, and M.-T. Velluet. Multiple-aperture optical telescopes: cophasing sensor testbed. In *5th International Conference On Space Optics*, volume SP-554, pages 479–484, Toulouse, France, 2004. CNES/ESA, ESA.

O. N. Strand. Theory and methods related to the singular-function expansion and Landwebers's iteration for integral equations of the first kind. *SIAM J. Numer. Anal.*, 11(4):798–815, September 1974.

B. J. Thelen, D. A. Carrara, and R. G. Paxman. Fine-resolution imagery of extended objects observed through volume turbulence using phase-diverse speckle. In M. C. Roggemann and L. R. Bissonnette, editors, *Propagation and imaging through the atmosphere II*, volume 3763, pages 102–111. Soc. Photo-Opt. Instrum. Eng., 1999a.

B. J. Thelen, D. A Carrara, and R. G. Paxman. Pre- and post-detection correction of turbulence-induced space-variant blur. In M. C. Roggemann, editor, *Propagation and imaging through the atmosphere IV*, volume 4125. Soc. Photo-Opt. Instrum. Eng., 2000.

B. J. Thelen, R. G. Paxman, D. A. Carrara, and J. H. Seldin. Maximum a posteriori estimation of fixed aberrations, dynamic aberrations, and the object from phase-diverse speckle data. *J. Opt. Soc. Am. A*, 16(5):1016–1025, 1999b.

E. Thiébaut and J.-M. Conan. Strict *a priori* constraints for maximum-likelihood blind deconvolution. *J. Opt. Soc. Am. A*, 12(3):485–492, 1995.

C. R. Vogel. A limited memory BFGS method for an inverse problem in atmospheric imaging. In P.C. Hansen, B.H. Jacobsen, and K. Mosegaard, editors, *Methods and Applications of Inversion. Lecture Notes in Earth Sciences*, pages 292–304. Springer-Verlag, 2000.

C. R. Vogel, T. Chan, and R. Plemmons. Fast algorithms for phase-diversity-based blind deconvolution. In D. Bonaccini and R. K. Tyson, editors, *Adaptive Optical System Technologies*, volume 3353, pages 994–1005. Soc. Photo-Opt. Instrum. Eng., March 1998.

O. Von der Lühe. Speckle imaging of solar small scale structure. I–methods. *Astron. Astrophys.*, 268(1):374–390, 1993.

V. Zarifis et al. The Multi Aperture Imaging Array. In S. Unwin and R. Stachnik, editors, *Working on the Fringe: optical and IR interferometry from ground and space*, volume 194 of *Astron. Soc. Pacific Conf. Series*, pages 278–285, 1999.

Annexe C

Article Blanco & Mugnier, 2011 [A2]

Marginal blind deconvolution of adaptive optics retinal images

L. Blanco[1,2,3,*] **and L. M. Mugnier**[1,3]

[1]*ONERA - The French Aerospace Lab, F-92322 Châtillon, France*
[2]*LESIA-Observatoire de Paris, CNRS, UPMC Univ Paris 06, Univ. Paris-Diderot, France*
[3]*Groupement d'Intérêt Scientifique PHASE (Partenariat Haute résolution Angulaire Sol Espace) between ONERA, Observatoire de Paris, CNRS and Université Paris Diderot, France*
[*]*leonardo.blanco@onera.fr*

Abstract: Adaptive Optics corrected flood imaging of the retina has been in use for more than a decade and is now a well-developed technique. Nevertheless, raw AO flood images are usually of poor contrast because of the three-dimensional nature of the imaging, meaning that the image contains information coming from both the in-focus plane and the out-of-focus planes of the object, which also leads to a loss in resolution. Interpretation of such images is therefore difficult without an appropriate post-processing, which typically includes image deconvolution. The deconvolution of retina images is difficult because the point spread function (PSF) is not well known, a problem known as blind deconvolution. We present an image model for dealing with the problem of imaging a 3D object with a 2D conventional imager in which the recorded 2D image is a convolution of an invariant 2D object with a linear combination of 2D PSFs. The blind deconvolution problem boils down to estimating the coefficients of the PSF linear combination. We show that the conventional method of joint estimation fails even for a small number of coefficients. We derive a marginal estimation of the unknown parameters (PSF coefficients, object Power Spectral Density and noise level) followed by a MAP estimation of the object. We show that the marginal estimation has good statistical convergence properties and we present results on simulated and experimental data.

OCIS codes: (100.1455) Blind deconvolution; (170.4470) Ophthalmology; (010.1080) Adaptive optics; (100.3190) Inverse problems; (100.6890) Three-dimensional image processing.

References and links

1. J. Liang, D. R. Williams, and D. T. Miller, "Supernormal vision and high-resolution retinal imaging through adaptive optics," J. Opt. Soc. Am. A **14**, 2884–2892 (1997).
2. M. Glanc, E. Gendron, F. Lacombe, D. Lafaille, J.-F. Le Gargasson, and P. Léna, "Towards wide-field retinal imaging with adaptive optics," Opt. Commun. **230**, 225–238 (2004).
3. J. Rha, R. S. Jonnal, K. E. Thorn, J. Qu, Y. Zhang, and D. T. Miller, "Adaptive optics flood-illumination camera for high speed retinal imaging," Opt. Express **14**, 4552–4569 (2006).
4. A. Roorda, F. Romero-Borja, I. William Donnelly, H. Queener, T. Hebert, and M. Campbell, "Adaptive optics scanning laser ophthalmoscopy," Opt. Express **10**, 405–412 (2002).
5. L. Blanc-Féraud, L. Mugnier, and A. Jalobeanu, "Blind image deconvolution," in "Inverse Problems in Vision and 3D Tomography," , A. Mohammad-Djafari, ed. (ISTE / John Wiley, London, 2010), chap. 3, pp. 97–121.
6. G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution and its applications," Opt. Lett. **13**, 547–549 (1988).
7. L. M. Mugnier, T. Fusco, and J.-M. Conan, "Mistral: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images," J. Opt. Soc. Am. A **21**, 1841–1854 (2004).

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

8. R. J. A. Little and D. B. Rubin, "On jointly estimating parameters and missing data by maximizing the complete-data likelihood," The American Statistician **37**, 218–220 (1983).
9. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," J. Opt. Soc. Am. A **21**, 1393–1401 (2004).
10. G. Harikumar and Y. Bresler, "Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms," IEEE Trans. Image Processing **8**, 202–219 (1999).
11. J. Idier, L. Mugnier, and A. Blanc, "Statistical behavior of joint least square estimation in the phase diversity context," IEEE Trans. Image Processing **14**, 2107–2116 (2005).
12. E. Lehmann, *Theory of point estimation* (John Wiley, New York, NY, 1983).
13. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035–1045 (2003).
14. Y. Goussard, G. Demoment, and J. Idier, "A new algorithm for iterative deconvolution of sparse spike," in "ICASSP," (1990), pp. 1547–1550.
15. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
16. É. Thiébaut, "Optimization issues in blind deconvolution algorithms," in "Astronomical Data Analysis II," , vol. 4847, J.-L. Starck and F. D. Murtagh, eds. (Proc. Soc. Photo-Opt. Instrum. Eng., 2002), vol. 4847, pp. 174–183.

1. Introduction

Early detection of retinal pathologies such as glaucoma, age related macula degeneration (AMD) or retinitis pigmentosa (RP) is crucial in dealing with these conditions and calls for in vivo eye fundus imaging with a cellular level resolution, typically to be able to visualize and count the retina photoreceptors. Adaptive optics (AO) flood illumination retinal imaging allows for such a high resolution imaging and has now been used for more than a decade [1, 2, 3].

However, AO flood imaging suffers from an intrinsic limitation that decreases image quality and makes both automatic post-processing (photoreceptor counting, blood vessel diameter measurements...) and visual interpretation difficult: the three-dimensional nature of the object and of the imaging process. Indeed, information from both the in-focus plane and out-of-focus planes in front of and behind the in-focus plane contribute to the final 2D image, which creates an important background that reduces image contrast and leads to a loss in resolution.

A hardware solution to this problem is Adaptive Optics Scanning Laser Ophtalmoscopy [4] (AO-SLO): by using a confocal pinhole, one selects only the photons coming from a specific layer of the tissue under examination.

An alternative software solution for mitigating these effects without any setup modification is image deconvolution. Retinal image deconvolution is difficult for two reasons:

- imaging is fundamentally 3D and we only record 2D images. This aspect should be taken into account in the image model in order to enable a high-quality deconvolution ;
- the point spread function (PSF) is not well known, therefore we must estimate the PSF together with the object, a technique known as blind deconvolution.

In this paper, we focus on the imaging of the photoreceptor layer of the retina. In order to deal with the lack of information associated with recording only 2D images of a 3D object, we propose an imaging model in which the photoreceptor layer is assumed to be approximately shift invariant along the optical axis of the imaging system (*i.e.*, the photoreceptor size does not vary significantly over the depth of focus of the instrument and the photoreceptor are more or less parallel to the optical axis). We show that this hypothesis, although it is a simplifying one, is very effective on experimental AO retinal images with a visible and measurable effect on the lateral resolution of the images. Section 2 presents the imaging model and the PSF parameterization we will use. In Section 3, we describe the joint estimation of the object and the PSF before showing, both on simulation and theoretically, that it is not suited for our problem.

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

In Section 4, we derive a marginal estimator and show its performance on simulation. In Section 5, we show results of blind marginal deconvolution of experimental in vivo retinal images. Section 6 summarizes the results.

2. Imaging model and PSF parameterization

The object and the imaging process are both three-dimensional (3D). If we record a stack i_{3D} of 2D images focused at different depths in the object, a reasonable image formation model, after background subtraction, can be written as a 3D convolution:

$$\mathbf{i}_{3D} = \mathbf{h}_{3D} *_{3D} \mathbf{o}_{3D} + \mathbf{n}, \tag{1}$$

where $\mathbf{i}_{3D}$ is the 3D image, $\mathbf{o}_{3D}$ is the 3D object, $*_{3D}$ denotes the 3D convolution operator, $\mathbf{h}_{3D}$ is the 3D PSF and $\mathbf{n}$ is the noise.
We assume that our object is shift invariant along the optical axis:

$$o_{3D}(x,y,z) = o_{2D}(x,y)\,\alpha(z), \tag{2}$$

where $\alpha(z)$ is the normalized flux emitted by the plane at depth z ($\int \alpha(z)\mathrm{d}z = 1$).

Strictly speaking, this assumption means that our object must be shift invariant in z over an infinite range. However, in practice this invariance must only be verified over the depth of focus of the instrument ($\approx 50\mu m$ for an AO flood imager, $\approx 10-15\mu m$ for a confocal imager). Indeed, planes farther than the depth of focus from the image plane contribute to the image with a PSF that has a very narrow spectrum thus their contribution is almost a constant background.

In our case, we assume that the lateral size of the photoreceptors does not vary significantly and that the photoreceptors are almost parallel to the optical axis. Additionally, the depth of focus is about the length of a cone photoreceptor. Hence, the structures in front and behind the photoreceptor layer (pigment epithelium or inner retina layers) are way out of focus and only contribute as a background.

Current flood imaging systems only record data in one plane of interest. Using Eq. (1) and Eq. (2), it is easy to show that, in plane $z = 0$ for instance:

$$\begin{aligned} i(x,y) &\triangleq i_{3D}(x,y,z)|_{z=0} \\ &= \iiint \alpha(-z')\, h_{3D}(x',y',z')\, o_{2D}(x-x',y-y')\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z' + n(x,y) \\ &= (h_{2D} o_{2D} *_{2D})(x,y) + n(x,y)\,, \end{aligned} \tag{3}$$

with h_{2D} an *effective* 2D PSF which depends on the longitudinal brightness distribution of the object $\alpha(z)$ and on the 3D PSF:

$$h_{2D}(x,y) = \int \alpha(-z)\, h_{3D}(x,y,z)\,\mathrm{d}z.$$

The 2D image $i(x,y)$ at the focal plane of the instrument is the 2D convolution of a 2D object and a global PSF h which is the linear combination of the individual 2D PSFs (each one conjugated with a different plane of the object) weighted by the back-scattered flux at each plane.

After discretization and using Riemann sum to approximate the integral:

$$h_{2D}(x,y) \approx \sum_j \alpha_j\, h_j(x,y)\,, \tag{4}$$

with $h_j(x,y) \triangleq h_{3D}(x,y,z_j)$ the 2D lateral PSF at depth z_j and $\alpha_j = \alpha(z_j)\Delta z_j$ where Δz_j is the effective thickness of the jth layer. We define $\alpha = \{\alpha_j\}_j$ as the vector of unknowns that

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

parameterize the PSF. α is normalized ($\sum \alpha_j = 1$) and each parameter is positive ($\alpha_j \geq 0$). We search for $h_{2\mathrm{D}}$ as a linear combination of a basis of PSF's, each corresponding to a given plane. In the following, we consider short-exposure diffractive PSF's so that each h_j can be computed from the residual aberrations measured with a WFS and the knowledge of the defocus of plane z_j.

3. Joint estimation

There is a large body of work on blind deconvolution, originating in good part from astronomy (see, *e.g.* , Blanc-Féraud [5]). The conventional blind deconvolution approach is to perform an estimation of both the object and the PSF, jointly (see, *e.g.* , Ayers [6] for pioneering works and Mugnier [7] for more recent results on astronomical data).

3.1. Method

The joint estimation can be cast in a Bayesian framework as the computation of the joint maximum *a posteriori* (jmap) estimator:

$$(\hat{\mathbf{o}}, \hat{\alpha}) = \arg\max_{o,\alpha} p(\mathbf{i}, \mathbf{o}, \alpha; \theta) \quad (5)$$

$$= \arg\max_{o,\alpha} p(\mathbf{i}|\mathbf{o}, \alpha; \theta) \times p(\mathbf{o}; \theta) \times p(\alpha; \theta) \quad (6)$$

where, $p(\mathbf{i}, \mathbf{o}, \alpha; \theta)$ is the joint probability density of the data (**i**), of the 2D object (**o**), and of the PSF decomposition coefficients (α). It may depend on set of regularization parameters or hyperparameters (θ). $p(\mathbf{i}|\mathbf{o}, \alpha; \theta)$ is the likelihood of the data **i**, $p(\mathbf{o}; \theta)$ is the *a priori* probability density function of the object **o** and $p(\alpha; \theta)$ is the *a priori* probability density function of the coefficients α. In the following, we will not use any regularization on the set of coefficients α because we don not have any probability law for the PSF coefficients. However, since we only need to estimate a small number of these coefficients, this is not a problem.
The noise on the images is mainly photon noise which has a Poisson distribution. However, AO retinal images are dominated by a strong and quite homogeneous background. In the following, we will therefore assume that the noise is stationary white Gaussian with a variance σ^2. For the object, we choose a stationary Gaussian prior probability distribution with a mean value $\mathbf{o}_\mathrm{m}$ and a covariance matrix $\mathbf{R}_\mathrm{o}$. The set of hyperarameters is therefore $\theta = (\sigma^2, \mathbf{o}_\mathrm{m}, \mathbf{R}_\mathrm{o})$. Under these assumptions, we have:

$$p(\mathbf{i}, \mathbf{o}, \alpha; \theta) = \frac{1}{(2\pi)^{\frac{N^2}{2}} \sigma^{N^2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{i} - \mathbf{H}\mathbf{o})^t(\mathbf{i} - \mathbf{H}\mathbf{o})\right)$$
$$\times \frac{1}{(2\pi)^{\frac{N^2}{2}} \det(\mathbf{R}_\mathrm{o})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{o} - \mathbf{o}_\mathrm{m})^t \mathbf{R}_\mathrm{o}^{-1}(\mathbf{o} - \mathbf{o}_\mathrm{m})\right),$$

where **H** is the operator performing the convolution by the PSF **h**, $\det(x)$ is the determinant of matrix x and N^2 is the number of pixels in the image. $\hat{\mathbf{o}}$ and $\hat{\alpha}$ can therefore be defined as the estimated object and coefficients that minimize a criterion $J(\mathbf{o}, \alpha)$ defined as follows:

$$J_\mathrm{jmap}(\mathbf{o}, \alpha) = J_\mathrm{i}(\mathbf{o}, \alpha) + J_\mathrm{o}(\mathbf{o}, \alpha), \quad (7)$$

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

where $J_{\mathrm{i}}(\mathbf{o},\alpha) = -\ln p(\mathbf{i}|\mathbf{o},\alpha;\theta)$ (data-fidelity) and $J_{\mathrm{o}} = -\ln p(\mathbf{o};\theta)$ (regularization term). The criterion to be minimized reads:

$$\begin{aligned} J_{\mathrm{jmap}}(\mathbf{o},\alpha) &= -\ln p(\mathbf{i}|\mathbf{o},\alpha;\theta) - \ln p(\mathbf{o};\theta) \quad (8)\\ J_{\mathrm{jmap}}(\mathbf{o},\alpha) &= \frac{1}{2}N^2 \ln\sigma^2 + \frac{1}{2\sigma^2}(\mathbf{i}-\mathbf{H}\mathbf{o})^t(\mathbf{i}-\mathbf{H}\mathbf{o}) \\ &+ \frac{1}{2}\ln\det(\mathbf{R}_{\mathrm{o}}) + \frac{1}{2}(\mathbf{o}-\mathbf{o}_{\mathrm{m}})^t\mathbf{R}_{\mathrm{o}}^{-1}(\mathbf{o}-\mathbf{o}_{\mathrm{m}}) + C, \quad (9)\end{aligned}$$

where C is a constant. By cancelling the derivative of $J(\mathbf{o},\alpha)$ with respect to the object, we obtain an analytical expression of the object $\hat{\mathbf{o}}(\alpha;\theta)$ that minimizes the criterion for a given $(\alpha;\theta)$:

$$\hat{\mathbf{o}}(\alpha,\theta) = (\mathbf{H}^t\mathbf{H} + \sigma^2\mathbf{R}_0^{-1})^{-1}(\mathbf{H}^t\mathbf{i} + \sigma^2\mathbf{R}_0^{-1}\mathbf{o}_{\mathrm{m}}) \quad (10)$$

Since the matrices $\mathbf{H}$ (convolution operator) and $\mathbf{R}_{\mathrm{o}}$ (covariance matrix of an object with a stationary probability density) are Toeplitz-block-Toeplitz, we can write the joint criterion J_{jmap} and the analytical expression of the object $\hat{\mathbf{o}}(\alpha,\theta)$ in the Fourier domain with a circulant approximation:

$$\begin{aligned} J_{\mathrm{jmap}}(\mathbf{o},\alpha) = \frac{1}{2}N^2\ln S_{\mathrm{n}} + \frac{1}{2}\sum_{\nu}\frac{|\tilde{i}(\nu)-\tilde{h}(\nu)\tilde{o}(\nu)|^2}{S_{\mathrm{n}}} \\ + \frac{1}{2}\sum_{\nu}\ln S_{\mathrm{o}}(\nu) + \frac{1}{2}\sum_{\nu}\frac{|\tilde{o}(\nu)-\tilde{o}_{\mathrm{m}}(\nu)|^2}{S_{\mathrm{o}}(\nu)} \end{aligned} \quad (11)$$

$$\text{and } \hat{\tilde{\mathbf{o}}}(\alpha) = \frac{\tilde{h}^*(\nu)\tilde{i}(\nu) + \frac{S_{\mathrm{n}}}{S_{\mathrm{o}}(\nu)}\tilde{o}_{\mathrm{m}}(\nu)}{|\tilde{h}(\nu)|^2 + \frac{S_{\mathrm{n}}}{S_{\mathrm{o}}(\nu)}}, \quad (12)$$

where S_{n} is the noise power spectral density (PSD), S_{o} is the object PSD (the new set of hyperparameters in the Fourier domain is $\{S_{\mathrm{n}}, S_{\mathrm{o}}\}$), ν is the spatial frequency and $\tilde{x}$ denotes the two-dimensional Fast Fourier Transform of x.
$\hat{\tilde{\mathbf{o}}}(\alpha)$ is the estimated object after classical Wiener filtering of the image $\mathbf{i}$ and is easily computed.
If we substitute Eq. (12) into Eq. (11), we obtain a new expression of J_{jmap} that does not depend explicitly on the object:

$$\begin{aligned} J'_{\mathrm{jmap}}(\alpha) = \frac{1}{2}N^2\ln S_{\mathrm{n}} + \frac{1}{2}\sum_{\nu}\ln S_{\mathrm{o}}(\nu) \\ + \frac{1}{2}\sum_{\nu}\frac{1}{S_{\mathrm{o}}(\nu)}\frac{|\tilde{i}(\nu)-\tilde{h}(\nu)\tilde{o}_{\mathrm{m}}(\nu)|^2}{|\tilde{h}(\nu)|^2 + \frac{S_{\mathrm{n}}}{S_{\mathrm{o}}(\nu)}}. \end{aligned} \quad (13)$$

The joint MAP solution is thus the pair $(\hat{o}(\alpha),\alpha)$ for the value of α that minimizes Eq. 13.

3.2. Simulation results

The following simulation was performed to evaluate the performance of the joint estimator in our problem.

A simulated image is built in the following manner:

$$\mathbf{i} = (\alpha * \mathbf{h}_{\mathrm{foc}} + (1-\alpha)\mathbf{h}_{\mathrm{defoc}}) * \mathbf{o} + \mathbf{n}, \quad (14)$$

where

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

- the global PSF is the sum of only two weighted PSF's, the first one $\mathbf{h}_{\text{foc}}$ being focused and the second one $\mathbf{h}_{\text{defoc}}$ defocused. We assume that the focused PSF has no aberration (AO correction is perfect). The defocus is equal to π radian RMS;
- The object used is a 128×128 pixel portion of an experimental AO image obtained with the XV-XX retinal imager developed by the Observatoire de Paris [2];
- Noise $\mathbf{n}$ is stationary Gaussian with a standard deviation $\sigma = 0.01 * \max(\mathbf{o})$, corresponding roughly to photon noise for an average of 10000 photons/pixel;
- $\alpha = 0.3$.

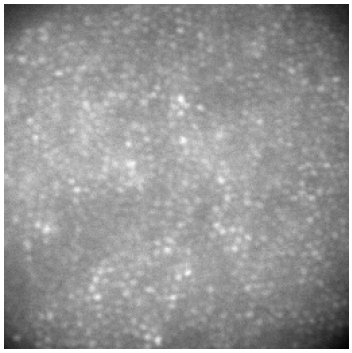


Fig. 1. Simulated object

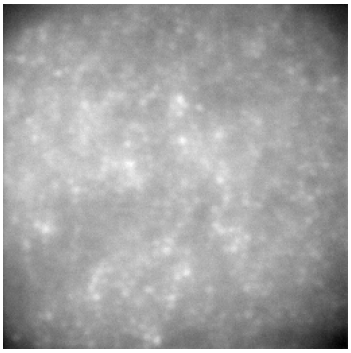


Fig. 2. Simulated image

We assume for the sake of this simulation that the object PSD S_{o} and the noise PSD S_{n} are known although it is not the case in practice. Therefore, we perform a so-called "supervised" estimation of α: we compute the joint criterion $J_{\text{jmap}}(\alpha; S_{\text{o}}, S_{\text{n}})$ (see Eq. 13)) for values of α ranging from 0 to 1 to find the value of α that minimizes the joint criterion. Figure 3 shows the result of such a computation.

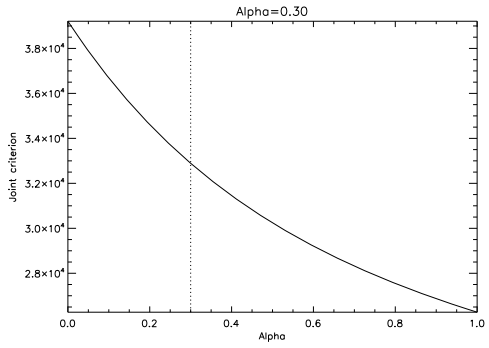


Fig. 3. Joint criterion for $0 \leq \alpha \leq 1$

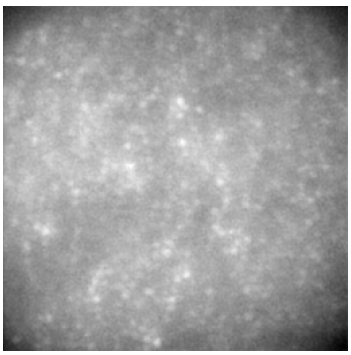


Fig. 4. Jointly estimated object

We see that the joint estimator is minimum for $\alpha = 1$ whereas the real value of α is 0.3. The joint estimation fails to retrieve the actual value even in this very simple case (two point spread functions, known hyperparameters). Figure 4 shows the restored object for the value of

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

α that minimizes $J_{\text{jmap}}(\alpha; S_\text{o}, S_\text{n})$. The image is poorly deconvolved since the estimated PSF is perfectly focused whereas the actual global PSF is only 30% focused.

The joint estimator does not work well in the case of myopic deconvolution of retinal images, even in the most simple case of only two PSF's with known S_n and S_o. Actually, a close look at Eq. (13) helps us understand why this joint estimator is actually degenerate in this case: if, for instance, the mean object we use to compute the joint criterion is constant ($\mathbf{o}_\text{m} = \beta$), and since the PSF and the set of parameters are both normalized, then the numerator does not depend on the set of parameters α. Minimizing J'_{jmap} is equivalent to maximizing this denominator, *i.e.*, to choosing the PSF with the highest MTF $|\tilde{h}|$, which is the most focused PSF.

One might wonder why the joint estimation, although known to show poor results for blind deconvolution [8] is actually quite used in other contexts such as astronomical imaging. The joint estimation works fairly well thanks to constraints such as PSF support or positivity (which effectively acts as an object support constraint) that help remove ambiguities between the object and the PSF. In our case, since we cannot use such constraints (the photoreceptor signal is superimposed over a strong background and the object extends far beyond the recorded field of view of the system), the joint estimator is not well suited for retinal image deconvolution.

Multi-frame joint deconvolution [9] can help since it increases the number of data for the same object but is only effective if the PSFs are different enough [10], such as in the case of phase diversity [11]. Therefore, another estimator with better statistical properties would be preferable, ideally capable of restoring the PSF on a single frame.

4. Marginal estimation

The poor results of the joint estimation led us to propose another estimator for our imaging problem. The estimator proposed is the marginal estimator, which has better properties [12] and has never been used previously in retinal images deconvolution although it has already been proposed in the literature in other contexts including estimation of aberrations by use of phase diversity [13]. The principle of marginal estimation is to integrate the object $\mathbf{o}$ out of the problem (*i.e.*, marginalize the posterior likelihood [5]). We integrate the joint probability of the object $\mathbf{o}$ and the PSF parameters α over all the possible values of object $\mathbf{o}$.

$$\hat{\alpha} = \arg\max_{\alpha} \int p(\mathbf{i}, \mathbf{o}, \alpha; \theta) \mathrm{d}\mathbf{o}. \tag{15}$$

Marginalization reduces the number of unknowns to be retrieved (from the total number of pixels of the image + the PSF parameters in the joint estimation case to just a few PSF parameters) and gives us a true maximum likelihood or maximum a posteriori (depending on the prior on the estimated parameters) estimator of the parameters of interest (namely, the PSF parameters). After estimation of the PSF parameters α, the object is restored by Wiener filtering of the image with the estimated global PSF and hyperparameters.

4.1. Marginal criterion

$$\hat{\alpha}_\text{ML} = \arg\max_{\alpha} p(\mathbf{i}, \alpha; \theta) = \arg\max_{\alpha} p(\mathbf{i}|\alpha; \theta) p(\alpha; \theta). \tag{16}$$

We keep the assumptions made for the joint estimation: a stationary white Gaussian noise with variance σ^2, stationary Gaussian prior probability distribution with a mean value $\mathbf{o}_\text{m}$ and covariance matrix $\mathbf{R}_\text{o}$ for the object. Since $\mathbf{i}$ is a linear combination of a Gaussian object and a Gaussian noise, it is also Gaussian. Its associated probability density reads:

$$p(\mathbf{i}|\alpha; \theta) = A(\det \mathbf{R}_\text{i})^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{i} - \mathbf{i}_\text{m})^t \mathbf{R}_\text{i}^{-1} (\mathbf{i} - \mathbf{i}_\text{m})\right), \tag{17}$$

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

where A is a constant, $\mathbf{R}_\mathrm{i}$ is the image covariance matrix and $\mathbf{i}_\mathrm{m} = \mathbf{H}\mathbf{o}_\mathrm{m}$. Since we only need to estimate a small number of parameters, there is no need to regularize the solution over α. We therefore use a Maximum Likelihood (ML) estimator rather than a Maximum *A Posteriori* (MAP) estimator.
Maximizing $p(\mathbf{i}|\alpha;\theta)$ is equivalent to minimizing the opposite of its logarithm:

$$J_\mathrm{ML}(\alpha) = \frac{1}{2}\ln\det(\mathbf{R}_\mathrm{i}) + \frac{1}{2}(\mathbf{i}-\mathbf{i}_\mathrm{m})^t\mathbf{R}_\mathrm{i}^{-1}(\mathbf{i}-\mathbf{i}_\mathrm{m}) + B \tag{18}$$

where B is a constant and $\mathbf{R}_\mathrm{i} = \mathbf{H}\mathbf{R}_\mathrm{o}\mathbf{H}^t + \sigma^2 I_d$ (I_d is the identity matrix). The marginal criterion can be written in the Fourier domain as follows:

$$\begin{aligned} J_\mathrm{ML}(\alpha) = \frac{1}{2}\sum_\nu \ln S_\mathrm{o}(\nu) + \frac{1}{2}\sum_\nu \ln\left(|\tilde{h}(\nu)|^2 + \frac{S_\mathrm{n}}{S_\mathrm{o}(\nu)}\right) \\ + \frac{1}{2}\sum_\nu \frac{1}{S_\mathrm{o}(\nu)} \frac{|\tilde{i}(\nu) - \tilde{h}(\nu)\tilde{o}_\mathrm{m}(\nu)|^2}{|\tilde{h}(\nu)|^2 + \frac{S_\mathrm{n}}{S_\mathrm{o}(\nu)}} + B'. \end{aligned} \tag{19}$$

Using Eq. 13 and Eq. 19, we obtain the following relationship between the marginal estimator and the joint estimator:

$$J_\mathrm{ML}(\alpha) = J'_\mathrm{jmap}(\alpha) + \frac{1}{2}\sum_\nu \ln\left(|\tilde{h}(\nu)|^2 + \frac{S_\mathrm{n}}{S_\mathrm{o}(\nu)}\right) - \frac{1}{2}N^2 \ln S_\mathrm{n}. \tag{20}$$

The marginal estimator is therefore very similar to the joint estimator in its mathematical expression (as shown by Goussard when the hyperparameters are known [14] and Blanc [13] for unknown hyperparameters. Its properties, as we will show in the following, are nevertheless significantly different to those of the joint estimator.

4.2. Marginal estimation results

We now present the results of the marginal estimation on simulated data. The same simulation as in the joint estimation was performed ($\mathbf{i} = (\alpha * \mathbf{h}_\mathrm{foc} + (1-\alpha)\mathbf{h}_\mathrm{defoc}) * \mathbf{o} + \mathbf{n}$, with $\alpha = 0.3$). The marginal criterion of Eq. (19) was computed for $0 \leq \alpha \leq 1$ with known hyperparameters. The object is restored by Wiener filtering. Results of the computation are shown on figure 5 and the restored object on figure 6.

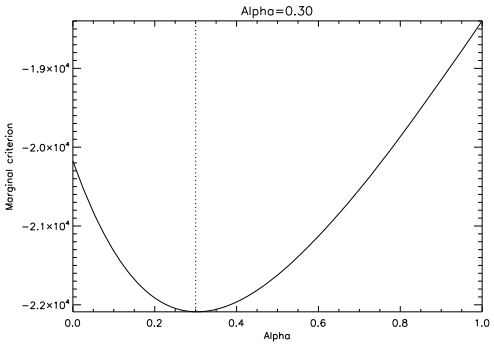


Fig. 5. Marginal criterion for $0 \leq \alpha \leq 1$

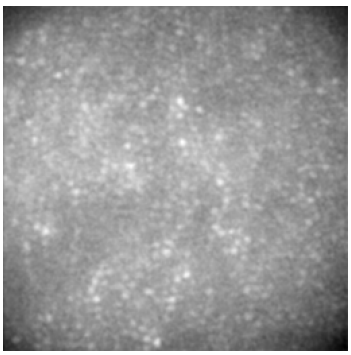


Fig. 6. Marginaly estimated object

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

Figure 5 shows that marginal criterion is minimum for $\alpha = 0.3$, which is the true value of α used in the simulation. The marginal estimator accurately estimates the parameter of interest in our simulation. As a result, the restored object, shown in Figure 6, is much sharper than the simulated image and much closer to the actual object used in the simulation than the object restored with the joint estimation.

The marginal estimator thus enables supervised myopic deconvolution of retinal images with our image model. In practice, we do not have access to the true noise value and the true PSD of the object we are observing. Fortunately, the marginal estimator allows us to estimate these PSDs together with the PSF coefficients, as shown in the next paragraph.

4.3. Hyperparameter estimation

The marginal estimator allows us to estimate the set of hyperparameters θ (actually the object PSD S_o and noise PSD S_n) together with the PSF coefficients in an automatic manner. This method is called unsupervised estimation:

$$(\hat{\alpha}, \hat{S_\mathrm{n}}, \hat{S_\mathrm{o}}) = \underset{\alpha, S_\mathrm{n}, S_\mathrm{o}}{\arg\max}\, f(\mathbf{i}, \alpha; S_\mathrm{n}, S_\mathrm{o}). \tag{21}$$

In order to reduce the number of hyperparameters we must estimate, we choose to model the object PSD S_o in the following way [15]:

$$S_\mathrm{o}(\nu) = \frac{k}{1 + \left(\frac{\nu}{\nu_0}\right)^p}. \tag{22}$$

Such PSD parametrization has been successfully used in various imaging applications such as astronomical imaging or earth observation from satellites. Since the noise is assumed to be Gaussian and homogeneous, $S_\mathrm{n} =$ constant. The criterion $J_\mathrm{ML}(\alpha)$ becomes $J_\mathrm{ML}(\alpha, S_\mathrm{n}, k, \nu_\mathrm{o}, p)$ and must now be minimized versus the PSF coefficients α and the hyperparameters $S_\mathrm{n}, k, \nu_\mathrm{o}$ and p.

With the change of variable $\mu = S_\mathrm{n}/k$, if we cancel the derivative of the criterion with respect to k, we obtain an analytical expression for $\hat{k}(\alpha, \mu, \nu_\mathrm{o}, p)$ that minimizes the criterion for a given value of the other parameters therefore only four hyperparameters remain, $\hat{\mu}$, $\hat{\nu_0}$, $\hat{p}$ and S_n [13].

PSF coefficients and hyperparameters are estimated in an alternate way. We initialize the algorithm with a perfectly focused global PSF.

There is no analytical expression for the minimum value of the criterion so the minimization has to be done numerically. In our case, the minimization is performed with a Variable Metric with Limited Memory, Bounded (VMLM-B) method developed by E. Thiébaut [16].

4.4. Asymptotic properties

The maximum likelihood estimator is known to be a consistent estimator, *i.e.*, it tends toward the actual values of the estimated parameters as the noise tends toward zero or as the size of data tends toward infinity. It is also known to be asymptotically normal [12] so that the neg-log-likelihood is asymptotically quadratic thus convex.

Extensive simulations were performed to validate the statistical behavior of our unsupervised marginal estimation. The simulation conditions are the same as previously:

- $\mathbf{i} = (\alpha * \mathbf{h}_\mathrm{foc} + (1-\alpha)\mathbf{h}_\mathrm{defoc}) * \mathbf{o} + \mathbf{n}$, with $\alpha = 0.3$;
- Noise RMS varies from 1% of the maximum value of the image to 20% of the maximum value of the image;

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

- 50 noise realizations were computed for each noise RMS value;
- The simulation was performed on 3 different subimages varying in size: a 32×32 pixel central region of image **i**, a 64×64 pixel central region of image **i** and the whole 128×128 pixel image **i**.

Figure 7 shows the RMS error on estimation of the PSF coefficients for the different values of noise and the varying data size, both in the supervised and unsupervised cases.
For a given data size and both in the supervised and unsupervised estimation, the marginal estimator RMS error tends towards zero (*i.e.*, the estimated parameters α tends towards the exact value) when noise decreases. Even more interestingly, for a given noise value, error tends towards zero as the size of data increases. In particular, for a 128×128 pixel image and for noise $\sigma = 5\%$ of the max value of the image, the RMS error on the PSF coefficient α estimation is less than 3%. For a noise RMS value of 1% of the image maximum, the unsupervised estimator basically shows the same performance as the supervised estimation.

This simulation shows that the unsupervised marginal estimator exhibits, in practice, its appealing theoretical properties, which opens the way to its use on experimental images.

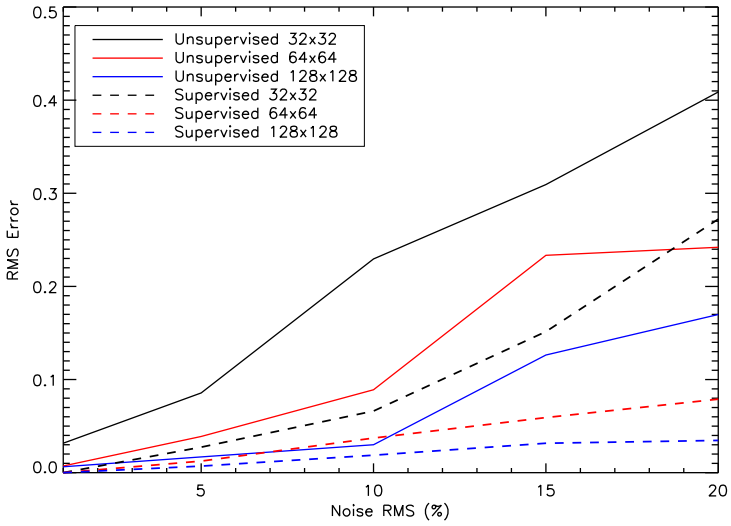


Fig. 7. RMS error on the estimation of the PSF coefficients as a function of noise level in percent (noise standard deviation over image maximum). The black, red and blue lines correspond, respectively, to 32×32, 64×64 and 128×128 pixels images. Supervised case is in dashed lines, unsupervised case in solid line.

5. Preliminary experimental results

We now show experimental results of the marginal blind deconvolution on *in vivo* retinal images:

- The experimental image (Figure 8) is a 256×256 pixel image recorded on a healthy subject with the AO eye-fundus imager of the Center for Clinical Investigation of the Quinze-Vingts Hospital in Paris, developed by the Observatoire de Paris-Meudon [2];

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

- We model the global PSF as a linear combination of 3 PSFs, the first one being focused, the second one being defocused with a focus $\phi_1 = \pi/2$ rad RMS and the third one being defocused by $\phi_2 = \pi$ rad RMS.
- We assume that the adaptive optics has perfectly corrected the wavefront and that the focused PSF is a Airy disk.

We must estimate $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$.

The unsupervised marginal estimation gives $\alpha = \{0.24, 0.22, 0.54\}$, the resulting estimated PSF is shown on Figure 9. For this image, the main contribution (more than half of the energy) comes from the most out-of-focus plane and only a little less than 25% of the energy comes from the in-focus-plane.

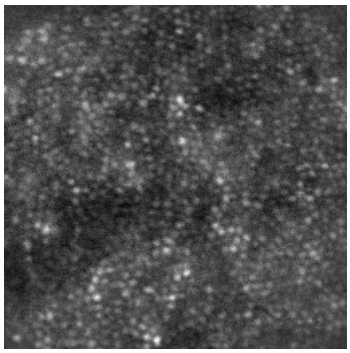


Fig. 8. Experimental image

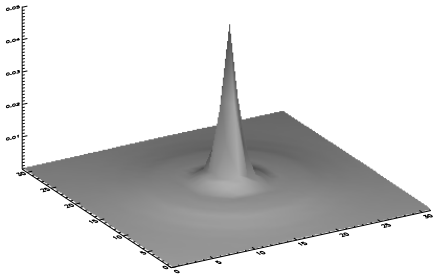


Fig. 9. Estimated PSF

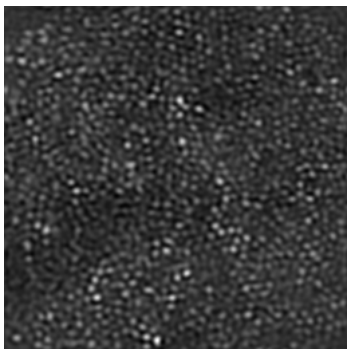


Fig. 10. Restored object

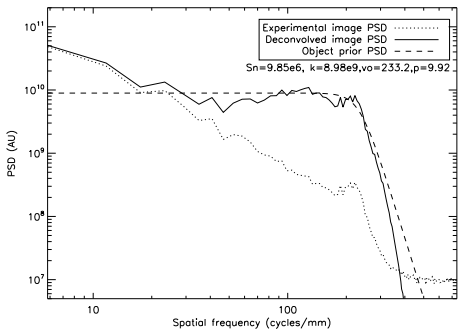


Fig. 11. PSD comparison between experimental image (dotted line), restored object (solid line) and object prior PSD (dashed line) with the estimated hyperparameters

The object is computed thanks to Eq. (12). Figure 10 shows the restored object after unsupervised marginal estimation. It is clearly visible that the restored object is much sharper than

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

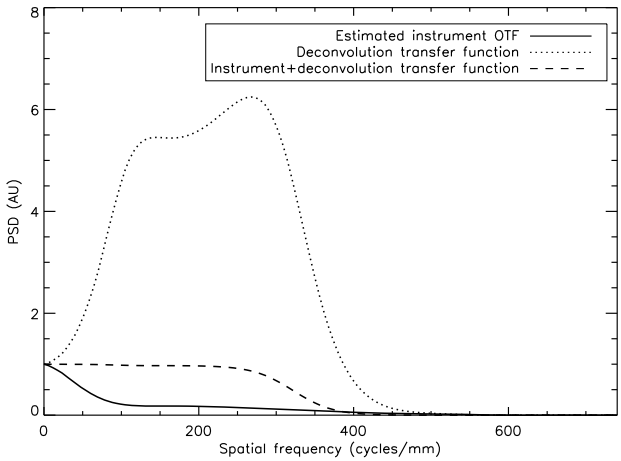


Fig. 12. Radial average of the estimated instrument optical transfer function, deconvolution transfer function and global (instrument+deconvolution) transfer function.

the original image. The photoreceptors have a much better contrast and can be seen clearly throughout the image. The restored object also is much less noisy than the original image.

Figure 11 shows a comparison of the power spectral densities of the experimental image and of the restored object: we can clearly see an improvement at the medium-high spatial frequency, with an improvement of a factor of 5 around 200-250 cycles/mm, *i.e.*, the spatial frequency corresponding to the cone photoreceptor size and separation. This frequency enhancement is clearly visible on Figure 12 that shows, in solid line, the estimated Optical Transfer Function (OTF) of our instrument (AO Flood imager+eye) as well as the deconvolution transfer function (dotted line) and the global (instrument+deconvolution) transfer function (dashed line). The deconvolution restores the spatial frequencies damped by the instrument transfer function up to 300 cycles/mm, a frequency that is beyond the spatial frequency of the cone photoreceptors in our image. These preliminary results show that our image model and the marginal estimator are well adapted to the deconvolution of adaptive optics corrected photoreceptor images. Motion artifacts due to eye movement during image acquisition and resulting in blurred images could possibly be addressed by changing the PSF basis to include motion induced PSFs and not only purely diffractive PSFs.

6. Conclusion

Blind deconvolution is a much needed tool for the interpretation and further processing of AO-corrected retinal images. We have proposed a reasonable imaging model to deal with the problem of only having 2D images of a 3D object that results in a useful PSF parameterization. We have shown, both analytically and through simulations, that the classical blind joint estimation of object and PSF is not suited for this problem. We have derived a marginal estimator of the PSF and extended it to estimate also the hyperparameters (object and noise PSDs), *i.e.*, to perform an unsupervised estimation. We have showed on simulations that this estimator is capable of restoring the PSF accurately even in the unsupervised case. The good statistical properties of the unsupervised marginal estimation have also been demonstrated.

Finally, we have shown preliminary results on experimental data, showing the efficiency

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

of the marginal estimator for myopic deconvolution of adaptive optics retinal images, with a measurable improvement of the contrast at the spatial frequencies corresponding to the cone photoreceptors. Although developed in the context of AO flood illumination retinal imaging, this marginal blind deconvolution method could also be applicable to other kinds of data such as confocal retinal imaging or more general microscopy data, or even astronomical data, mainly by changing the PSF basis used.

Acknowledgements

The authors are grateful to Marie Glanc for her support and her guidance on the use of the LESIA AO bench she designed and built.

#150979 - $15.00 USD Received 12 Jul 2011; revised 16 Sep 2011; accepted 17 Sep 2011; published 1 Nov 2011

Annexe D

Article Mocœur et coll., 2009 [A9]

November 15, 2009 / Vol. 34, No. 22 / OPTICS LETTERS 3487

Analytical solution to the phase-diversity problem for real-time wavefront sensing

Isabelle Mocœur,[1,2,*] Laurent M. Mugnier,[1] and Frédéric Cassaing[1]

[1]*Office National d'Etudes et de Recherche Aérospatiales, Optics Department, BP 72, 92322 Châtillon Cedex, France*
[2]*Centre National d'Etudes Spatiales, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 4, France*
[*]*Corresponding author: mocoeur_astro@orange.fr*

Received May 26, 2009; revised October 3, 2009; accepted October 8, 2009;
posted October 14, 2009 (Doc. ID 111870); published November 5, 2009

High-resolution optical systems require a very accurate control of the optical paths. For the measurement of aberrations on extended objects, several iterative phase-diversity algorithms have been developed, based on aberration estimation from focal-plane intensity measurements. Here we present an analytical estimator in the case of small aberrations. Under this assumption, a quadratic criterion is derived that allows us to express the solution (phase and object) under a simple analytical form. We also compare the performance of our algorithm with the iterative phase diversity, demonstrating that the analytic estimator is appropriate for closed-loop operation. © 2009 Optical Society of America

OCIS codes: 010.7350, 100.3190, 110.5100, 100.1830, 100.3020.

The image $\boldsymbol{i}$ recorded at the focal plane of an instrument is modeled by the discrete and noisy convolution of the point-spread function (PSF) $\boldsymbol{h}$ with the observed object $\boldsymbol{o}$, which defines the mathematical model of the data. Generally, the PSF is degraded by aberrations $\boldsymbol{a}=\{a_1,\cdots,a_k\}$ such as turbulence; therefore, these aberrations must be estimated in order to be corrected. The phase-diversity technique [1] uses simultaneous acquisition of the focal-plane image and at least a second image differing by a known set of aberrations $\boldsymbol{a_d}$, conventionally a small defocus. In this Letter, we consider two images, $\boldsymbol{i_1}$ and $\boldsymbol{i_2}$; the first one is acquired at the focal plane, whereas the second one is obtained in a plane defocused by a distance d. The corresponding PSFs are then given by $\boldsymbol{h_1}=\boldsymbol{h}(\boldsymbol{a})$ and $\boldsymbol{h_2}=\boldsymbol{h}(\boldsymbol{a}+\boldsymbol{a_d})$.

To retrieve both the aberrations and the object that are most compatible with the measurements, Gonsalves [1] first proposed to use a least-square approach; an extension of this method is the joint maximum *a posteriori* estimation of $(\hat{\boldsymbol{a}},\hat{\boldsymbol{o}})$ that are most compatible with the measurements by using statistical information on the data [2,3]. If the noise is assumed to be a stationary white Gaussian distribution with constant variance σ^2, the joint criterion J_{I} to be minimized can be written in the Fourier domain as

$$J_{\mathrm{I}}(\boldsymbol{o},\boldsymbol{a})=N_\nu\ln\sigma^2+\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}\sum_{d=1}^{2}|\tilde{\boldsymbol{i}}_{\boldsymbol{d}}(\nu)-\tilde{\boldsymbol{h}}_{\boldsymbol{d}}(\boldsymbol{a},\nu)\tilde{\boldsymbol{o}}(\nu)|^2+\sum_{\nu=1}^{N_\nu}\frac{|\tilde{\boldsymbol{o}}(\nu)-\tilde{\boldsymbol{o}}_{\boldsymbol{m}}(\nu)|^2}{2S_o(\nu)}+\mathrm{R}(\boldsymbol{a}),\tag{1}$$

where $\tilde{}$ denotes Fourier transformation and where $\tilde{\boldsymbol{h}}$, which is the Fourier transform of $\boldsymbol{h}$, is the optical transfer function (OTF); N_ν is the number of pixels in the image; and d is the dth diversity plane (1 for focal plane, 2 for extrafocal); the two last terms can be used to introduce possible prior knowledge on the aberrations and/or on the object: S_o is the power spectral density model of $\boldsymbol{o}$, $\boldsymbol{o_m}$ is the mean object, and $\mathrm{R}(\boldsymbol{a})$ is the phase regularization term. Here we choose $\boldsymbol{o_m}=0$; we take $\mathrm{R}(\boldsymbol{a})=0$, which means that we do not regularize the estimation of the aberrations explicitly. An implicit regularization is achieved by expanding the phase on a finite, small set of Zernike polynomials and is enough to obtain good results for the considered noise levels.

Criterion J_{I} must be minimized with respect to both the object $\boldsymbol{o}$ and the aberrations $\boldsymbol{a}$. However, while the object is unknown, it can be estimated for given aberrations,

$$\hat{\tilde{\boldsymbol{o}}}(\boldsymbol{a},\nu)=\frac{\tilde{\boldsymbol{i}}_{\boldsymbol{1}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{1}}^{*}(\boldsymbol{a},\nu)+\tilde{\boldsymbol{i}}_{\boldsymbol{2}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{2}}^{*}(\boldsymbol{a},\nu)}{|\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\tilde{\boldsymbol{a}},\nu)|^2+|\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)|^2+\dfrac{\sigma^2}{S_o(\nu)}},\tag{2}$$

where * denotes complex conjugate. We set $\sigma^2/S_o=\varepsilon$ close to 0, which means we under-regularize the inversion, because it has been shown [4] that doing so leads to a consistent estimator for the aberrations. Then, by introducing the estimated object of Eq. (2) into Eq. (1), we obtain a criterion that explicitly depends on the aberrations only,

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}\frac{|\tilde{\boldsymbol{i}}_{\boldsymbol{1}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)-\tilde{\boldsymbol{i}}_{\boldsymbol{2}}(\nu)\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\boldsymbol{a},\nu)|^2}{|\tilde{\boldsymbol{h}}_{\boldsymbol{1}}(\boldsymbol{a},\nu)|^2+|\tilde{\boldsymbol{h}}_{\boldsymbol{2}}(\boldsymbol{a},\nu)|^2+\varepsilon}+\mathrm{Cst}.\tag{3}$$

To derive the aberrations $\boldsymbol{a}$, J_{II} is usually minimized by an iterative gradient-based method. But although iterative estimators are optimal in terms of performance [5], they are time consuming, since each iteration costs $2N_d$ FFTs, where N_d is the number of diversity planes (here, $N_d=2$).

During the past 15 years, efforts have been made toward noniterative algorithms: first by proposing better numerical algorithms [6,7], then by modifying the criterion used to estimated the aberrations from the data [8,9]. However, none of these methods is

0146-9592/09/223487-3/$15.00

truly single iteration, since each of them requires at least two iterations to converge [8,10].

To derive an analytical solution for the aberrations, we use the small phase assumption in the two following ways [11]:

- We consider the denominator of Eq. (3) as a weighting term, at $\boldsymbol{a}=0$, for instance (or at the last estimate $\boldsymbol{a_0}$ for $\boldsymbol{a}$).
- We linearize the expression of the OTF in each diversity plane by a first-order Taylor expansion, obtaining an affine expression of $\tilde{\boldsymbol{h}}_1$ and $\tilde{\boldsymbol{h}}_2$ as a function of $\boldsymbol{a}$ for each frequency ν,

$$\tilde{h}_d(\boldsymbol{a},\nu)=\boldsymbol{\alpha}_d(\nu)\boldsymbol{a}+\beta_d(\nu), \tag{4}$$

where $\boldsymbol{\alpha}_d(\nu)$ is a row vector of size $k_{\max}$ and $\beta_d(\nu)$ is a scalar. The differentiation is done at $\boldsymbol{a}=0$ or at $\boldsymbol{a_0}$ for each of these OTFs. So $\beta_d(\nu)=\tilde{h}_d(\boldsymbol{a_0},\nu)$, $\boldsymbol{\alpha}_d(\nu)^T=\partial h_d/\partial a(\boldsymbol{a_0},\nu)$, where ^{T} denotes transposition.

We obtain a new expression of the criterion that is quadratic and can be written as

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\sum_{\nu=1}^{N_\nu}|\boldsymbol{A}(\nu)\boldsymbol{a}-B(\nu)|^2+\mathrm{Cst}, \tag{5}$$

with

$$\boldsymbol{A}(\nu)=\frac{\tilde{i}_2(\nu)\boldsymbol{\alpha}_1(\nu)-\tilde{i}_1(\nu)\boldsymbol{\alpha}_2(\nu)}{\sqrt{|\tilde{h}_1(0,\nu)|^2+|\tilde{h}_2(0,\nu)|^2+\varepsilon}}, \tag{6}$$

$$B(\nu)=\frac{-\tilde{i}_2(\nu)\beta_1(\nu)+\tilde{i}_1(\nu)\beta_2(\nu)}{\sqrt{|\tilde{h}_1(0,\nu)|^2+|\tilde{h}_2(0,\nu)|^2+\varepsilon}}. \tag{7}$$

We define the matrix $\boldsymbol{A}$ of size $N_\nu\times k_{\max}$ as the stack of all row vectors $\boldsymbol{A}(\nu)$ of Eq. (6). Similarly, we define vector $\boldsymbol{B}$ of size N_ν as the stack of the scalar $B(\nu)$ of Eq. (7). Equation (5) can be then rewritten as

$$J_{\mathrm{II}}(\boldsymbol{a})=\frac{1}{2\sigma^2}\|\boldsymbol{A}\boldsymbol{a}-\boldsymbol{B}\|^2+\mathrm{Cst}. \tag{8}$$

To minimize $J_{\mathrm{II}}(\boldsymbol{a})$, its gradient is derived with respect to the aberrations, which leads to a linear equation depending on $\boldsymbol{a}$. The resulting aberration vector is given by

$$\hat{\boldsymbol{a}}=[\Re(\boldsymbol{A}^H\boldsymbol{A})]^\dagger\Re(\boldsymbol{A}^H\boldsymbol{B}), \tag{9}$$

with $\Re$ the real part operator and $\dagger$ the generalized inverse of a matrix.

The resulting algorithm is much faster than the iterative one, requiring only $N_d=2$ FFTs. The inversion of the $\Re(\boldsymbol{A}^H\boldsymbol{A})$ matrix is not critical, since it is a square matrix of size $k_{\max}\ll N_\nu$.

Once the phase has been computed, the object can be then restored as well by introducing $\hat{\mathbf{a}}$ in Eq. (2).

To study the properties of the analytical estimator, we consider the specific low-order aberrations of a phased-array optical system, which are the positioning errors between the apertures, namely, the pistons and tip/tilts. To solve the inverse problem, we suppose that the instrument pupil is composed by N_T identical apertures. Each aperture n has a complex transmission $\boldsymbol{p_n}$, its phase being expanded on an orthonormalized basis, here a set of $k_{\max}$ scaled Zernike polynomials $\boldsymbol{Z_k}$,

$$\boldsymbol{p}_n(\boldsymbol{u})=\Pi(\boldsymbol{u})\exp\left[j\sum_{k=1}^{k_{\max}}a_{kn}\boldsymbol{Z_k}(\boldsymbol{u})\right], \tag{10}$$

where its modulus is described by the disk function Π,

$$\Pi(\boldsymbol{u})=\begin{cases}1 & \text{for } 0\leqslant|\boldsymbol{u}|\leqslant R_n\\ 0 & \text{elsewhere}\end{cases}. \tag{11}$$

In Eq. (10), $j^2=-1$ and a_{kn} is the rms amplitude of the kth mode over the nth subaperture. The corresponding aberration vector $\boldsymbol{a}$ is of size $N_Tk_{\max}$. As we deal only with piston ($k=1$) and tip/tilt ($k=2,3$), we will consider $k_{\max}=3$ in the following.

Two kinds of tests are reported here with two observation planes, one at $\boldsymbol{a_d}=0$ and the other with a 1 rad rms defocus. The object is an urban scene; monochromatic images of size $N_{\mathrm{pix}}=256\times256$ pixels are simulated with photon noise plus a read-out noise of 10 electrons per pixel and are sampled at the Shannon rate. The signal-to-noise ratio (SNR) is defined as the ratio between the mean value of the photon number and the noise standard deviation (per pixel).

Using three apertures in an equilateral configuration, we first consider a piston linearity test, applying at high flux (SNR=87) a 51-point ramp of $[-2\pi,+2\pi]$ on a given subaperture. One pair of the corresponding images is represented in Fig. 1.

The graph in Fig. 2 compares, for the aberrated subaperture, the piston estimated by our analytical method and the iterative algorithm with the introduced piston. First, we note that the reconstruction made with the analytic algorithm is excellent between $[-\pi/2;\pi/2]$ with an accuracy below $\lambda/60$ and a bias almost zero ($<10^{-3}$ rad) at the origin. In addition, the piston is reconstructed between ±1 rad with slope coefficients equal to 1 whatever the algorithm considered. Beyond 1 rad, where the small phase assumption is no longer valid, the bias increases rapidly. We also note that near $a=\pm\pi$, a wrapping occurs; since our imaging model is monochromatic, phase information is obtained only modulo 2π.

To test phase estimation in a case of small phase perturbation, we apply a set of random piston and tip/tilt listed in Table 1 to all subapertures (where to-

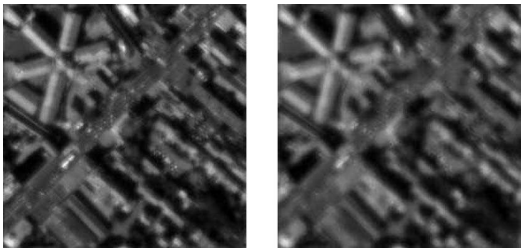


Fig. 1. Focal (left) and extrafocal (right) images obtained when a piston of 1 rad is applied on a subaperture.

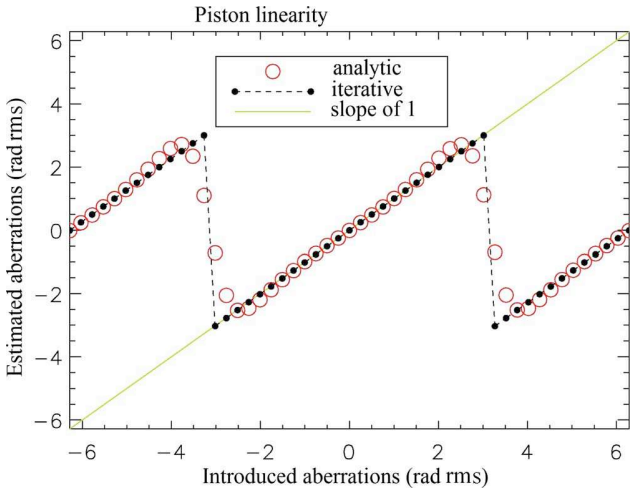


Fig. 2. (Color online) Piston linearity obtained when a piston ramp is applied on an subaperture.

tal absolute amplitude is equal to 0.53 rad rms, i.e., $\lambda/12$) for different levels of source brightness. For each data set, increasing fluxes are considered, ranging from $N_{\mathrm{bph}}=3\times10^6$ photoelectrons (phe$^-$) per image (SNR=4) to 3×10^{10} phe$^-$ (SNR=677). For each level, a data set of fifty images is simulated.

Figure 3 presents the total rms error obtained by both estimators versus the level of source brightness. Between 3×10^6 phe$^-$ and 1×10^9 phe$^-$, the two algorithms present the same behavior, following a law in $1/\sqrt{N_{\mathrm{bph}}}$. For higher flux, the error associated with the analytic algorithm remains constant around 2×10^{-3} rad. This observation is not surprising, since the affine approximation of the OTF is valid only for aberrations close to zero. The plateau we observe is then due to the inherent approximation of the linearization we made. However, the aberration estimations accuracy in the photon-noise regime remains better than 3×10^{-2} rad or $\lambda/200$ (equivalent to 10% of the total absolute amplitude introduced in piston), which is suitable for many applications.

Table 1. Aberrations Applied over the Configuration (rad rms)

	Piston	Tip	Tilt
Subaperture 1	−0.084	−0.017	0.019
Subaperture 2	0.161	−0.018	0.065
Subaperture 3	−0.055	0.019	0.092

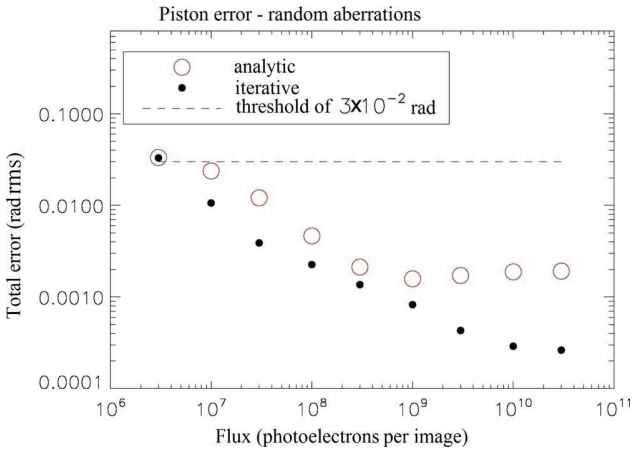


Fig. 3. (Color online) Error estimated over the subapertures when a set of random aberrations is applied.

As a conclusion, we have developed, under the small phase assumption, a noniterative focal plane algorithm for the phase diversity wavefront sensor that requires only 2 FFTs. This algorithm can typically estimate piston aberrations up to $|\pi/2|$ rad rms with an error of $\lambda/60$ and can be used for real-time correction of phase disturbances.

References

1. R. A. Gonsalves, Opt. Eng. **21**, 829 (1982).
2. O. M. Bucci, A. Capozzolo, and G. D'Elia, J. Opt. Soc. Am. A **16**, 1759 (1999).
3. A. Blanc, L. M. Mugnier, and J. Idier, J. Opt. Soc. Am. A **20**, 1035 (2003).
4. J. Idier, L. M. Mugnier, and A. Blanc, IEEE Trans. Image Process. **14**, 2107 (2005).
5. L. M. Mugnier, A. Blanc, and J. Idier, in *Advances in Imaging and Electron Physics*, P. Hawkes, ed. (Elsevier, 2006), Vol. 141, pp. 1–76.
6. M. G. Löfdahl and A. L. Duncan, Proc. SPIE **3353**, 952 (1998).
7. C. R. Vogel, T. Chan, and R. Plemmons, Proc. SPIE **3353**, 994 (1998).
8. G. B. Scharmer, in *High Resolution Solar Physics: Theory, Observations and Techniques*, T. R. Rimmele, K. S. Balasubramaniam, and R. R. Radick, eds., ASP Conferences Series (1999), Vol. 183, pp. 330–341.
9. M. G. Löfdahl and G. B. Scharmer, Proc. SPIE **4013**, 737 (2000).
10. J. Seldin, R. Paxman, V. Zarifis, L. Benson, and R. Stone, Proc. SPIE **401**, 48 (2000).
11. I. Mocœur, L. M. Mugnier, and F. Cassaing, in *21ième Colloque GRETSI* (2007).

Annexe E

Article Mugnier et coll., 2009 [A10]

1326 J. Opt. Soc. Am. A/Vol. 26, No. 6/June 2009 Mugnier *et al.*

Optimal method for exoplanet detection by angular differential imaging

Laurent M. Mugnier,[1,2,*] Alberto Cornia,[1,2,3] Jean-François Sauvage,[1,3] Gérard Rousset,[2,3] Thierry Fusco,[1,3] and Nicolas Védrenne[1,3]

[1]*ONERA/DOTA (Département d'Optique Théorique et Appliquée), B.P. 72, 92322 Châtillon cedex, France*
[2]*LESIA (Laboratoire d'Etudes des Systemes Informatiques et Automatiques), Observatoire de Paris, 5 place Jules Janssen, 92195 Meudon, France*
[3]*Groupement d'Intérêt Scientifique PHASE (Partenariat Haute résolution Angulaire Sol Espace) between ONERA, Observatoire de Paris, CNRS, and Université Paris Diderot*
**Corresponding author: mugnier@onera.fr*

Received September 16, 2008; revised February 13, 2009; accepted March 24, 2009; posted April 9, 2009 (Doc. ID 101668); published May 12, 2009

We propose a novel method for the efficient direct detection of exoplanets from the ground using angular differential imaging. The method combines images appropriately, then uses the combined images jointly in a maximum-likelihood framework to estimate the position and intensity of potential planets orbiting the observed star. It takes into account the mixture of photon and detector noises and a positivity constraint on the planet's intensity. A reasonable detection criterion is also proposed based on the computation of the noise propagation from the images to the estimated intensity of the potential planet. The implementation of this method is tested on simulated data that take into account static aberrations before and after the coronagraph, residual turbulence after adaptive optics correction, and noise. © 2009 Optical Society of America

OCIS codes: 100.3190, 110.3010, 110.2970, 110.1080.

1. INTRODUCTION

The direct detection of exoplanets from the ground is a very promising field of astronomy today [1]. A goal is the characterization of the physical composition of the exoplanets by a spectral analysis of their emitted and/or reflected light. This observation from the ground is a technological challenge. Indeed, in order to be able to observe a sufficient number of targets, it is required to cope with an intensity ratio (also called contrast) between the star and its planet that may be as high as 10^6 in IR bands [2] at very small angular separations. Two consortia are currently building planet searchers based on direct imaging in the near-IR: SPHERE (Spectro-Polarimetry High-contrast Exoplanet Research) [3] for the Very Large Telescope of the European Southern Observatory (VLT) (ESO) and GPI (Gemini Planet Imager) [4] for GEMINI observatory. The goal of the European project SPHERE is to detect giant planets orbiting nearby stars up to 100 pc from the Sun. For instance these planets may present an atmosphere rich in methane [5], with interesting spectral signatures around 1.6 μm to be used in spectral imaging. The planets searched for have orbits typically between 5 and 100 astronomical units. This requires being able to resolve an angular separation between the planet and its parent star that can be as small as a few diffraction elements (λ/D).

The SPHERE instrument is a combination of several optical features, all of them optimized toward the final goal of exoplanet detection. First of all, an extreme adaptive optics system (XAO) concentrates the light into a coherent Airy pattern, performing a real-time correction of Earth's atmospheric turbulence [6]. Then, the coronagraphic stage strongly attenuates the star intensity and therefore significantly reduces the photon noise. The coronagraphs considered in the SPHERE project are a Lyot coronagraph [7], a four-quadrant phase mask [8], and an apodized Lyot coronagraph [9].

The final optical quality is a key factor in direct exoplanet detection, because the main limitation for the detection of faint objects is demonstrated to be the static speckles in the coronagraphic images of the star [10]. These speckles are the consequence of an imperfect correction of static aberrations before the coronagraph. Note that because we are dealing with long-exposure images, the turbulence residuals are averaged and form a halo around the position of the star in the focal plane; the speckles are due only to static aberrations.

The combination of XAO and a coronagraphic device is necessary to reduce both photon noise and speckle in the final image, but is not sufficient for the considered intensity ratios between star and planet. To attain the detection performance needed to detect a large number of planets, it is mandatory to combine the above-mentioned optical devices with an *a posteriori* processing of all the data. The main problem is to disentangle the potential planet signal from the quasi-static speckles, that are due to static aberrations and constitute a major noise source. These speckles present the same characteristic angular size as the planet signal λ/D. With no more information, it is impossible to discriminate between the speckles and the planet. In order to do so, the SPHERE instrument includes the ability to perform spectral and angular differential imaging.

Spectral differential imaging consists of acquiring si-

multaneous images of the star–planet system at different wavelengths [11,12]. The spectral signature of the exoplanet's atmosphere ensures that the planet signal will significantly vary between spectral images taken in two slightly different spectral bands, while the star signal and therefore the speckles remain the same. A subtraction of two such images brings a significant attenuation of the star signal while enhancing the planet signal. With the IRDIS (Infra-Red Dual-beam Imaging and Spectroscopy) instrument [13] of SPHERE, one can make use of two close spectral channels, e.g., between bands $H2=1.59$ μm and $H3=1.64$ μm, the latter corresponding to a methane absorption line.

Angular differential imaging is a method originally designed for the calibration of static speckles for the Hubble Space Telescope—see, e.g., [14,15] and references therein. The original idea was to perform a rotation of the entire telescope, and therefore of the observed field on the detector, while the telescope point-spread function (PSF), which is the star signal and includes the static speckles, would remain the same. This idea has been developed recently in the case of a ground-based observation [16,17].

For instance at the Nasmyth foci of an alt–az mount telescope, both field and pupil rotate during the tracking of the target in the sky. The observer has the choice of implementing a de-rotation of the field image or of the pupil one. For angular differential imaging, we choose a de-rotation of the pupil image, thus ensuring the best temporal stability of the quasi-static speckles. In an image series obtained in these conditions with the star on-axis, the additional information we have at hand is therefore the induced circular trajectory of the planet through the images of the series.

In Section 2, we present our processing method, which aims at exploiting the field rotation in a ground-based imager dedicated to exoplanet detection. Note that the method could actually also be used to detect weak companion stars. The estimation of the planet position and intensity is done on the differential data, through a maximum-likelihood approach as presented in Section 3. An associated detection criterion is proposed in Section 4, and the complete method is tested in Section 5. Section 6 concludes the paper.

2. PRINCIPLE OF THE PROPOSED METHOD

If both angular and spectral data are available, as is the case with the SPHERE instrument, then one may first combine each pair of simultaneous spectral images into one image so as to enhance the planet signal by a partial suppression of the static speckles of the star signal, and then use the resulting combined image series as angular data.

In this paper, we do not consider that we record simultaneous spectral images. We assume that we have a series of angular images, and we investigate their joint processing. These images correspond to different recording times and therefore to different positions of the planet signal in the focal plane due to the field rotation. At least two approaches are possible for this problem:

- jointly estimate the star signal (coronagraphic response in our case) and the planet position and intensity [18,19] from the angular image series;
- first, subtract two by two the angular images to remove the star signal (the speckles) and produce angular differential data; second, estimate the planet [20,21] only from this new angular differential data series.

In the framework of the SPHERE project, the static aberrations are likely to slowly evolve during observing time, and if the first option above were chosen, the estimation of the star signal should therefore be done several times during the night. We therefore choose the second option, which consists of removing the star signal numerically by performing an angular image subtraction between image pairs taken at time intervals that are long enough to obtain the peculiar signature of the planet's apparent rotation but shorter than the evolution time of the static aberrations. This signature is illustrated in Fig. 1. Let $\boldsymbol{i}_t$ be the raw image at time t; the new data are image differences $\Delta(\boldsymbol{r},t_1,t_2)$:

$$\Delta(\boldsymbol{r},t_1,t_2) \triangleq \boldsymbol{i}_{t_1}(\boldsymbol{r}) - \boldsymbol{i}_{t_2}(\boldsymbol{r}), \tag{1}$$

where t_1 and t_2 are chosen times and $\boldsymbol{r}$ the position in the focal plane.

In the angular image series, there is a very large number of possible couples (t_1,t_2). In this paper, we select the time couples so that (a) all images are selected at least

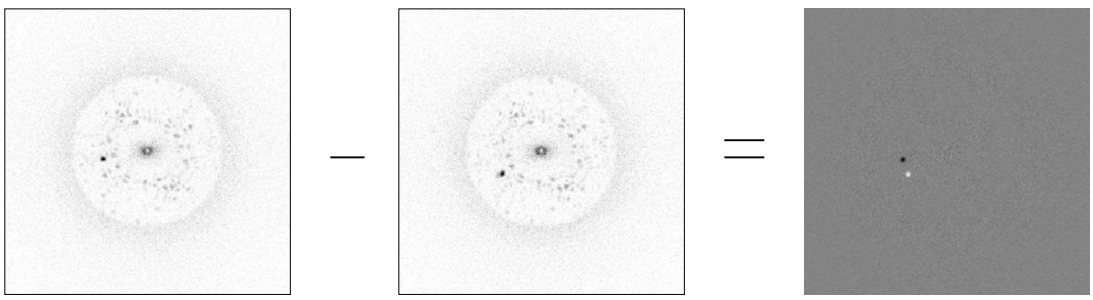


Fig. 1. Illustration of the angular differences performed on the raw data. Left and center: two noiseless raw coronagraphic images $\boldsymbol{i}_{t_1}$ and $\boldsymbol{i}_{t_2}$ of a star with a very bright planet (10^3 intensity ratio between star and planet). Right: difference $\Delta(\boldsymbol{r},t_1,t_2)$, used as our new data, which completely removes the star signal in this case where the quasi-static aberrations have not evolved. The colormap is inverted for better legibility, with black corresponding to the maximum value.

1328 J. Opt. Soc. Am. A/Vol. 26, No. 6/June 2009 Mugnier *et al.*

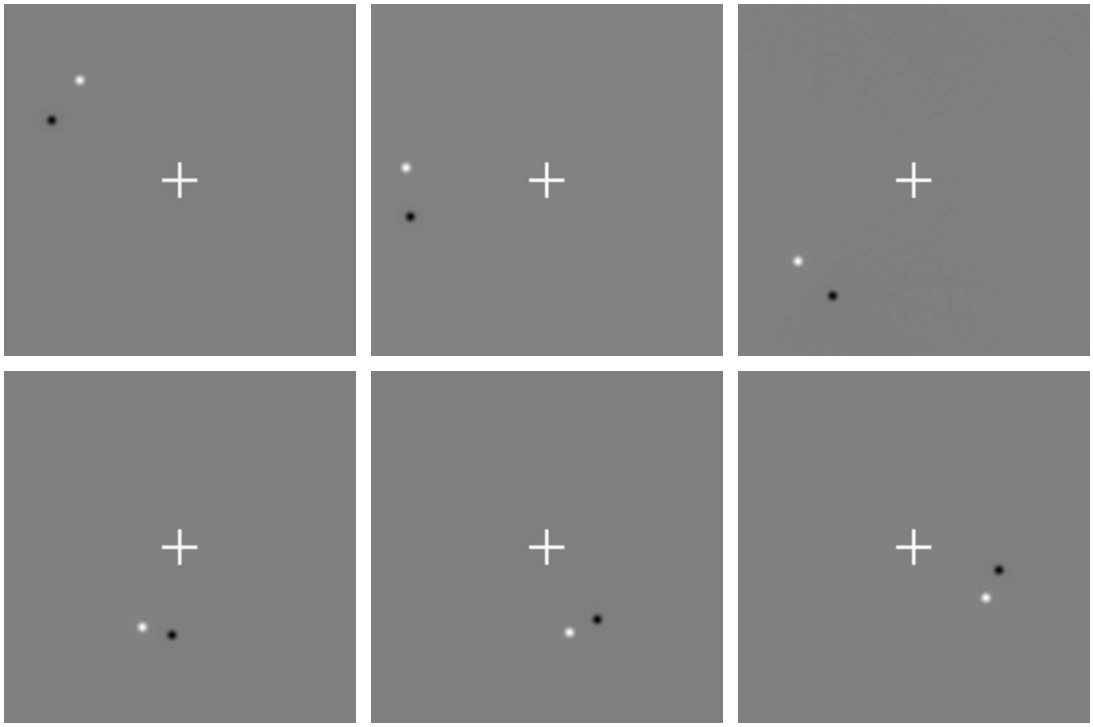


Fig. 2. Examples of planet signatures $\boldsymbol{p}(\boldsymbol{r},k;\boldsymbol{r}_0)$. Top: for one value of $\boldsymbol{r}_0$ and three values of the time index k. Bottom: for another value of $\boldsymbol{r}_0$ and the three same values of the time index k. The cross in the middle represents the center of the field rotation.

once, (b) at least 95% of the planet signal is preserved in the image difference (which can be shown to imply that the planet must have moved by at least $1.5\lambda/D$), and (c) if possible, the static aberrations are very similar in the two subtracted images. Note that the best compromise may be difficult to achieve because it depends on many factors, i.e., both system parameters, such as the evolution rate of static aberrations, and observational parameters, such as the field rotation speed. In any case, a set of new data are computed using all the raw images. These new data are the ones to be used subsequently to detect the planet(s).

The new data consist of $k_{\max}$ image differences denoted by $\Delta(\boldsymbol{r},k)$, where k is an index used to reference the time couples, and is referred to as the time index in the sequel. Assuming that a planet is present, the data model at each pixel $\boldsymbol{r}$ of image difference k is

$$\Delta(\boldsymbol{r},k) = a\boldsymbol{p}(\boldsymbol{r},k;\boldsymbol{r}_0) + \boldsymbol{n}(\boldsymbol{r},k), \tag{2}$$

where scalar a is the unknown planet intensity, the two-component vector $\boldsymbol{r}_0$ is the unknown planet position at the beginning of the observation (in the first image of the series), $\boldsymbol{n}(\boldsymbol{r},k)$ denotes the noise, and $\boldsymbol{p}(\boldsymbol{r},k;\boldsymbol{r}_0)$ is a synthetic (i.e., noiseless) pattern, which is the precomputed theoretical planet signature (PS) for a planet at an assumed $\boldsymbol{r}_0$ initial position.

The PS of index k is simply the difference of two theoretical noiseless planet images (i.e., PSFs) suitably positioned in the field and corresponding to the field rotations at the two times t_1 and t_2 used in $\Delta(\boldsymbol{r},k)$. It can be seen as the space- and time-varying PSF of our new data $\Delta(\boldsymbol{r},k)$. It is important to emphasize that the PS directly depends on $\boldsymbol{r}_0$; therefore, a set of PSs must be computed for all the candidate positions of a planet in order to be used later in the detection method. Figure 2 shows such PSs $\boldsymbol{p}(\boldsymbol{r},k;\boldsymbol{r}_0)$ for two different initial positions $\boldsymbol{r}_0$ and three values of the time index k.

3. MAXIMUM-LIKELIHOOD ESTIMATION FOR POSITION AND INTENSITY OF THE PLANET

The maximum-likelihood approach consists of searching for $(\hat{\boldsymbol{r}}_0,\hat{a})$ that maximize the likelihood $L(\boldsymbol{r}_0,a)$. In the following we assume that the noise is non-homogeneous (i.e., non-stationary), Gaussian, white in both time k and in space $\boldsymbol{r}$, and with variance $\sigma^2(\boldsymbol{r},k)$. This assumption is reasonable and allows us to take into account both the photon and the detector noise, as is done in AO-corrected image restoration [22]: for the intensity levels considered here, the Poisson statistics of photon noise is well approximated by a Gaussian probability density. Its variance map is estimated from the set of images, e.g., as an empirical variance of the image series at each pixel, and is considered known in the following expressions. As for the detector read-out noise, it is reasonably homogeneous white Gaussian and its variance can be estimated beforehand.

The likelihood is given by

$$L(\boldsymbol{r}_0,a) \propto \exp\left\{-\frac{1}{2}\sum_k \sum_{\boldsymbol{r}} \frac{|\Delta(\boldsymbol{r},k) - ap(\boldsymbol{r},k;\boldsymbol{r}_0)|^2}{2\sigma^2(\boldsymbol{r},k)}\right\}. \tag{3}$$

Maximizing this likelihood with respect to $(\boldsymbol{r}_0,a)$ is equivalent to maximizing the following metric, which is equal to the log-likelihood up to unimportant constants:

$$J(\boldsymbol{r}_0,a) \triangleq -a^2\sum_{k,\boldsymbol{r}}\frac{p^2(\boldsymbol{r},k;\boldsymbol{r}_0)}{\sigma^2(\boldsymbol{r},k)} + 2a\sum_{k,\boldsymbol{r}}\frac{p(\boldsymbol{r},k;\boldsymbol{r}_0)\Delta(\boldsymbol{r},k)}{\sigma^2(\boldsymbol{r},k)} \tag{4}$$

$$= 2\ln L(\boldsymbol{r}_0,a) + \text{const}. \tag{5}$$

The optimal value $\hat{a}(\boldsymbol{r}_0)$ of a for each given $\boldsymbol{r}_0$ is computable analytically:

$$\hat{a}(\boldsymbol{r}_0) = \frac{\sum_{k,\boldsymbol{r}} p(\boldsymbol{r},k;\boldsymbol{r}_0)\Delta(\boldsymbol{r},k)/\sigma^2(\boldsymbol{r},k)}{\sum_{k,\boldsymbol{r}} p^2(\boldsymbol{r},k;\boldsymbol{r}_0)/\sigma^2(\boldsymbol{r},k)}. \tag{6}$$

The numerator of this expression can be seen as a scalar product (correlation) between the PS $\boldsymbol{p}(\boldsymbol{r},k;\boldsymbol{r}_0)$ and the image differences $\Delta(\boldsymbol{r},k)$, with weights given by the noise variance. The denominator is simply a normalization constant.

If we insert this optimal value for the intensity into metric J, we obtain an expression of the latter that depends, explicitly at least, only on the sought planet position:

$$J'(\boldsymbol{r}_0) \triangleq J[\boldsymbol{r}_0,\hat{a}(\boldsymbol{r}_0)] = \frac{\left[\sum_{k,\boldsymbol{r}} p(\boldsymbol{r},k;\boldsymbol{r}_0)\Delta(\boldsymbol{r},k)/\sigma^2(\boldsymbol{r},k)\right]^2}{\sum_{k,\boldsymbol{r}} p^2(\boldsymbol{r},k;\boldsymbol{r}_0)/\sigma^2(\boldsymbol{r},k)}. \tag{7}$$

This criterion J' can be computed for each possible initial planet position on a grid that can be chosen as the original pixel grid of the images or as a finer grid if it is useful. The most likely initial planet position is then $\hat{\boldsymbol{r}}_0 = \arg\min J'(\boldsymbol{r}_0)$, and the most likely intensity is $\hat{a}(\hat{\boldsymbol{r}}_0)$ as computed with Eq. (6).

This estimator can be improved by constraining the estimated intensity to be positive. Indeed, the value of $\hat{a}(\boldsymbol{r}_0)$ of Eq. (6) is not necessarily positive, whereas the true intensity is. Additionally, because the estimation of $\hat{a}(\boldsymbol{r}_0)$ is a one-dimensional optimization, the optimal intensity subject to the positivity constraint is simply

$$\hat{a}_{pos}(\boldsymbol{r}_0) = \max\{\hat{a}(\boldsymbol{r}_0),0\}. \tag{8}$$

If we now insert this value for the planet intensity into metric J, it is easy to show that we obtain the criterion

$$J''(\boldsymbol{r}_0) \triangleq J[\boldsymbol{r}_0,\hat{a}_{pos}(\boldsymbol{r}_0)] = \begin{cases} J'(\boldsymbol{r}_0), & \text{if } \hat{a}(\boldsymbol{r}_0) > 0 \\ 0, & \text{if } \hat{a}(\boldsymbol{r}_0) \leq 0 \end{cases}, \tag{9}$$

where $\hat{a}(\boldsymbol{r}_0)$ is given by Eq. (6) and $J'(\boldsymbol{r}_0)$ by Eq. (7).

Figures 3 and 4 illustrate the usefulness of the positivity constraint in the idealized case of a noiseless image: the left and right images of Fig. 3 show the maps $\hat{a}(\boldsymbol{r}_0)$ and $\hat{a}_{pos}(\boldsymbol{r}_0)$, respectively. These maps are the planet intensity estimated at each position $\boldsymbol{r}_0$ without and with the positivity constraint, respectively. In the case without positivity the map is very similar to the auto-correlation of the PS, which explains its shape with two negative bumps. Figure 4 shows the corresponding maps $J'(\boldsymbol{r}_0)$ and $J''(\boldsymbol{r}_0)$ of the log-likelihood obtained, respectively, without [Equation (7)] and with [Equation (9)] the positivity constraint on the intensity. Clearly, J'' is very different from a thresholded version of J' and has fewer local maxima: the positivity constraint removes the sidelobes of the log-likelihood, and hence it should contribute to removing false detections in a noisier case.

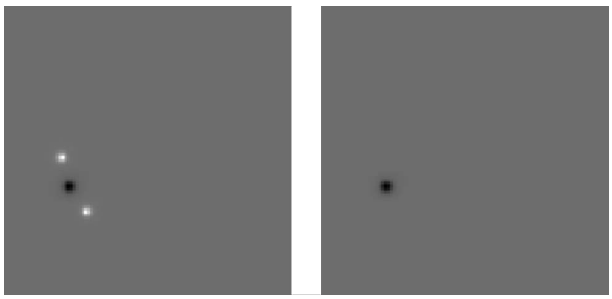


Fig. 3. Intensity maps $\hat{a}(\boldsymbol{r}_0)$ and $\hat{a}_{pos}(\boldsymbol{r}_0)$ estimated without (left) and with (right) the positivity constraint, respectively, in a noiseless case. The colormap is inverted for better legibility: the white sidelobes have the lowest, and negative, value; the gray background has zero value, and black corresponds to the highest value.

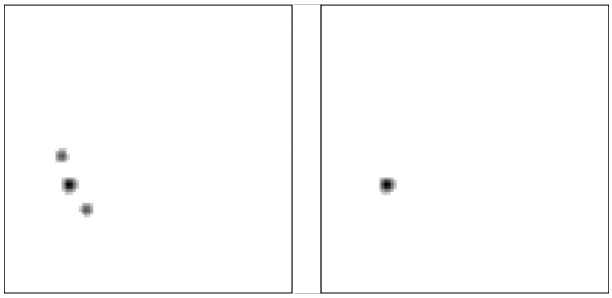


Fig. 4. Log-likelihood maps of the position of the planet in the noiseless case of Fig. 3 without [$J'(\boldsymbol{r}_0)$, left] and with [$J''(\boldsymbol{r}_0)$, right] the positivity constraint on the estimated intensity. The colormap is inverted for better legibility: white is the lowest, and zero, value; black corresponds to the highest value.

4. DETECTION CRITERION

Once the likelihood and intensity maps are computed, the main problem is to decide which peaks are true planets and which ones are not. One way to do so is to additionally compute the standard deviation of the estimated intensity $\sigma[\hat{a}(\boldsymbol{r}_0)]$ for each possible planet position $\boldsymbol{r}_0$, i.e., to compute how the noise propagates from the images to our intensity estimator.

We define the signal-to-noise ratio (SNR) of the estimated planet intensity as

$$\text{SNR}[\hat{a}(\boldsymbol{r}_0)] \triangleq \hat{a}(\boldsymbol{r}_0)/\sigma[\hat{a}(\boldsymbol{r}_0)]. \tag{10}$$

A possible detection criterion is then to decide that all positions where this SNR is greater than some threshold are true detections. In the Gaussian setting assumed in this paper, this detection criterion can be linked to the probability of false alarm.

The variance of the estimated intensity for a given position $\boldsymbol{r}_0$ is computed by means of Eq. (6) using the abovementioned property that the noise in our images $\Delta(k,\boldsymbol{r})$ is white, both temporally and spatially:

$$\sigma^2[\hat{a}(\boldsymbol{r}_0)] = \frac{1}{\left[\sum_{\boldsymbol{r}',k'} \frac{p^2(\boldsymbol{r},k';\boldsymbol{r}_0)}{\sigma^2(\boldsymbol{r},k')}\right]^2} \sum_{\boldsymbol{r},k} \left[\frac{p(\boldsymbol{r},k;\boldsymbol{r}_0)}{\sigma^2(\boldsymbol{r},k)}\right]^2 \sigma^2(\boldsymbol{r},k),$$

which can be simplified into

$$\sigma^2[\hat{a}(\boldsymbol{r}_0)] = \left[\sum_{\boldsymbol{r},k} \frac{p^2(\boldsymbol{r},k;\boldsymbol{r}_0)}{\sigma^2(\boldsymbol{r},k)}\right]^{-1}. \quad (11)$$

Interestingly, the SNR of the estimated intensity is linked very directly to the log-likelihood J' of Eq. (7):

$$J'(\boldsymbol{r}_0) = \{\hat{a}(\boldsymbol{r}_0)/\sigma[\hat{a}(\boldsymbol{r}_0)]\}^2 = \{\mathrm{SNR}[\hat{a}(\boldsymbol{r}_0)]\}^2. \quad (12)$$

Consequently, maximizing (resp. thresholding) the likelihood is actually equivalent to maximizing (resp. thresholding) the SNR of the estimated intensity.

Implementation details. The computation of the estimated intensity maps and of the likelihood maps is quite straightforward. It is performed by means of Eqs. (6)–(9). The main burden lies in the computation of the PSs $p(\boldsymbol{r},k;\boldsymbol{r}_0)$ for all time indices k and all searched initial planet positions $\boldsymbol{r}_0$. Each of these PSs is the difference of two theoretical noiseless planet images that must be suitably positioned in the field. This requires shifting each of these two theoretical planet images by a non-integer amount, whereas only integer shifts can be performed at very low computational cost. To solve this problem, we pre-compute a library of all the possible shifted theoretical planet images for x and y shifts regularly spaced between zero and one, typically every 1/20th of a pixel. This notably alleviates the computational burden. The total computing time currently remains of the order of three hours for the simulations presented below, which involve a hundred 256×256 images and a 200×200 grid for the planet position search.

5. TESTS ON SIMULATED DATA

A. Simulation Conditions

We use a Fourier-based simulation method that describes the AO via the spatial power spectrum of the residual phase [23] and is presented in [24]. It takes the following set of parameters, which are representative of the SPHERE/IRDIS instrument on the VLT:

- an 8 m telescope, a seeing of 0.8″, and a wind speed of 12.5 m/s;
- a SAXO-like AO system [6]: 41×41 actuators, a 40×40 sub-aperture Hartmann–Shack wavefront sensor, a sampling frequency of 1200 Hz;
- static aberrations with a standard deviation of $\sigma_{\phi_u}=35$ nm upstream of the coronagraph and $\sigma_{\phi_d}=100$ nm downstream of the coronagraph. We have assumed a pupil-stabilized mode with static aberrations kept constant during the simulated run.

A hundred 256×256 images are simulated at an imaging wavelength of $\lambda=1.593$ μm with Poisson noise. The image sampling corresponds to that of the SPHERE/IRDIS instrument, which is Shannon-sampled for a wavelength of 0.95 μm; the images are thus oversampled by a factor $1.593/0.95=1.677$, and their field is $76\lambda/D$ wide. The image of the star is computed by means of the analytical expression for the long-exposure AO-corrected coronagraphic image of a star [25] and is shown on the left part of Fig. 5.

We have simulated seven planets which lie aligned at distances that are multiples of $4\lambda/D$ from the central star. The long-exposure AO-corrected image of a planet is computed using the static aberrations and the phase structure function of the AO-corrected residual phase, assuming that the planets do not "see" the coronagraph. Such an image is presented on the right part of Fig. 5. For each of the seven simulated planets, this long-exposure planet image is then added to each of the star images at the appropriate planet locations. With the currently foreseen integration time of SPHERE/IRDIS images, the smearing of a planet due to field rotation during an individual exposure remains negligible, even for planets far from their star, and is thus not taken into account in the simulation.

The star intensity at the entrance aperture of the telescope is 2.67×10^7 ph/s and the planet intensity is 28.5 ph/s; the intensity ratio is thus 9.36×10^5. The telescope+instrument transmission is 0.09 without the coronagraph. The coronagraph further attenuates the star light with a transmission factor of 0.13. Depending on the simulation data, the total exposure time is either 1 or 2 h. In the 1 h case the total star intensity is 1.127×10^7 ph/image (or 172 ph/pixel on average). The total planet intensity is 93 ph/image in any of the 100 images. This corresponds to a maximum planet intensity of $\simeq$4.8 ph/pixel, the exact maximum value depending on whether or not the planet falls at the center of a pixel in the image.

For the purpose of testing our method, we have simulated the field rotation in the following simplified way:

- 50 images are simulated before the star crosses the meridian, and 50 images afterward; the set of 100 parallactic angles of the star is centered on the meridian.
- The step between two consecutive images is constant; for 100 images it is 1°.
- There is a gap of 20° around the meridian, to prevent the overlapping of the planet signals of two images that are to be subtracted. As a result, the angle between the first and the last image is 120°.

The image combination scheme we chose in these simulations is to associate each image with its symmetrical

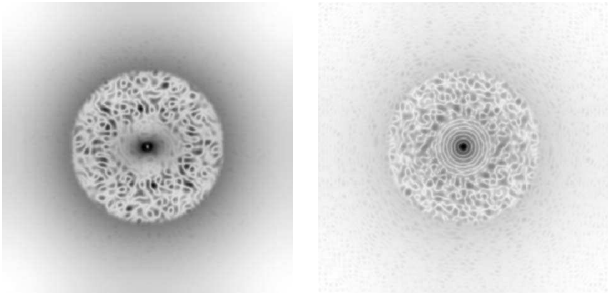


Fig. 5. Simulated PSFs with (left) and without (right) coronagraph in logarithmic scale. Note that for legibility the colormap is inverted, and each of these two images is represented with its own gray-level scale (with black corresponding to its maximum value).

one with respect to the meridian (the first image with the last, the second with the last-but-one, etc.). This way, the position of the achromatic dispersion compensator of the AO system is the same in the two images of each image couple, which should minimize the differential aberrations if the achromatic dispersion compensator is the main contributor to the evolution of the static aberrations.

B. Impact of the Proposed Positivity Constraint and of the Noise Variance Map

Figure 6 shows the likelihood maps obtained with 100 images and an exposure time of 1 h, with and without the positivity constraint and the non-homogeneous noise variance. The use of a noise variance map can be seen as taking into account some prior knowledge: if the noise variance map is unknown, one will use a homogeneous (i.e., constant) noise variance map, which cancels out in all the above expressions. As can be seen in Fig. 6, the positivity constraint on the planet's intensity and the use of a noise variance map both improve the likelihood map and thus the detection: going from the homogeneous noise variance without positivity (top left) to the homogeneous noise variance with positivity (top right) notably decreases low-level peaks of the likelihood as well as some high-level peaks that correspond to false alarms (for instance, one in the middle, below the center of the image). These are the peaks due to the negative values of the estimated intensity.

Additionally, going from the homogeneous noise variance with positivity (top right) to the inhomogeneous noise variance with positivity (bottom right) further improves the likelihood map by dimming some other spurious peaks (for instance, one on the top left part of the image, at about 45° from the star).

To better quantify the improvement brought by positivity and by the use of an inhomogeneous noise variance map, Fig. 7 shows the SNR of the estimated intensity [defined by Equation (10)] thresholded to values from 3 to 6, in the photon-starved case of a 1 h total observation time. In the two cases where a homogeneous noise is assumed in the processing, the noise variance has been taken equal to the spatial average of the empirical variance of each pixel in time.

As seen on the first two lines of images of Fig. 7, the use of the positivity constraint removes a very substantial number of false alarms. On the same images, one notes that the inhomogeneous noise model also reduces the false alarm rate.

For the case where both the positivity and the inhomo-

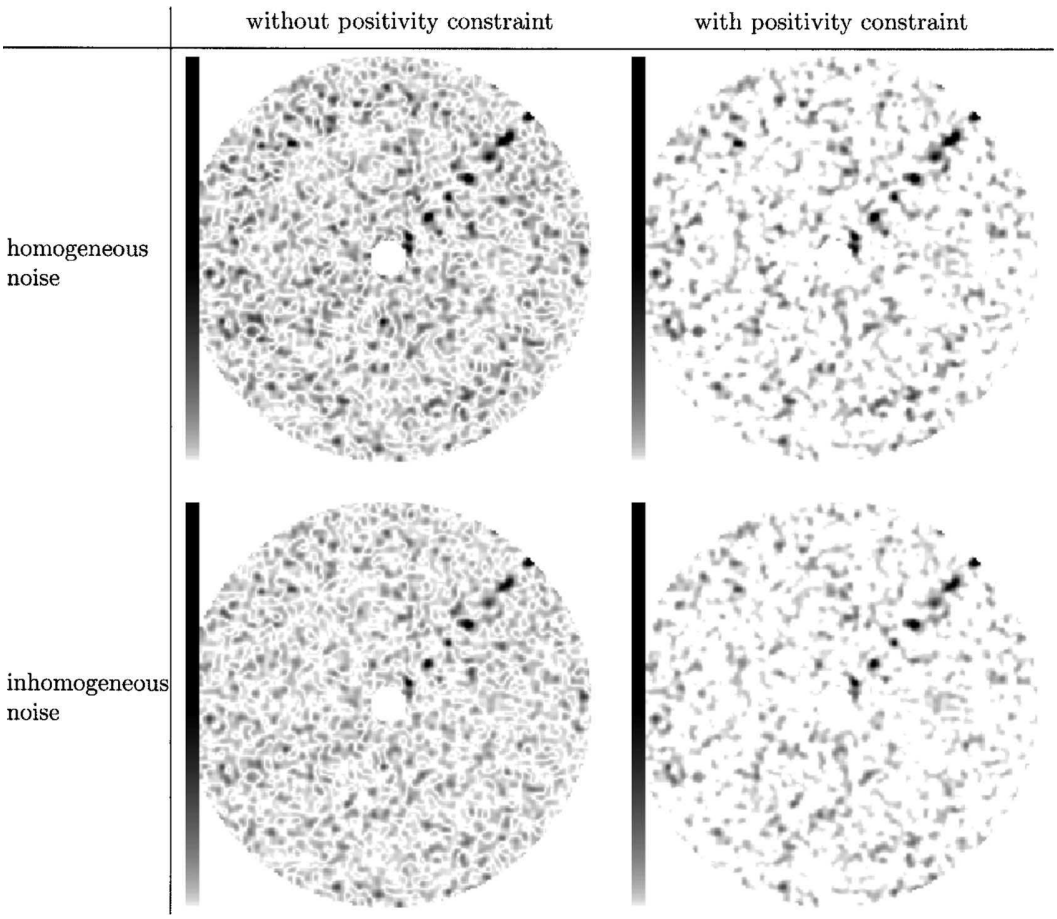


Fig. 6. Likelihood maps with 100 images and an exposure time of 1 h. Top row: homogeneous noise. Bottom row: inhomogeneous noise. Left column: without positivity constraint. Right column: with positivity constraint. For legibility the colormap is inverted: white corresponds to the minimum value of zero, and black to the highest value.

geneous noise variance map are used, there exists in this simulation a threshold (of 4) for which all the true planets are detected and no false alarm is present. The corresponding detection map is the boxed one of Fig. 7. The image immediately on the left, obtained without the positivity constraint, has a false alarm at the very right of the field.

For the three other cases, whatever the chosen threshold, in this simulation there are either false alarms (for low threshold values) or undetected planets (for high threshold values).

One can also note that in the homogeneous cases there are more detected planets for high threshold values, as well as more false alarms for low threshold values, compared to the inhomogeneous cases. For instance, five planets are detected in the homogeneous case with positivity and a 6σ detection (i.e., a threshold of 6) and only three in the inhomogeneous case. To detect the same number of planets in the inhomogeneous case one must set the threshold to 5 instead of 6. The homogeneous detection maps thus appear to be, so to speak, shifted toward the high thresholds. We conjecture that this is due to the noise standard deviation adopted in the homogeneous case being somewhat arbitrary because of the mismatch between the true noise model and the one used in the detection.

C. Impact of the Exposure Time

Figure 8 illustrates the influence of the exposure time on the likelihood maps and on the detection maps for several threshold values. As expected, for 2 h of total exposure time instead of 1, the likelihood map has fewer spurious sidelobes and thresholding is easier. This can be seen in the detection maps: there are fewer false detections for the low threshold of 3 (second line of Fig. 8). Additionally, all planets are detected even for higher threshold values, up to a value of 5 (last line of Fig. 8). The 2 h exposure time thus results in a more reliable detection; in the Gaussian setting assumed here, and assuming that the quasi-static aberrations have been perfectly removed by the image differentiation, a 4σ detection such as the one obtained on Fig. 7 corresponds to a 6×10^{-5} probability of false alarm, whereas the 5σ detection of Fig. 8 corresponds to a 6×10^{-7} probability of false alarm.

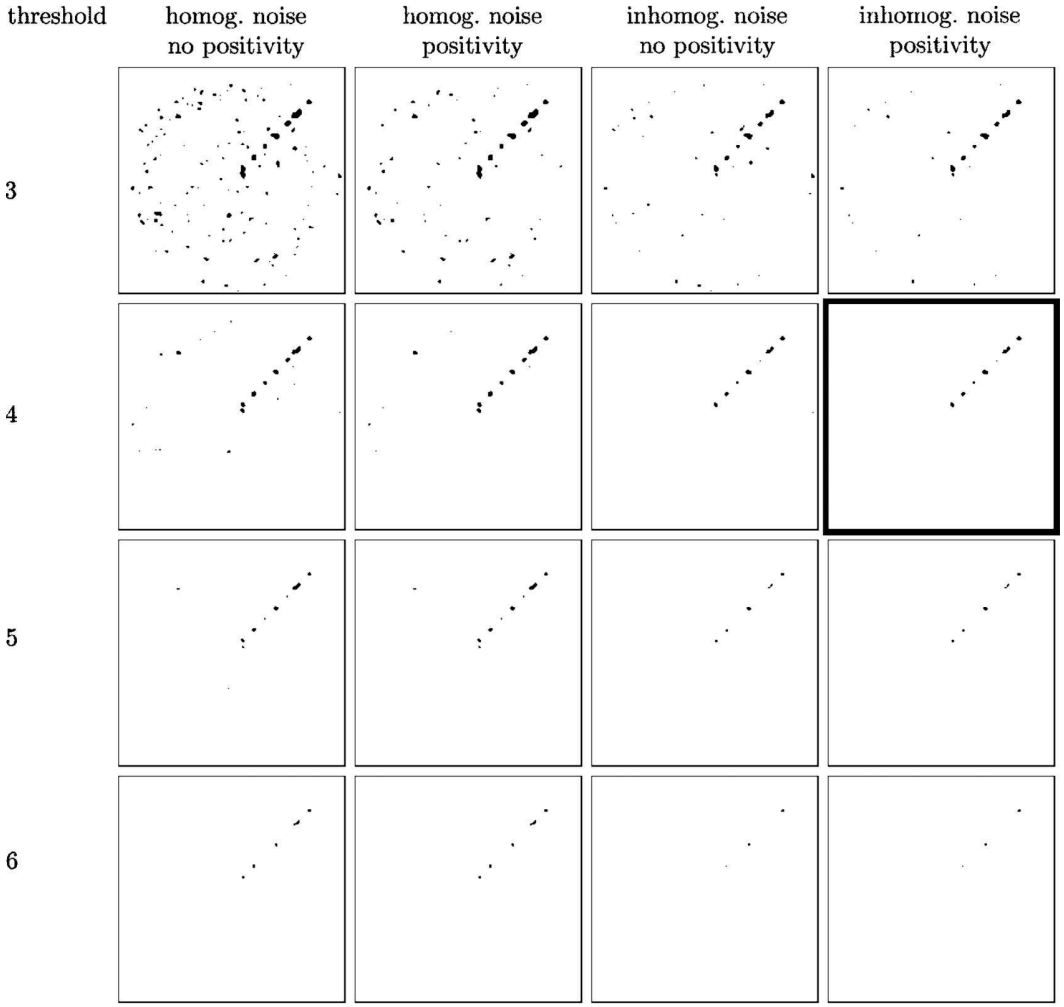


Fig. 7. Detection maps obtained by thresholding the maps of the SNR of the estimated intensity of Fig. 6, for various thresholds mentioned in the left column. From left to right: homogeneous noise, no positivity; homogeneous noise and positivity; inhomogeneous noise, no positivity; inhomogeneous noise and positivity.

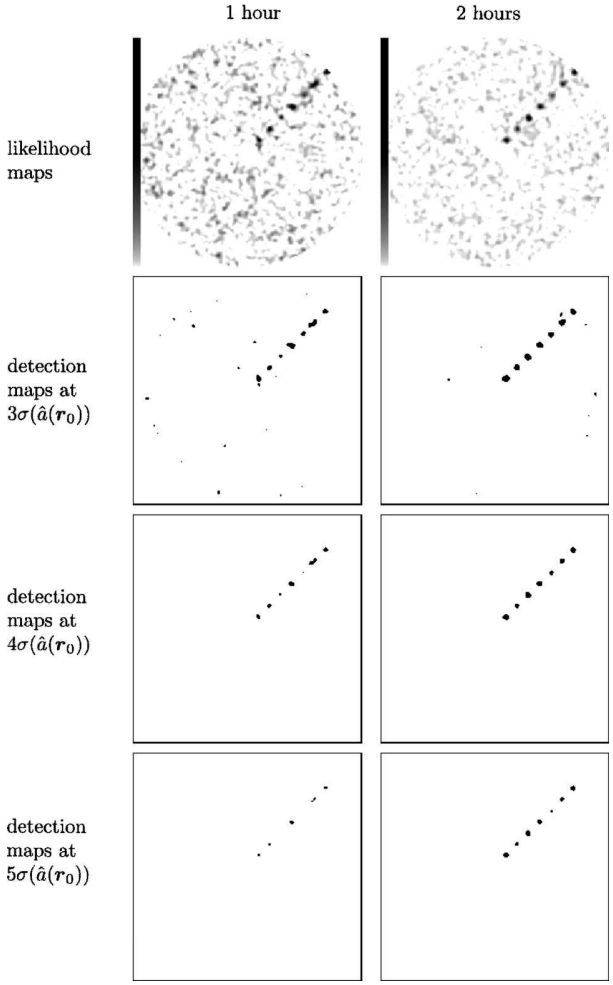


Fig. 8. Likelihood and detection maps for different exposure times: 1 h (left) and 2 h (right) with 100 images in each case. The estimation is done with the inhomogeneous noise model and a positivity constraint.

6. DISCUSSION

We have presented a method based on maximum likelihood for exoplanet detection with ground-based instruments such as SPHERE, and tested it by means of simulations. This method makes use of the temporal diversity of the images brought by field rotation in order to disentangle planets from speckles. It can enforce a positivity constraint on the estimated intensity and can use the noise variance map of the images, the beneficial influence of which has been demonstrated. A reasonable detection criterion has also been proposed and tested; it is based on the computation of the noise propagation from the images to the estimated intensity of the potential planet. As an example, under the assumptions and simulation conditions described in the paper, a reliable detection is obtained with 2 h of data for a 10^6 intensity ratio between the star and the planet.

While the results of our simulations are very encouraging, these simulations should be made more realistic by taking into account the variations of turbulence strength and of static aberrations during the night. The effect of these variations on the images will be partially compensated for by performing a spectral difference between simultaneous image channels and using these spectral difference images as inputs $\boldsymbol{i}_t$ for the method proposed in this paper. Yet, this compensation will be only partial. Preliminary simulations (not presented herein) suggest that the subtraction between images taken at different times must incorporate a scaling factor that can be optimized for each image pair. Additionally, the optimal scaling factor varies with the distance to the star, which suggests performing the detection in different annuli with a potentially different scaling factor for each annulus.

One short-term perspective is thus to assess the performance of the method in the case of slowly evolving aberration and turbulence parameters using both spectral channels of the IRDIS instrument. Additionally, because the pair-wise image combination is quite flexible, this combination should be adapted to the variability of these parameters in order to optimize the detection. Another short-term perspective is to optimize the code in order to reduce the computation time. Additionally, the design of a more elaborate detection criterion taking into account the non-Gaussianity of the noise also deserves further studies. As a final note we mention that this method, which has been named ANDROMEDA, for ANgular Differential OptiMal Exoplanet Detection Algorithm, is likely to be implemented in the data pipeline of the SPHERE/IRDIS instrument, so that it will be available to astronomers for processing the bi-spectral images provided by this instrument.

ACKNOWLEDGMENTS

The authors thank David Mouillet of Laboratoire d'Astrophysique de l'Observatoire de Grenoble (LAOG) for sharing his expertise on the SPHERE system, and the people from the SPHERE consortium for several fruitful discussions, notably Marcel Carbillet, Anthony Boccaletti, Maud Langlois, Kjetil Dohlen, and Jean-Luc Beuzit.

REFERENCES

1. J. R. P. Angel, "Groundbased imaging of extrasolar planets using adaptive optics," Nature (London) **368**, 203–207 (1994).
2. A. Burrows and D. Sudarsky, "Models of irradiated extrasolar giant planets," in *Scientific Frontiers in Research on Extrasolar Planets*, Vol. 294 of Astronomical Society of the Pacific Conference Series, D. Deming and S. Seager, eds. (Astronomical Society of the Pacific, 2003), pp. 491–498.
3. J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, and F. Wildi, "SPHERE: a 'planet finder' instrument for the VLT," Proc. SPIE **7014**, 701418 (2008).
4. B. A. Macintosh, J. R. Graham, D. W. Palmer, R. Doyon, J. Dunn, D. T. Gavel, J. Larkin, B. Oppenheimer, L. Saddlemyer, A. Sivaramakrishnan, J. K. Wallace, B. Bauman, D. A. Erickson, C. Marois, L. A. Poyneer, and R. Soummer, "The Gemini Planet Imager: from science to design to construction," Proc. SPIE **7015**, 701518 (2008).
5. M. R. Swain, G. Vasisht, and G. Tinetti, "The presence of methane in the atmosphere of an extrasolar planet," Nature (London) **452**, 329–331 (2008).
6. T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper, and P. Puget, "High-order adaptive optics requirements for direct detection of extra-solar planets: Application to the

SPHERE instrument," Opt. Express **14**, 7515–7534 (2006).
7. B. Lyot, "The study of the solar corona and prominences without eclipses (George Darwin Lecture, 1939)," Mon. Not. R. Astron. Soc. **99**, 580–594 (1939).
8. D. Rouan, P. Riaud, A. Boccaletti, Y. Clénet, and A. Labeyrie, "The Four-Quadrant Phase-Mask Coronagraph. I. Principle," Publ. Astron. Soc. Pac. **112**, 1479–1486 (2000).
9. R. Soummer, C. Aime, and P. E. Falloon, "Stellar coronagraphy with prolate apodized circular apertures," Astron. Astrophys. **397**, 1161–1172 (2003).
10. R. Racine, G. A. Walker, D. Nadeau, and C. Marois, "Speckle noise and the detection of faint companions," Publ. Astron. Soc. Pac. **112**, 587–594 (1999).
11. C. Marois, R. Doyon, D. Nadeau, R. Racine, and G. A. H. Walker, "Effects of quasi-static aberrations in faint companion searches," in *Astronomy with High-Contrast Imaging*, Vol. 8 of EAS Publications Series, C. Aime and R. Soummer, eds. (EDP Sciences, 2003), pp. 233–243.
12. C. Marois, R. Doyon, D. Nadeau, R. Racine, M. Riopel, P. Vallée, and D. Lafrenière, "TRIDENT: an infrared differential imaging camera optimized for the detection of methanated substellar companions," Publ. Astron. Soc. Pac. **117**, 745–756 (2005).
13. K. Dohlen, M. Langlois, and M. Saisse, "The infrared dual imaging and spectrograph for SPHERE: design and performance," Proc. SPIE **7014**, 70143L (2008).
14. A. Labeyrie, "Detection of extra-solar planets," in *Formation of Planetary Systems*, A. Brahic, ed. (Cepadues-Editions, Toulouse, 1982), pp. 883, 885–888.
15. M. Mueller and G. Weigelt, "High-resolution astronomical imaging by roll deconvolution of space telescope data," Astron. Astrophys. **175**, 312–318 (1987).
16. C. Marois, D. Lafrenière, R. Doyon, B. Macintosh, and D. Nadeau, "Angular differential imaging: A powerful high-contrast imaging technique," Astrophys. J. **641**, 556–564 (2006).
17. D. Lafrenière, C. Marois, R. Doyon, D. Nadeau, and É. Artigau, "A new algorithm for point-spread function subtraction in high-contrast imaging: A demonstration with angular differential imaging," Astrophys. J. **660**, 770–780 (2007).
18. I. Smith, A. Ferrari, and M. Carbillet, "Detection algorithm of exoplanets in field-rotated images: preliminary results on simulated SPHERE coronographic images," in *Semaine de l'Astrophysique Française*, SF2A (EDP Sciences, 2007).
19. I. Smith, M. Carbillet, A. Ferrari, D. Mouillet, A. Boccaletti, and K. Dohlen, "Simulation of moving exoplanets detection using the VLT instrument SPHERE/IRDIS," Proc. SPIE **7015**, 70156F (2008).
20. L. M. Mugnier, J.-F. Sauvage, T. Fusco, and G. Rousset, "Multi-channel planet detection algorithm for angular differential imaging," in *Adaptive Optics: Analysis and Methods*, OSA Technical Digest (OSA, 2007), paper ATu A4.
21. J.-F. Sauvage, L. Mugnier, A. Woelfflé, T. Fusco, and G. Rousset, "Multi-channel algorithm for exoplanets detection by angular differential imaging," in *Semaine de l'Astrophysique Française*, SF2A (EDP Sciences, 2007).
22. L. M. Mugnier, T. Fusco, and J.-M. Conan, "MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images," J. Opt. Soc. Am. A **21**, 1841–1854 (2004).
23. L. Jolissaint, J.-P. Véran, and R. Conan, "Analytical modeling of adaptive optics: foundations of the phase spatial power spectrum approach," J. Opt. Soc. Am. A **23**, 382–394 (2006).
24. R. Conan, T. Fusco, G. Rousset, D. Mouillet, J.-L. Beuzit, M. Nicolle, and C. Petit, "Modeling and analysis of XAO systems. Application to VLT-Planet Finder," Proc. SPIE **5490**, pp. 602–608 (2004).
25. J.-F. Sauvage, "Calibration et méthodes d'inversion en imagerie haute dynamique pour la détection directe d'exoplanètes," Ph.D. thesis (Université Paris VII, 2007).

Annexe F

Article Meimon et coll., 2009 [A7]

108 J. Opt. Soc. Am. A/Vol. 26, No. 1/January 2009 Meimon *et al.*

Self-calibration approach for optical long-baseline interferometry imaging

Serge Meimon,[1,*] Laurent M. Mugnier,[1] and Guy Le Besnerais[2]

[1]*Département d'Optique Théorique et Appliquée, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon cedex, France*

[2]*Département Traitement de l'Information et Modélisation, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon cedex, France*

[*]*Corresponding author: meimon@onera.fr*

Received May 28, 2008; revised October 14, 2008; accepted October 17, 2008; posted October 23, 2008 (Doc. ID 96631); published December 19, 2008

Current optical interferometers are affected by unknown turbulent phases on each telescope. In the field of radio interferometry, the self-calibration technique is a powerful tool to process interferometric data with missing phase information. This paper intends to revisit the application of self-calibration to optical long-baseline interferometry (OLBI). We cast rigorously the OLBI data processing problem into the self-calibration framework and demonstrate the efficiency of the method on a real astronomical OLBI data set. © 2008 Optical Society of America

OCIS codes: 120.3180, 100.3020, 100.3190.

1. INTRODUCTION

Optical long-baseline interferometry (OLBI) aims to combine light collected by widely separated telescopes to access angular resolutions beyond the diffraction limit of each individual aperture. Long-baseline interferometers measure a discrete set of spatial frequencies of the observed object, or Fourier data. Due to instrumental complexity, current interferometers recombine only a few telescopes, and even several nights of observation lead to a very limited number of Fourier data; moreover, due to the atmospheric turbulence, it is very difficult to get reliable phase information from ground-based interferometry [1]. Hence OLBI has to deal with severe underdetermination and missing phase information.

The classical answer to underdetermination is to use a parametric approach, i.e., to search for an object entirely described by a small set of parameters (for instance, a circular object with a parametric attenuation profile). With a "good model," such an approach allows a reliable and precise estimation of astrophysical parameters. A good model should limit as much as possible the number of free parameters, while allowing a description of all the object's features, because parametric inversion cannot reveal unguessed features. The χ^2 fit is often used as a model quality diagnosis, since an inadequate model will often result in a poor fit to the data, thus revealing that a new model (with more parameters or different parameters) is needed. However, it does not reveal which new model must be adopted.

As progress in instrumental issues gives access to better frequency coverage, i.e., to potentially finer descriptions of the object, the choice of the model becomes more difficult. An alternate and complementary approach is then nonparametric reconstruction, which we will call "optical long-baseline interferometric imaging" (OLBII). Imaging means that the object is described by a large set of parameters, such as coefficients of the object's decomposition in some spatial functional basis, while underdetermination is tackled by regularization tools. Imaging is useful to understand the structure of a complex object when prior information is limited.

From the beginning, OLBII has been influenced by the remarkable techniques developed in radio interferometry with very large baselines (VLBI) [2]. For instance, the "WIPE" OLBII technique of Lannes *et al.* [3] is inspired by the well-known CLEAN method [4]. As regards the missing phase problem, the self-calibration technique proposed in radio interferometry by Cornwell and Wilkinson [5] underlies recent work in OLBII [6].

This paper intends to revisit the application of self-calibration to OLBI. Our contribution is threefold:

1. We cast rigorously the OLBI data processing problem into the self-calibration framework, with consideration of the second-order statistics of the noise.

2. We propose WISARD (for Weak-phase Interferometric Sample Alternating Reconstruction Device), a self-calibration algorithm dedicated to OLBII, which uses the proposed data model within a Bayesian regularization approach.

3. We demonstrate the efficiency of WISARD on a real astronomical OLBI data set.

The paper is organized as follows: Section 2 describes the observation model of OLBI, briefly presents a Bayesian approach, and discusses the main problems that are encountered because of the incomplete OLBI data. Section 3 is devoted to the derivation of a specific myopic model, which achieves a good approximation of the data model and leads to self-calibration techniques. One such technique, WISARD, is proposed in Section 4. Results of WISARD on simulated and real astronomical data sets are

1084-7529/09/010108-13/$15.00

presented in Section 5. Our conclusions are given in Section 6. Most mathematical derivations are gathered in the appendixes.

2. REALISTIC OBSERVABLES IN OPTICAL LONG-BASELINE INTERFEROMETRY

A. Ideal Interferometric Data

Here we describe the ideal data, i.e., without aberrations, noise, or turbulence effects, produced by an N_t-telescope interferometer observing a monochromatic source with wavelength λ. The brightness distribution of the source is denoted $x(\boldsymbol{\xi})$, $\boldsymbol{\xi}$ being angular coordinates on the sky. Individual telescopes T_k of the interferometer are located at three-space positions $\vec{OT}_k$, and we denote $\boldsymbol{r}_k(t)$ the projection of $\vec{OT}_k$ onto $\mathcal{P}$, the plane normal to the pointing direction. Because of the Earth's rotation, the pointing direction changes during an observing night, so these projected vectors are time dependent.

Each pair (T_k, T_l) of telescopes yields a fringe pattern with a 2D spatial frequency $\boldsymbol{\nu}_{kl}(t) \triangleq \frac{\boldsymbol{u}_{kl}(t)}{\lambda}$, where $\boldsymbol{u}_{kl}(t)$ is the baseline

$$\boldsymbol{u}_{kl}(t) \triangleq \boldsymbol{r}_l(t) - \boldsymbol{r}_k(t), \tag{1}$$

that is, the projection of the vector $\vec{T_kT_l}$ onto $\mathcal{P}$.

Measuring the position and contrast of these fringes yields a phase $\phi^{kl}_{\text{data}}(t)$ and an amplitude $a^{kl}_{\text{data}}(t)$, which can be grouped together in a complex visibility:

$$y^{\text{data}}_{kl}(t) \triangleq a^{\text{data}}_{kl}(t) e^{i\phi^{\text{data}}_{kl}(t)}. \tag{2}$$

According to the Van Cittert–Zernike theorem [7], complex visibilities are *ideally* linked to the normalized Fourier transform (FT) of $x(\boldsymbol{\xi})$ at the 2D spatial frequency $\boldsymbol{\nu}_{kl}(t)$ through

$$y^{\text{data}}_{kl}(t) = \eta_{kl}(t) \frac{\text{FT}[x(\boldsymbol{\xi})](\boldsymbol{\nu}_{kl}(t))}{\text{FT}[x(\boldsymbol{\xi})](\boldsymbol{0})}. \tag{3}$$

The instrumental visibility $\eta_{kl}(t)$ accounts from the many potential sources of visibility loss: residual perturbations of the wavefront at each telescope, differential tilts between telescopes, differential polarization effects, nonzero spectral width, etc. In practice, the instrumental visibility is calibrated on a star reputed to be unresolved by the interferometer before the object of interest is observed and is compensated for in the preprocessing of the raw data. Thanks to this calibration step, we replace $\eta_{kl}(t)$ with 1 in Eq. (3).

For the sake of clarity, we consider a *complete* N_t-telescope array in what follows, i.e., one in which all the possible two-telescope baselines can be formed simultaneously, and a nonredundant interferometer configuration, where each baseline provides a different spatial frequency. Extension to incomplete and redundant settings is straightforward. Thus, at each time t, there are

$$N_b = \binom{N_t}{2} = \frac{N_t(N_t-1)}{2} \tag{4}$$

complex observation equations such as Eq. (3).

Let us briefly introduce the discretized observation model. The sought brightness distribution x is represented by the coefficients $\boldsymbol{x}$ of its projection onto some convenient spatial basis (box functions, sinc's, wavelets, prolate spheroidal functions, etc). The normalized discrete-continuous Fourier matrix $\boldsymbol{H}(t)$ maps the chosen discrete spatial representation into the real-valued instantaneous frequency coverage $\{\boldsymbol{\nu}_{kl}(t)\}_{1\leqslant k<l\leqslant N_t}$, and we further define

$$\begin{cases} \boldsymbol{a}(\boldsymbol{x},t) \triangleq |\boldsymbol{H}(t)\boldsymbol{x}|, \\ \boldsymbol{\phi}(\boldsymbol{x},t) \triangleq \arg\{\boldsymbol{H}(t)\boldsymbol{x}\}. \end{cases} \tag{5}$$

B. Effect of Atmospheric Turbulence on Short-Exposure Measurements

At optical wavelengths, atmospheric turbulence affects phase measurements through path-length fluctuations. The statistics of these fluctuations can be described by a time-scale parameter, the coherence time τ_0, typically around 10 ms, and by a space-scale parameter, the Fried parameter r_0 [[8]]. We assume that the diameter of the elementary apertures is small relative to the Fried parameter or that each telescope is corrected from the effects of turbulence by adaptive optics. The remaining turbulent effects on the interferometric measurements can be seen as a delay line between the two telescopes T_k and T_l, which affects short-exposure phase measurements through an additive differential piston $\varphi_l(t) - \varphi_k(t)$:

$$\phi^{\text{data}}_{kl}(t) = \phi_{kl}(\boldsymbol{x},t) + \varphi_l(t) - \varphi_k(t) + \text{noise}[2\pi] \tag{6}$$

or, in a matrix formulation:

$$\boldsymbol{\phi}^{\text{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{B}\boldsymbol{\varphi}(t) + \text{noise}[2\pi], \tag{7}$$

where $N_b \times N_t$ operator $\boldsymbol{B}$, called the baseline operator, is defined in Appendix A.

Because the differential pistons are zero mean, one might think that the object phase $\boldsymbol{\phi}(\boldsymbol{x},t)$ could be recovered from Eq. (7) by averaging over many realizations of the atmosphere. However, for a long baseline relative to the Fried parameter, the optical path difference between apertures introduced by turbulence may be very much greater than the observation wavelength and thus lead to random pistons much larger than 2π. The 2π-wrapped perturbation that affects phase (7) is then practically uniformly distributed in $[0, 2\pi]$. In consequence, averaging the short-exposure phase measurements (7) does not improve the signal-to-noise ratio (SNR).

In phase referencing techniques (see [9]), the turbulent pistons are measured in order to subtract them in Eq. (7). However powerful and promising, these methods require specific hardware and are not feasible for all sources. The only other way to obtain exploitable long-exposure data then is to form piston-free short-exposure observables *before* the averaging.

C. Piston-Free Short-Exposure Observables

Piston-free short-exposure phase observables are quantities $f(\boldsymbol{\phi}^{\text{data}}(t))$ in which the turbulent term $\boldsymbol{B}\boldsymbol{\varphi}(t)$ cancels out:

$$f(\phi^{\text{data}}(t)) = f(\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{B}\boldsymbol{\varphi}(t)) = f(\boldsymbol{\phi}(\boldsymbol{x},t)). \quad (8)$$

For an interferometric array of three telescopes or more, the closure phases [10] are one famous example, in which f is a linear operator performing triplewise summation of the phases. For any set of three telescopes (T_k, T_l, T_m), the short-exposure visibility phase data are

$$\begin{cases} \phi_{kl}^{\text{data}}(t) = \phi_{kl}(\boldsymbol{x},t) + \varphi_l(t) - \varphi_k(t) + \text{noise}[2\pi], \\ \phi_{lm}^{\text{data}}(t) = \phi_{lm}(\boldsymbol{x},t) + \varphi_m(t) - \varphi_l(t) + \text{noise}[2\pi], \\ \phi_{mk}^{\text{data}}(t) = \phi_{mk}(\boldsymbol{x},t) + \varphi_k(t) - \varphi_m(t) + \text{noise}[2\pi], \end{cases} \quad (9)$$

and the turbulent pistons cancel out in the closure phase defined by

$$\begin{aligned} \beta_{klm}^{\text{data}}(t) &\triangleq \phi_{kl}^{\text{data}}(t) + \phi_{lm}^{\text{data}}(t) + \phi_{mk}^{\text{data}}(t) + \text{noise}[2\pi] \\ &= \phi_{kl}(\boldsymbol{x},t) + \phi_{lm}(\boldsymbol{x},t) + \phi_{mk}(\boldsymbol{x},t) + \text{noise}[2\pi] \\ &\triangleq \beta_{klm}(\boldsymbol{x},t) + \text{noise}[2\pi]. \end{aligned} \quad (10)$$

We have the following properties:

- The set of all three-telescope closure phases that can be formed using a complete array is generated by the $(N_t-1)(N_t-2)/2$ closure phases $\beta_{1kl}^{\text{data}}(t)$, $k<l$, i.e., the closure phase that includes telescope T_1 (indeed, $\beta_{klm}^{\text{data}} = \beta_{1kl}^{\text{data}} + \beta_{1lm}^{\text{data}} - \beta_{1km}^{\text{data}}$). In what follows, these canonical closure phases are grouped together in a vector $\boldsymbol{\beta}^{\text{data}}$, and $\boldsymbol{C}$ denotes the linear closure operator such that $\boldsymbol{C}\boldsymbol{\phi}^{\text{data}} = \boldsymbol{\beta}^{\text{data}}$ (see Appendix A).
- If f is a continuous differentiable function verifying property (8), then

$$f(\boldsymbol{\phi}) = g(\boldsymbol{C}\boldsymbol{\phi}), \quad (11)$$

where g is some continuous differentiable function. In other terms, there is essentially *no operator other than the closure operator* that cancels out the effect of turbulence on short-exposure visibility phases (this property holds only in the monochromatic case).

The proof of the second property is given in Appendix B.

D. Long-Exposure Observables Data Model

To minimize the effect of noise, one is led to average short-exposure measurements into long-exposure observables, chosen so that they are asymptotically unbiased. The averaging time must be short enough with respect to the Earth's rotation so that the baseline does not change, and long enough to reach an acceptable SNR. The averaged quantities are generally these:

- averaged *squared amplitudes* $\boldsymbol{s}^{\text{data}}(t) = \langle \boldsymbol{a}^{\text{data}}(t+\tau)^2 \rangle_\tau$,
- averaged *bispectra* $\boldsymbol{V}_{1kl}^{\text{data}}(t) = \langle y_{1k}^{\text{data}}(t+\tau) \cdot y_{kl}^{\text{data}}(t+\tau) \cdot y_{l1}^{\text{data}}(t+\tau) \rangle_\tau$, $k<l$.

Squared amplitudes are preferred to amplitudes because their bias can be estimated and subtracted from the data. Short-exposure bispectra are continuous differentiable functions verifying property (8) and so correspond to a particular choice of g in Eq. (11). In the absence of noise, the averaged bispectrum amplitudes are redundant with the averaged squared amplitudes. Although they should be useful in low-SNR conditions, averaged bispectrum amplitudes are not considered in what follows. The averaged bispectrum phases $\beta_{1kl}^{\text{data}}(t)$, $k<l$ constitute unbiased long-exposure closure phase estimators. As such, they are linked to the object phases $\boldsymbol{\phi}(\boldsymbol{x},t)$ through

$$\boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \text{noise}[2\pi]. \quad (12)$$

It is shown in Appendix A that the kernel of the closure operator $\boldsymbol{C}$ is of dimension (N_t-1). Hence Eq. (12) implies that optical interferometry through turbulence has to deal with partial phase information. This result can also be obtained by counting up phase unknowns for each instant of measurement t: there are $N_t(N_t-1)/2$ unknown object visibility phases and $(N_t-1)(N_t-2)/2$ observable independent closure phases, which results in (N_t-1) missing phase data. As is well known in the radio interferometric community, the greater the number of apertures in the array, the smaller the proportion of missing phase information.

The long-exposure observables considered in this paper are noisy squared amplitudes $\boldsymbol{s}^{\text{data}}(t)$ and closure phases $\boldsymbol{\beta}^{\text{data}}(t)$. The only statistics usually available are the variances for each observable (as, for instance, in the OIFITS data exchange format [11]). The assumed noise distribution is consequently zero-mean white Gaussian:

$$\begin{cases} \boldsymbol{s}^{\text{data}}(t) = \boldsymbol{a}^2(\boldsymbol{x},t) + \boldsymbol{s}^{\text{noise}}(t), & \boldsymbol{s}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{s}(t)}), \\ \boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}^{\text{noise}}(t)[2\pi], & \boldsymbol{\beta}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{\beta}(t)}). \end{cases} \quad (13)$$

The matrices $\boldsymbol{R}_{\boldsymbol{s}(t)}$ and $\boldsymbol{R}_{\boldsymbol{\beta}(t)}$ are diagonal, with variances related to the integration time, although correlations may be produced by the use of the same reference stars in the calibration process [12].

E. Bayesian Reconstruction Methods

This approach first forms the anti-log-likelihood according to model (13):

$$J^{\text{data}}(\boldsymbol{x}) = \sum_t J^{\text{data}}(\boldsymbol{x},t) = \sum_t \chi^2_{\boldsymbol{s}(t)}(\boldsymbol{x}) + \chi^2_{\boldsymbol{\beta}(t)}(\boldsymbol{x}), \quad (14)$$

where $\chi^2_{\boldsymbol{s}(t)}(\boldsymbol{x}) = \boldsymbol{s}^{\text{data}}(t)$ denotes the classical χ^2 statistic $(\boldsymbol{s}^{\text{data}}(t) - \boldsymbol{\alpha}^2(\boldsymbol{x},t))^{\mathrm{T}} \boldsymbol{R}^{-1}_{\boldsymbol{s}(t)} (\boldsymbol{s}^{\text{data}}(t) - \boldsymbol{\alpha}^2(\boldsymbol{x},t))$. Closure terms $\chi^2_{\boldsymbol{\beta}(t)}(\boldsymbol{x})$ are a weighted quadratic distance between complex phasors [13] instead of a χ^2 statistic over closure phase residuals. One then associates J^{data} with a regularization term to account for the incompleteness of the data in such inverse problems and minimizes the composite criterion

$$J(\boldsymbol{x}) = J^{\text{data}}(\boldsymbol{x}) + J^{\text{prior}}(\boldsymbol{x}) \quad (15)$$

under the following constraints:

$$\begin{aligned} &\forall (p,q), \qquad x(p,q) \geqslant 0, \\ &\sum_{p,q} x(p,q) = 1. \end{aligned} \quad (16)$$

The first requires positivity of the sought object, and the second is a constraint of unit flux. Indeed, fringe visibilities are by definition flux-normalized quantities [i.e., normalized by the FT of the object at the null frequency; see Eq. (3)], so the data are independent of the total flux of the sought object (of course an interferometer is sensitive

to the total flux of the source, but this last value is not contained in the fringe visibility itself).

The regularization term J^{prior} is chosen to enforce some properties of the object that are known *a priori* (smoothness, spiky behavior, positivity, etc.) and should also ease the minimization. Simple and popular regularization terms are convex separable penalizations of the object pixels (i.e., white priors) or of the object spatial derivatives (for instance, first-order derivative or gradient). In what follows, we quickly describe the prior terms used in this paper. These priors are more extensively described and compared in [14]. For a general review on regularization, see [15].

Entropic priors belong to the family of white priors and often allow one to obtain a clean image while preserving its sharp spiky features, whereas quadratic penalization tends to soften the reconstructed map. The white quadratic-linear (or $L_2L_1^w$) penalization given by

$$L_2L_1^w(\boldsymbol{x}) = \delta^2 \sum_{p,q} \frac{\boldsymbol{x}(p,q)}{s\delta} - \ln\left(1 + \frac{\boldsymbol{x}(p,q)}{s\delta}\right) \tag{17}$$

that we use in Section 5 leads to a kind of entropic regularization, in the sense of [16]. We propose a nominal setting of the two parameters δ and s:

$$s = 1/N_{\text{pix}}; \qquad \delta = 1. \tag{18}$$

As regards regularization based on the object's spatial derivatives, we shall consider here only quadratic penalization, but convex quadratic-linear L_2L_1 penalization functions could also be invoked.

Reference [17] is one of the works that adopts such a Bayesian approach for processing OLBI using a constrained local descent method to minimize Eq. (15). A convex data criterion J, i.e., such that $J(k\cdot x_1+(1-k)\cdot x_2) \leqslant k\cdot J(x_1)+(1-k)\cdot J(x_2)$, $\forall x_1,x_2$, $\forall k \in [0,1]$, has no local minima, which makes the minimization much easier. Unfortunately, the criterion J is nonconvex. To be more precise, the difficulty of the problem can be summed up as follows:

(i) The small number of Fourier coefficients makes the problem underdetermined. Here the regularization term and the positivity constraint can help by limiting the high frequencies of the reconstructed object [6].

(ii) Closure phase measurements imply missing phase information and make the Fourier synthesis problem nonconvex. Adding a regularization term does not generally correct the problem [18].

(iii) Phase and modulus measurements with additive Gaussian noise lead to a non-Gaussian likelihood and a nonconvex log-likelihood with respect to $\boldsymbol{x}$. As a consequence, even with no missing phases, some approximation of the real observable statistics is necessary to get a convex data fidelity term. This data conversion from polar to Cartesian coordinates, which is commonly used in the field of radar processing [19], has been studied only recently in OLBI [20]; see Subsection 3.C.

These characteristics imply that optimizing J by a local descent algorithm can work only if the initialization selects the "right" valley of the criterion. The design of a good initial position is very case dependent and will not be extensively addressed here. The other key aspects are then the followed path, i.e., the minimization method, and the shape of the function to minimize, i.e., the behavior of the criterion $\boldsymbol{x} \mapsto J(\boldsymbol{x})$. This paper addresses both aspects:

- We design a specific OLBI criterion $\mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha})$ where two sets of variables appear explicitly, one in the spatial domain $\boldsymbol{x}$, describing the sought object, and another in the Fourier phase domain $\boldsymbol{\alpha}$, which accounts for the missing phase information. This specific criterion is designed to solve (iii), i.e., so that for a known $\boldsymbol{\alpha}$, the criterion is convex with respect to $\boldsymbol{x}$. In other words, if we had all the complex visibility phase measurements instead of just the closure phases, our criterion $\boldsymbol{x} \mapsto \mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha})$ *would be convex*;
- We adopt an alternate minimization method, working on the two sets of variables. This approach can be related to "myopic" approaches of some inverse problems, where missing data concerning the instrumental response are modeled and sought for during the inversion [21]. Alternate minimization methods are inspired by self-calibration methods in radio interferometry and have been used in optical interferometry by Lannes *et al.* [6]. However, the criterion used in [6] was essentially imported from radio interferometry and does not match OLBI data model [13]. Our main contribution is to derive a criterion that accounts for data model (13), while allowing an efficient alternate minimization. This construction is the subject of the next section.

3. EQUIVALENT MYOPIC MODEL FOR SELF-CALIBRATION

The aim of this section is to approximate the data model of Eq. (13):

$$\boldsymbol{s}^{\text{data}}(t) = \boldsymbol{a}^2(\boldsymbol{x},t) + \boldsymbol{s}^{\text{noise}}(t), \qquad \boldsymbol{s}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{s}(t)}), \tag{19}$$

$$\boldsymbol{\beta}^{\text{data}}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}^{\text{noise}}(t)[2\pi],$$

$$\boldsymbol{\beta}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{\beta}(t)}) \tag{20}$$

by a myopic linear model with additive complex Gaussian noise of the following form:

$$\boldsymbol{y}^{\text{data}}(t) = \mathcal{F}_{\boldsymbol{\alpha}(t)} \cdot \boldsymbol{H}(t)\boldsymbol{x} + \boldsymbol{y}^{\text{noise}}(t), \tag{21}$$

where the operator $\cdot$ denotes componentwise multiplication and $\mathcal{F}_{\boldsymbol{\alpha}(t)}$ is a vector of phasors depending on phase aberration parameters $\boldsymbol{\alpha}(t)$, which are defined in Subsection 3.B. This will be done in three steps:

- Subsection 3.A is devoted to the derivation of the observation model for the pseudo amplitude term $\boldsymbol{a}^{\text{data}}(t)$ from Eq. (19).
- Subsection 3.B is devoted to the derivation of the observation model for the pseudo phase term $\boldsymbol{\phi}^{\text{data}}(t)$ from Eq. (20).
- Subsection 3.C shows how to combine pseudo phase and pseudo amplitude models in a complex model such as Eq. (21) while solving problem (iii) of Subsection 2.E.

A. Pseudo Amplitude Data Model

In Eq. (19), we have assumed a Gaussian distribution for $\boldsymbol{s}^{\text{data}}(t)$ around $\boldsymbol{s}(\boldsymbol{x},t)$, which is questionable, since

squared amplitudes should be nonnegative. However, such a statistic model is acceptable provided that the probability of a negative component of $\boldsymbol{s}^{\text{data}}(t)$ is very weak. For uncorrelated measurements, this assumption corresponds to mean values much greater than the corresponding standard deviation. Appendix D shows how to build the mean and covariance matrix of the square root of such a distribution. The mean vector is taken as the pseudo amplitude data $\boldsymbol{a}^{\text{data}}(t)$ and the covariance matrix called $\boldsymbol{R}_{\boldsymbol{a}(t)}$.

Observation model (19) can then be approximated by the following pseudo amplitude data model:

$$\boldsymbol{a}^{\text{data}}(t)=\boldsymbol{a}(\boldsymbol{x},t)+\boldsymbol{a}^{\text{noise}}(t),\qquad \boldsymbol{a}^{\text{noise}}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{a}(t)}). \tag{22}$$

B. Pseudo Phase Data Model

We start from a generalized inverse solution to the phase closure equation of Eq. (20). The generalized inverse $\boldsymbol{C}^{\dagger}$ of $\boldsymbol{C}$, defined by $\boldsymbol{C}^{\dagger}\triangleq\boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1}$, is such that $\boldsymbol{C}\boldsymbol{C}^{\dagger}=\mathbf{I}\mathrm{d}$. By applying it on all the terms of Eq. (20), we obtain

$$\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{data}}(t)=\boldsymbol{C}^{\dagger}\boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x},t)+\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t)+2\pi\boldsymbol{C}^{\dagger}\boldsymbol{\kappa}, \tag{23}$$

where $\boldsymbol{\kappa}$ is a vector of integers to account for the fact that each phase component is measured modulo 2π. We define

$$\boldsymbol{\phi}^{\text{data}}(t)\triangleq\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{data}}(t), \tag{24}$$

$$\boldsymbol{\phi}^{\text{ker}}(t)\triangleq(\boldsymbol{C}^{\dagger}\boldsymbol{C}-\mathbf{I}\mathrm{d})\boldsymbol{\phi}(\boldsymbol{x},t)+2\pi\boldsymbol{C}^{\dagger}\boldsymbol{\kappa} \tag{25}$$

and obtain

$$\boldsymbol{\phi}^{\text{data}}(t)=\boldsymbol{\phi}(\boldsymbol{x},t)+\boldsymbol{\phi}^{\text{ker}}(t)+\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t). \tag{26}$$

Vector $\boldsymbol{\phi}^{\text{ker}}(t)$ belongs to the 2π-wrapped kernel of operator $\boldsymbol{C}$:

$$\boldsymbol{C}\boldsymbol{\phi}^{\text{ker}}(t)=(\underbrace{\boldsymbol{C}\boldsymbol{C}^{\dagger}}_{=\mathbf{I}\mathrm{d}}\boldsymbol{C}-\boldsymbol{C})\boldsymbol{\phi}(\boldsymbol{x},t)+2\pi\underbrace{\boldsymbol{C}\boldsymbol{C}^{\dagger}}_{=\mathbf{I}\mathrm{d}}\boldsymbol{\kappa}=2\boldsymbol{\kappa}\pi=\boldsymbol{0}[2\pi].$$

As shown in Appendix C, if $\boldsymbol{\phi}^{\text{ker}}=0[2\pi]$, there exists a real vector $\boldsymbol{\alpha}(t)$ of dimension $N_{\mathrm{t}}-1$ such that $\boldsymbol{\phi}^{\text{ker}}(t)=\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)[2\pi]$, where $\bar{\boldsymbol{B}}$ is obtained by removing the first column of operator $\boldsymbol{B}$. So we have

$$\boldsymbol{\phi}^{\text{data}}(t)=\boldsymbol{\phi}(\boldsymbol{x},t)+\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)+\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t)[2\pi]. \tag{27}$$

Now the problem is that $\boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{\text{noise}}(t)$ is a zero-mean random vector with a *singular covariance matrix*:

$$\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}\triangleq\boldsymbol{C}^{\dagger}\boldsymbol{R}_{\boldsymbol{\beta}(t)}\boldsymbol{C}^{\dagger\mathrm{T}}.$$

To obtain a strictly convex log-likelihood, we have to approximate this term by a proper Gaussian vector $\boldsymbol{\phi}^{\text{noise}}(t)$, with an invertible covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$ chosen so as to correctly fit the second-order statistics of the noise in phase closure measurement equation (20). This last requirement can be written as the following equation:

$$\boldsymbol{C}\boldsymbol{R}_{\boldsymbol{\phi}(t)}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{R}_{\boldsymbol{\beta}(t)}. \tag{28}$$

In other words, we are led to choose an invertible covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$ so as to mimic the statistical behavior of the closures, which is expressed by Eq. (28).

We propose to modify matrix $\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}$ by setting its nondiagonal components to 0, *i.e.*, to use the following diagonal matrix:

$$\{\boldsymbol{R}_{\boldsymbol{\phi}(t)}\}_{ij}=\begin{cases}3\cdot\{\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}\}_{ij} & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}. \tag{29}$$

The factor 3 allows us to preserve the total weight of the phase term in the log-likelihood by satisfying the condition

$$\sum_{i,j}|\{\boldsymbol{R}_{\boldsymbol{\phi}(t)}\}_{ij}|=\sum_{i,j}|\{\boldsymbol{R}^{0}_{\boldsymbol{\phi}(t)}\}_{ij}|.$$

There are several ways of choosing $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$, and we propose this particular choice without claiming it is optimal. Note that the myopic model derived in what follows can accommodate to any choice of a proper (*i.e.*, invertible) covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}(t)}$.

With Eqs. (24), (27), and (29), we obtain the visibility phase pseudo data model:

$$\boldsymbol{\phi}^{\text{data}}(t)=\boldsymbol{\phi}(\boldsymbol{x},t)+\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)+\boldsymbol{\phi}^{\text{noise}}(t)[2\pi],$$

$$\boldsymbol{\phi}^{\text{noise}}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}). \tag{30}$$

C. Pseudo Complex Visibility Data Model

Gathering Eqs. (22) and (30), we have finally approximated the data model [Eqs. (19) and (20)] by

$$\begin{cases}\boldsymbol{a}^{\text{data}}(t)=\boldsymbol{a}(\boldsymbol{x},t)+\boldsymbol{a}^{\text{noise}}(t),\\ \boldsymbol{\phi}^{\text{data}}(t)=\boldsymbol{\phi}(\boldsymbol{x},t)+\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)+\boldsymbol{\phi}^{\text{noise}}(t)[2\pi],\\ \text{with } \boldsymbol{a}^{\text{noise}}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{a}(t)}),\quad \boldsymbol{\phi}^{\text{noise}}(t)\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}).\end{cases} \tag{31}$$

We form pseudo complex visibility measurements $\boldsymbol{y}^{\text{data}}(t)$ defined by

$$\boldsymbol{y}^{\text{data}}(t)\triangleq\boldsymbol{a}^{\text{data}}(t)\cdot\mathrm{e}^{\mathrm{i}\boldsymbol{\phi}^{\text{data}}(t)}. \tag{32}$$

The approach proposed in [20], which we recall and generalize in Appendix E, is based on an approximated complex visibility data model:

$$\boldsymbol{y}^{\text{data}}(t)=\boldsymbol{H}(t)\boldsymbol{x}\cdot\mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}+\boldsymbol{y}^{\text{noise}}(t). \tag{33}$$

This is exactly the sought model stated at the beginning of this section in Eq. (21), with $\mathcal{F}_{\boldsymbol{\alpha}(t)}=\mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}$. We now define the myopic observation model as follows:

$$\boldsymbol{y}_{m}(\boldsymbol{x},\boldsymbol{\alpha}(t))\triangleq\boldsymbol{H}(t)\boldsymbol{x}\cdot\mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}. \tag{34}$$

As shown in Appendix E, the mean value $\bar{\boldsymbol{y}}^{\text{noise}}(t)$ and covariance matrix $\boldsymbol{R}_{\boldsymbol{y}^{\text{noise}}(t)}$ of the additive complex noise term $\boldsymbol{y}^{\text{noise}}(t)$ are carefully designed so that the corresponding data likelihood criterion is convex quadratic with respect to the complex $\boldsymbol{y}_{m}(\boldsymbol{x},\boldsymbol{\alpha}(t))$ while remaining close to the real nonconvex model. To illustrate these properties, we consider one complex visibility and plot in the complex plane the distribution of $\boldsymbol{y}^{\text{data}}(t)$ around $\boldsymbol{y}_{m}(\boldsymbol{x},\boldsymbol{\alpha}(t))$ for the true noise distribution—i.e., a polar Gaussian noise in phase and modulus—and our Cartesian Gaussian approximation (see Fig. 1) In particular, the "elliptic" covariance matrix we propose (which yields elliptic contour plots in Fig. 1) is preferable to the more classical

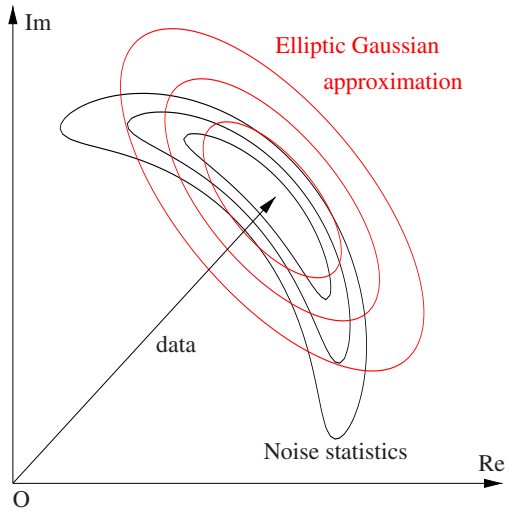


Fig. 1. (Color online) Contour plots of a polar Gaussian distribution and of its Cartesian Gaussian approximation.

"circular" approximation that appears in previous contributions on OLBI [22]. The latter can be described by half as many parameters as needed for the elliptic one (one radius for a circle, instead of a short axis and a long axis for an ellipsis), but it is clearly less accurate [20] (such a noise statistics description has also been investigated for the complex bispectra in the OIFITS data exchange format [11]).

From Eq. (33), we build Chi-2 statistics over real and imaginary parts of the observation equation

$$\chi^2_{\boldsymbol{y}(t)}(\boldsymbol{x},\boldsymbol{\alpha}(t)) \triangleq \begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}^{\mathrm{data}}(t)-\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))-\bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \\ \mathrm{Im}\{\boldsymbol{y}^{\mathrm{data}}(t)-\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))-\bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \end{bmatrix}^{\mathrm{T}} \times \boldsymbol{R}_{\boldsymbol{y}^{\mathrm{noise}}(t)}^{-1} \begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}^{\mathrm{data}}(t)-\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))-\bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \\ \mathrm{Im}\{\boldsymbol{y}^{\mathrm{data}}(t)-\boldsymbol{y}_m(\boldsymbol{x},\boldsymbol{\alpha}(t))-\bar{\boldsymbol{y}}^{\mathrm{noise}}(t)\} \end{bmatrix}.$$

And we finally propose the myopic goodness-of-fit criterion:

$$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_t \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}(t),t) = \sum_t \chi^2_{\boldsymbol{y}(t)}(\boldsymbol{x},\boldsymbol{\alpha}(t)). \quad (35)$$

We can now design a myopic Bayesian approach to the reconstruction problem by combining the data term with a regularization term along the lines of Subsection 2.E:

$$\mathcal{J}(\boldsymbol{x},\boldsymbol{\alpha}) = \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) + J^{\mathrm{prior}}(\boldsymbol{x}). \quad (36)$$

The next section describes an alternate minimization technique applied to regularized criterion (36).

4. WISARD

In this section, we describe WISARD, standing for Weak-phase Interferometric Sample Alternating Reconstruction Device, a self-calibration method for OLBII.

A. Global Structure of WISARD

WISARD is made of four major blocks:

- A first block recasts the raw data (i.e., closure phases and squared visibilities) in myopic data (i.e., phases and moduli) as described in Subsections 3.A and 3.B.
- A second "convexification block" computes a Gaussian approximation of the pseudo visibility data model as described in Subsection 3.C.
- A third block builds a guess for the object $\boldsymbol{x}$ and aberrations $\boldsymbol{\alpha}$ (i.e., a good starting point).
- Finally, the self-calibration block performs the minimization of regularized criterion (36), under constraints (16). It alternates optimization of the object for given aberrations and optimization of the aberrations for the current object.

The structure of WISARD is sketched in Fig. 2. The principles that underline the three first blocks of WISARD have been described in previous sections, while details on the self-calibration minimization are gathered in the next one.

B. Self-Calibration Block

In the following, we describe the three key components of the self-calibration block.

Minimization with respect to $\boldsymbol{x}$. The criterion $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha})$ we have derived is quadratic and hence convex with respect to the object $\boldsymbol{x}$. Hence the minimization versus $\boldsymbol{x}$ does not raise special difficulties.

Minimization with respect to $\boldsymbol{\alpha}$. $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha})$ is the sum of terms involving only measurements obtained at one time instant t [Eq. (35)]:

$$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_t \mathcal{J}^{\mathrm{data}}(\boldsymbol{x},\boldsymbol{\alpha}(t),t).$$

Because the time between two measurements is much greater than the turbulence coherent time (around

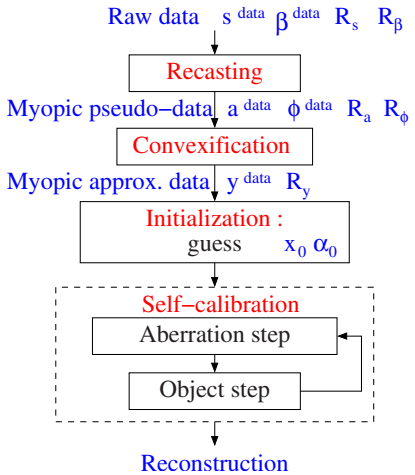


Fig. 2. (Color online) WISARD algorithm loop.

10 ms), aberrations $\boldsymbol{\alpha}(t)$ at two different instants are statistically independent. We can then solve separately for each set of $\boldsymbol{\alpha}(t)$, which dramatically reduces the complexity of the minimization. The number of $\boldsymbol{\alpha}(t)$ components to solve for is (N_t-1) and the minimization is delicate, as the criterion exhibits periodic structures that have been studied in [22].

However, exact minimization is affordable for a three-telescope interferometric array. In this case we have to perform several two-parameter minimizations, and each one can be efficiently initialized by an exhaustive search on a 2D grid, which ensures we avoid local minima. On the other hand, when N_t gets high enough, e.g., 6, then the number of $\boldsymbol{\alpha}(t)$ to solve for, e.g., 5, gets small compared to the number of closure phases available, e.g., 15. With a three-telescope array, 2/3 of the phase information is missing, whereas with a six-telescope array, only 1/3 of the phase information is missing. In this last case, which corresponds to the processing of synthetic data presented in Subsection 5.A, the reconstructions were straightforward, and no effects of the local minima in $\boldsymbol{\alpha}$ were witnessed.

In other words, coping with the ambiguities in $\boldsymbol{\alpha}$, for instance, with the specific criterion proposed in [22], may be necessary only for $N_t=4$ or $N_t=5$. For $N_t=3$, an exhaustive search is possible, and for $N_t\geqslant 6$, ambiguities in $\boldsymbol{\alpha}$ do not have, according to our experience, a major impact on reconstruction.

Starting point: object and aberration guess $\boldsymbol{x}_0$ *and* $\boldsymbol{\alpha}_0$. If a parametric model of the observed stellar source is not available, the object starting point is a mean square solution, from which we extract the positive part. The first step in the self-calibration block is a minimization with respect to $\boldsymbol{\alpha}$ for $\boldsymbol{x}=\boldsymbol{x}_0$.

5. RESULTS

This section presents some results of processing by the WISARD algorithm, with both synthetic and experimental data.

A. Processing of Synthetic Data

The first example takes synthetic interferometric data that were used in the international Imaging Beauty Contest organized by P. Lawson for the International Astronomical Union (IAU) [23]. These data simulate the observation of the synthetic object shown in Fig. 3 with the Navy Prototype Optical Interferometer (NPOI) [24] six-telescope interferometer. The corresponding frequency coverage, shown in Fig. 3, has a structure in arcs of circles typical of the supersynthesis technique, which consists in repeating the measurements over several nights of observation so that the same baselines access different measurement spatial frequencies because of the Earth's rotation. In total, there are 195 square visibility modules and 130 closure phases, together with the associated variances.

Six reconstructions obtained with WISARD are shown in Fig. 4. On the upper row is a reconstruction using a quadratic regularization based on a power spectral density model in $1/|u|^3$ for a weak, a strong, and a correct regularization parameter. The latter gives a satisfactory level of smoothing but does not restore the peak in the center of the object. The peak is visible in the under-regularized reconstruction on the left but at the cost of too high a residual variance.

The reconstruction presented on the lower row is a good trade-off between smoothing and restoration of the central peak thanks to the use of the white $L_2L_1^w$ prior term introduced in Subsection 2.E. The automatically set parameters [Eq. (18)] are very satisfactory (left), and a light tuning (center and right) allows an even better reconstruction. The goodness of fit of the $L_2L_1^w$ reconstruction can be appreciated in Fig. 5. The crosses (red online) show the reconstructed visibility moduli (i.e., of the FT of the reconstructed object at the measurement frequencies), and the squares (blue online) are the moduli of the measured visibilities. The difference between the two, weighted by 10 times the standard deviation of the moduli, is shown as the dotted curve. The mean value of this difference is 0.1, which shows a good fit (to within 1 σ).

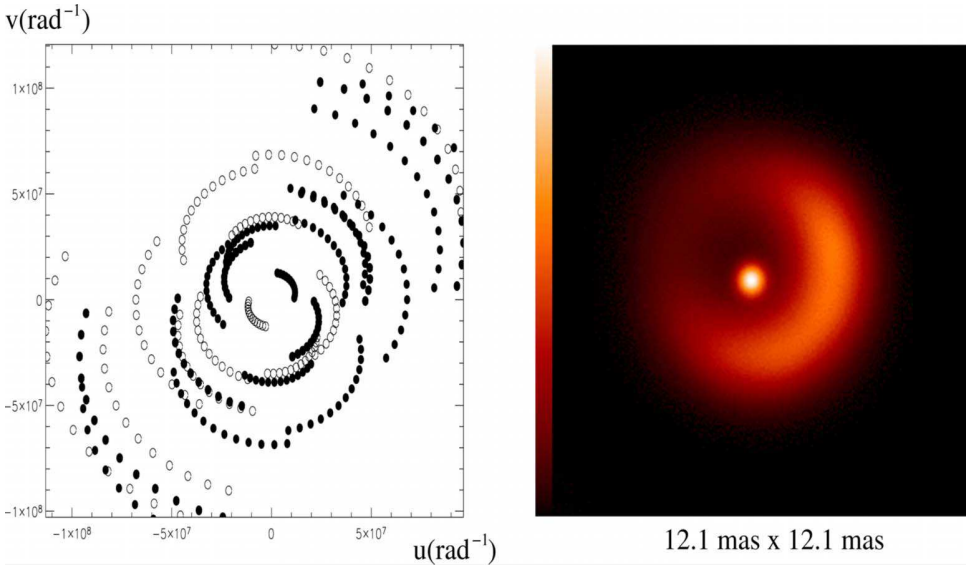


Fig. 3. (Color online) Synthetic object (right) and frequency coverage (left) from the Imaging Beauty Contest 2004.

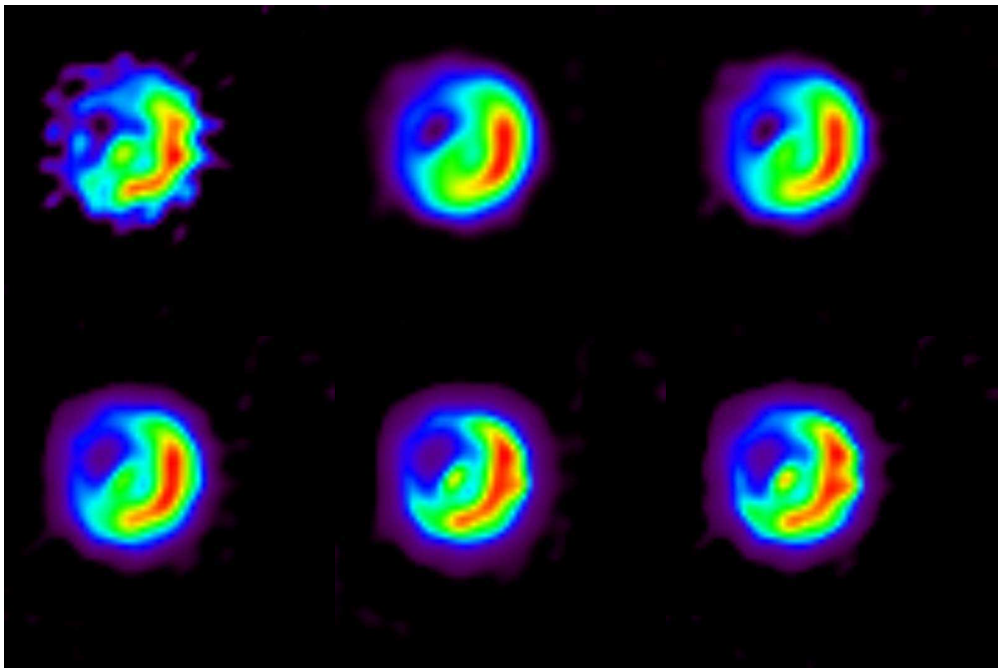


Fig. 4. (Color online) Reconstructions with WISARD. Upper row, under-regularized quadratic model (left), over-regularized quadratic model (center), quadratic model with correct regularization parameter (right). Lower row, white $L_2L_1^{uu}$ model with automatically set scale and delta parameters (left), white $L_2L_1^{uu}$ model with half-scale (center), white $L_2L_1^{uu}$ model with half-delta (right). Each image field is 12.1×12.1 mas.

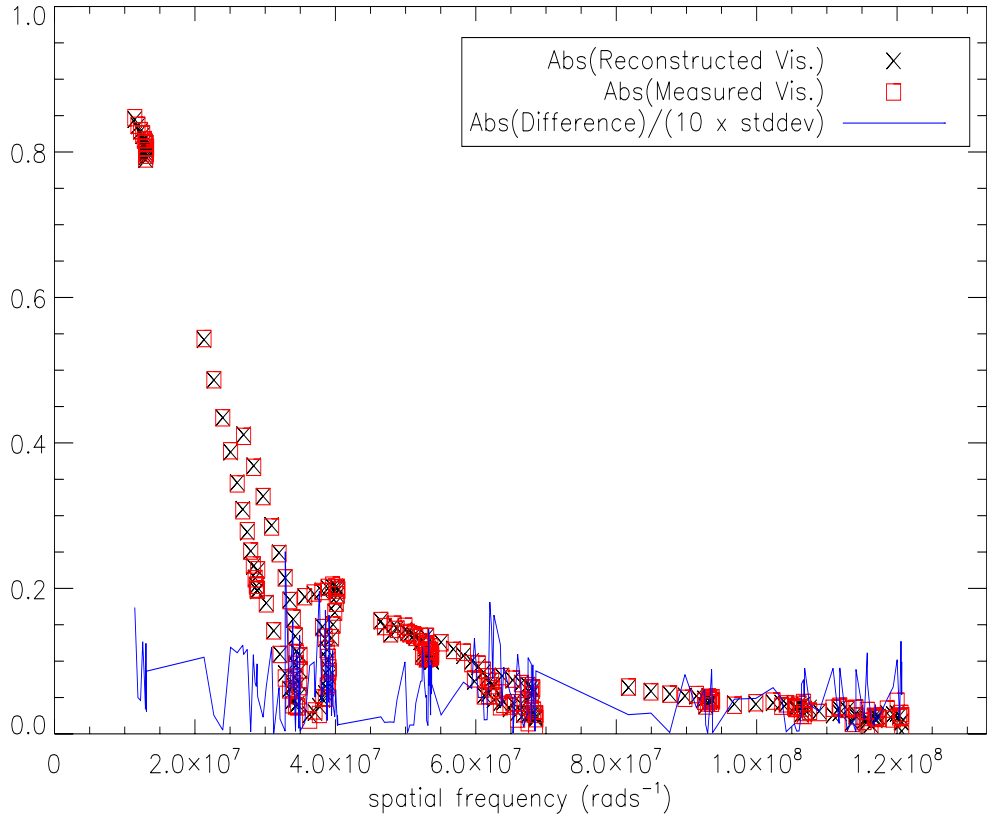


Fig. 5. (Color online) Goodness of fit at WISARD convergence.

B. Processing of Experimental Data

Here we present the reconstruction of the star χ Cygni from experimental data using the WISARD algorithm. The data were obtained by S. Lacour and S. Meimon under the leadership of G. Perrin during a measuring campaign on the IR/Optical Telescope Array (IOTA) interferometer [25] in May 2005. As already mentioned, each measurement has to be calibrated by observation of an object that acts as a point source at the instrument's resolving power. The calibrators chosen were HD 180450 and HD 176670.

The star χ Cygni is a Mira-type star, Mira itself being an example of such stars. Perrin *et al.* [26] propose a model of Mira-type stars, composed of a photosphere, an empty layer, and a thin molecular layer. The aim of the mission was to obtain images of χ Cygni in the H band ($1.65\ \mu\text{m} \pm 175$ nm) and, in particular, to highlight possible assymmetric features in the structure of the molecular layer.

Figure 6 shows, on the left, the $u-v$ coverage obtained, i.e., the set of spatial frequencies measured, multiplied by the observation wavelength. Because the sky is habitually represented with the west on the right, the coordinates used are, in fact, $-u, v$. The domain of the accessible $u-v$ plane is constrained by the geometry of the interferometer and the position of the star in the sky. The "hourglass" shape is characteristic of the IOTA interferometer, and entails nonuniform resolution that affects the image reconstruction, shown on the right. The reconstructed angular field has sides of 60 mas. In addition to the positivity constraint, the regularization term used is the $L_2L_1^w$ term described in Subsection 2.E. The interested reader will find an astrophysical interpretation of this result in [27].

6. CONCLUDING COMMENTS

We have proposed a complete and precise self-calibration approach to optical interferometry image reconstruction. After pointing out the data model specificities in the OLBI context, we have emphasized the sources of underdeterminations, which make a classical Bayesian criterion descent method critical. Namely, the main problems are the phase underdetermination caused by turbulence effects, and, as noted only recently, the polar coordinate structure of the data model.

We have built a specially designed approximate myopic data model in order to derive a self-calibration method. Special care was given to the design of the second-order statistics of the myopic model, an aspect that was ignored in previous related works.

We have extended our previous work on polar data conversion [20] and proposed a convex approximation of the noise model that reduces the number of local minima of the criterion to minimize.

We also addressed integer ambiguities induced by closure phase wrapping, which are classical when dealing with phase data, and have discussed their impact on the image reconstruction quality: for three-telescope data, we have proposed an exhaustive search method, and we have witnessed that these ambiguities do not raise any particular problem when processing the interferometer data of six or more telescopes. Concerning the remaining case of four to five telescopes, the work by Lannes [22] should be worth investigating. On the other hand, global minimization methods were left aside because of their intensive computation needs. As computer performance increases, these methods might be, in the years to come, an appropriate way to deal with local minima.

All these developments allowed us to propose WISARD, a self-calibration method for OLBII reconstruction and to demonstrate its efficiency on simulated data.

Finally, WISARD was also used to successfully process real astronomical OLBI data sets. These results were made possible thanks to a close partnership with the astronomers Sylvestre Lacour and Guy Perrin of the Observatoire de Paris Meudon, within the PHASE partnership (Partenariat Haute résolution Angulaire Sol-Espace). Indeed, an accurate astronomical model of the observed stellar object is a precious guideline for reconstructing a complex image from OLBI data. From the author's point

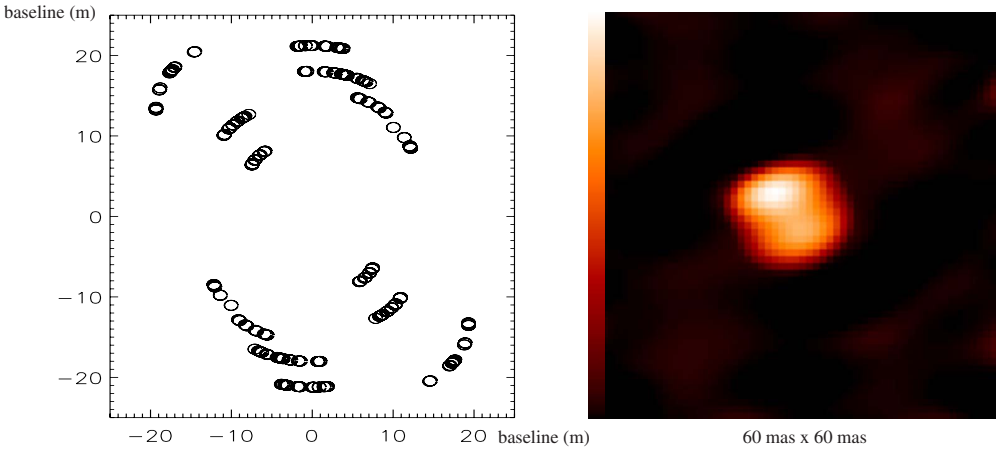


Fig. 6. (Color online) Frequency coverage (left) and reconstruction of the star χ Cygni (right).

of view, such a collaboration is essential to the success of OLBII techniques.

APPENDIX A: BASELINE AND CLOSURE OPERATORS C AND B

Let N_t be the number of telescopes of the interferometric array. We have the following definitions:

$$\boldsymbol{B}_2 \triangleq [-1 \quad 1], \tag{A1}$$

$$\boldsymbol{B}_{N_t} \triangleq \left[\begin{array}{cc} -\mathbf{1}_{N_t-1} & \mathrm{Id}_{N_t-1} \\ \hline \boldsymbol{O} & \boldsymbol{B}_{N_t-1} \end{array}\right], \tag{A2}$$

$$\boldsymbol{C}_{N_t} \triangleq \left[-\boldsymbol{B}_{N_t-1} \quad \mathrm{Id}_{[(N_t-1)(N_t-2)]/2}\right], \tag{A3}$$

for $N_t \geqslant 3$.

In what follows, we prove that $\ker \boldsymbol{C} = \mathrm{im}\, \boldsymbol{B}$.

We have $\boldsymbol{C}_{N_t}\boldsymbol{B}_{N_t} = 0$, so

$$\mathrm{im}\, \boldsymbol{B} \subset \ker \boldsymbol{C}. \tag{A4}$$

It is straightforward to prove by recurrence that $\boldsymbol{B}_{N_t} \cdot \mathbf{1}_{N_t} = 0$, which yields $\mathrm{rank}\, \boldsymbol{B}_{N_t} \leqslant N_t - 1$. Because $\boldsymbol{B}_{N_t}$ contains Id_{N_t-1}, we gather

$$\dim \mathrm{im}\, \boldsymbol{B} \triangleq \mathrm{rank}\, \boldsymbol{B} = N_t - 1. \tag{A5}$$

Here $\boldsymbol{C}_{N_t}$ contains $\mathrm{Id}_{(N_t-1)(N_t-2)/2}$, which yields $\mathrm{rank}\, \boldsymbol{C}_{N_t} \geqslant (N_t-1)(N_t-2)/2$, or

$$\dim \ker \boldsymbol{C}_{N_t} \leqslant N_t - 1. \tag{A6}$$

With Eqs. (A4)–(A6), we gather

$$\ker \boldsymbol{C} = \mathrm{im}\, \boldsymbol{B}. \tag{A7}$$

APPENDIX B: CHARACTERIZATION OF THE BASELINE PHASE-INDEPENDENT OPERATORS

Here we prove that any continuous differentiable function f verifying property (8)

$$f(\boldsymbol{\phi} + \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{\phi}), \quad \forall\, (\boldsymbol{\phi}, \boldsymbol{\varphi})$$

is such that $f(\boldsymbol{\phi}) = g(\boldsymbol{C}\boldsymbol{\phi})$, where $\boldsymbol{C}$ has more columns than rows, so its pseudo inverse is defined by $\boldsymbol{C}^\dagger \triangleq \boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1}$ and verifies

$$\boldsymbol{C}\boldsymbol{C}^\dagger = \mathrm{Id} \tag{B1}$$

and thus

$$\boldsymbol{C}\boldsymbol{C}^\dagger\boldsymbol{C} - \boldsymbol{C} = 0 \Rightarrow \boldsymbol{C}(\boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{\phi}) = 0,$$

$$\forall\, \boldsymbol{\phi} \overset{\mathrm{A7}}{\Rightarrow} \exists\, \boldsymbol{\varphi}, (\boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{\phi}) = \boldsymbol{B}\boldsymbol{\varphi},$$

$$\forall\, \boldsymbol{\phi} \Rightarrow \exists\, \boldsymbol{\varphi}, \boldsymbol{\phi} = \boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{B}\boldsymbol{\varphi}, \quad \forall\, \boldsymbol{\phi}.$$

With this we obtain that any f verifying property (8) is such that

$$f(\boldsymbol{\phi}) = f(\boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{\phi} - \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{\phi}) = g(\boldsymbol{C}\boldsymbol{\phi}).$$

APPENDIX C: WRAPPED KERNEL OF OPERATOR C

The kernel of operator $\boldsymbol{C}$ is given by $\ker \boldsymbol{C} = \mathrm{im}\, \boldsymbol{B}$ [Eq. (A7)]. With dimensional arguments, it is easy to see that

$$\mathrm{im}\, \boldsymbol{B} = \mathrm{im}\, \bar{\boldsymbol{B}},$$

where $\bar{\boldsymbol{B}}$ is obtained by removing the first column of operator $\boldsymbol{B}$, so we have

$$\ker \boldsymbol{C} = \mathrm{im}\, \bar{\boldsymbol{B}}. \tag{C1}$$

Let us now characterize the set of $\boldsymbol{\phi}^{\ker}$ such that

$$\boldsymbol{C}\boldsymbol{\phi}^{\ker} \equiv 0[2\pi].$$

Because $\boldsymbol{C}$ has integer components, $\boldsymbol{\phi}^{\ker}$ can be considered modulo 2π. With Eq. (C1), we obtain

$$\exists\, \boldsymbol{\alpha}_1,\ \boldsymbol{\phi}^{\ker} \equiv \boldsymbol{C}^\dagger(0[2\pi]) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}_1[2\pi]. \tag{C2}$$

Because $\bar{\boldsymbol{B}}$ has integer components, $\boldsymbol{\alpha}_1$ can be considered modulo 2π. The issue here is to evaluate the $\boldsymbol{C}^\dagger(0[2\pi])$ term, i.e., the value of $\boldsymbol{C}^\dagger(2\pi\boldsymbol{\kappa})$, with $\boldsymbol{\kappa}$ any integer vector.

Equations (A1) show that $\boldsymbol{C} = [\boldsymbol{M}|\mathrm{Id}]$. The integer vector

$$\boldsymbol{\mu} \triangleq \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix}$$

is then such that

$$\boldsymbol{C}\boldsymbol{\mu} = [*|\mathrm{Id}]\begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix} = \boldsymbol{\kappa}.$$

Then we have

$$\boldsymbol{C}\boldsymbol{\mu} = \boldsymbol{\kappa} \Rightarrow \boldsymbol{C}\boldsymbol{\mu}' = \boldsymbol{C}\boldsymbol{C}^\dagger\boldsymbol{\kappa} \Rightarrow \boldsymbol{C}(\boldsymbol{C}^\dagger\boldsymbol{\kappa} - \boldsymbol{\mu}) = 0 \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$\boldsymbol{C}^\dagger\boldsymbol{\kappa} - \boldsymbol{\mu} = \boldsymbol{B}\boldsymbol{\alpha}_2 \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa} = 2\pi\boldsymbol{\mu} + \boldsymbol{B}(2\pi\boldsymbol{\alpha}_2) \Rightarrow \exists\, \boldsymbol{\alpha}_2,$$

$$\boldsymbol{C}^\dagger(0[2\pi]) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}_1 \equiv \boldsymbol{B}\underbrace{(2\pi\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_1)}_{\boldsymbol{\alpha}}[2\pi].$$

So Eq. (C2) yields

$$\exists\, \boldsymbol{\alpha}, \boldsymbol{\phi}^{\ker} \equiv \bar{\boldsymbol{B}}\boldsymbol{\alpha}[2\pi]. \tag{C3}$$

APPENDIX D: SQUARE ROOT OF A GAUSSIAN DISTRIBUTION

Let us assume we measure the squared value s of a positive value a, with an additive Gaussian noise:

$$s^{\mathrm{data}} = a^2 + s^{\mathrm{noise}}, \tag{D1}$$

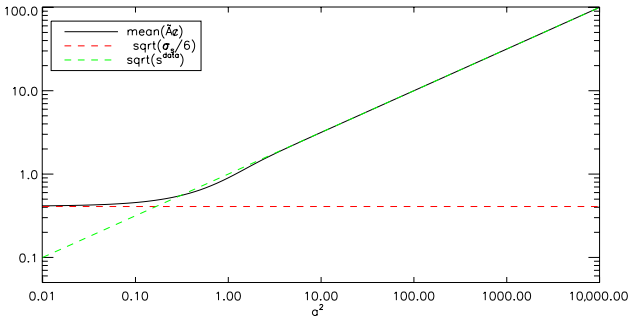


Fig. 7. (Color online) Behavior of $\langle \hat{a} \rangle$ in function of a^2 with a unit σ_s.

with s^{noise} being zero-mean Gaussian with the variance σ_s^2. Let $\hat{a}$ be the estimator of a from s^{data} defined by

$$\hat{a} = \begin{cases} \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} > 0 \\ 0 & \text{else} \end{cases},$$

where $\hat{a}$ can be seen as pseudo data. The data model of $\hat{a}$ derived from Eq. (D1) is not additive Gaussian. As will be shown in Appendix E, an optimal Gaussian approximation of the data model of $\hat{a}$ would be

$$\hat{a} = a + a^{\text{noise}}, \tag{D2}$$

with a^{noise} a Gaussian noise with a mean equal to $\langle \hat{a} \rangle$ and a standard deviation $\sqrt{\text{Var}(\hat{a})}$.

We have studied the behavior of the mean $\langle \hat{a} \rangle$ and standard deviation $\sqrt{\text{Var}(\hat{a})}$ of this estimator for various values of a^2, with a unit σ_s (see Figs. 7 and 8). We can distinguish two regimes for $\langle \hat{a} \rangle$:

- A low-mean regime, where $a^2 \leqslant \sigma_s/6$: a nonnegligible part of the distribution of s^{data} around a^2 is in the negative domain. Because $\hat{a}$ estimates a null value for a when s^{data} is negative, its mean will depend mainly on the width of the Gaussian wings. A good approximation of $\langle \hat{a} \rangle$ is $\sqrt{\sigma_s/6}$.
- A high-mean regime, where $a^2 \geqslant \sigma_s/6$: most of the distribution of s^{data} around a^2 is in the positive domain. The fact that $\hat{a}$ estimates a null value for a when $s^{\text{data}} < 0$ does not affect its mean $\langle \hat{a} \rangle$, which is close to a. Because a is not known, we choose $\langle \hat{a} \rangle = \sqrt{s^{\text{data}}}$. We can distinguish the same two regimes for $\sqrt{\text{Var}(\hat{a})}$. However, the transition is around σ_s:
- When $a^2 \leqslant \sigma_s$, the fact that $\hat{a}$ estimates a null value for a when s^{data} is negative tends to diminish its standard deviation, which we approximate by $\sqrt{\text{Var}(\hat{a})} \simeq \sqrt{\sigma_s}/2$.
- In the high-mean regime, where $a^2 \geqslant \sigma_s$, most of the distribution of s^{data} around a^2 is in the positive domain, and $\sqrt{\text{Var}(\hat{a})}$ is close to the classical expression. This expression corresponds to a first-order expansion in σ_a:

$$(a + \sigma_a)^2 = a^2 + \sigma_s \Rightarrow 2a\sigma_a \simeq \sigma_s,$$

where $\sigma_s/2a$. Because a is not known, we choose $\sqrt{\text{Var}(\hat{a})} = \sigma_s/2\sqrt{s^{\text{data}}}$. We then propose the pseudo data model

$$a^{\text{data}} = a + a^{\text{noise}},$$

with

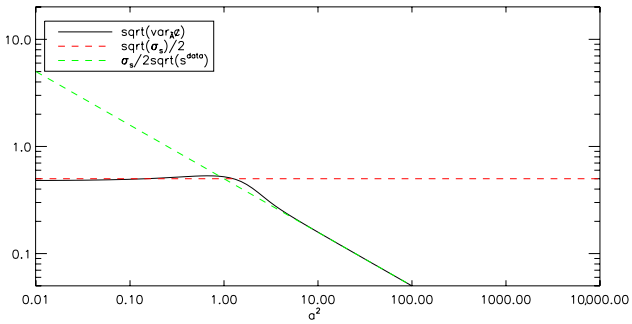


Fig. 8. (Color online) Behavior of $\sqrt{\text{Var}(\hat{a})}$ in function of a^2 with a unit σ_s.

$$a^{\text{data}} = \begin{cases} \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} > 0 \\ 0 & \text{else} \end{cases}$$

and a^{noise} a Gaussian noise with mean and standard deviation defined by

$$\bar{a} = \begin{cases} \sqrt{\sigma_s/6} & \text{if } s^{\text{data}} \leqslant \sigma_s/6, \\ \sqrt{s^{\text{data}}} & \text{if } s^{\text{data}} \geqslant \sigma_s/6 \end{cases},$$

$$\sigma_a = \begin{cases} \sqrt{\sigma_s}/2 & \text{if } s^{\text{data}} \leqslant \sigma_s \\ \dfrac{\sigma_s}{2\sqrt{s^{\text{data}}}} & \text{if } s^{\text{data}} \geqslant \sigma_s \end{cases}.$$

We also decide to discard the data such that $s^{\text{data}} \leqslant -\sigma_s$.

APPENDIX E: CARTESIAN GAUSSIAN APPROXIMATION TO A POLAR GAUSSIAN DISTRIBUTION

If we define

$$\boldsymbol{y}_{\boldsymbol{\alpha}(t)}(\boldsymbol{x},t) \triangleq \boldsymbol{H}(t)\boldsymbol{x} \cdot \mathrm{e}^{\mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t)}, \tag{E1}$$

Eq. (31) reads

$$\begin{cases} \boldsymbol{a}^{\text{data}}(t) = |\boldsymbol{y}_{\boldsymbol{\alpha}(t)}|(\boldsymbol{x},t) + \boldsymbol{a}^{\text{noise}}(t), & \boldsymbol{a}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{a}(t)}), \\ \boldsymbol{\phi}^{\text{data}}(t) \overset{2\pi}{\equiv} \arg \boldsymbol{y}_{\boldsymbol{\alpha}(t)}(\boldsymbol{x},t) + \boldsymbol{\phi}^{\text{noise}}(t), & \boldsymbol{\phi}^{\text{noise}}(t) \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{\phi}(t)}). \end{cases} \tag{E2}$$

1. General Expression

We consider a polar distribution of a Gaussian vector $\boldsymbol{y}$ of modulus $\boldsymbol{a}$ and phase $\boldsymbol{\phi}$:

$$\boldsymbol{\phi}^{\text{data}} = \bar{\boldsymbol{\phi}} + \boldsymbol{\phi}^{\text{noise}}, \tag{E3}$$

$$\boldsymbol{a}^{\text{data}} = \bar{\boldsymbol{a}} + \boldsymbol{a}^{\text{noise}}, \tag{E4}$$

where $\boldsymbol{\phi}^{\text{noise}}$ and $\boldsymbol{a}^{\text{noise}}$ are zero-mean real Gaussian vectors of covariance matrices $\boldsymbol{R}_{\boldsymbol{a}}$ and $\boldsymbol{R}_{\boldsymbol{\phi}}$ (the vectors $\boldsymbol{\phi}^{\text{noise}}$ and $\boldsymbol{a}^{\text{noise}}$ are assumed uncorrelated).

With the definitions

$$\begin{cases} \bar{\boldsymbol{y}} \triangleq \bar{\boldsymbol{a}} \exp \mathrm{i}\bar{\boldsymbol{\phi}}, \\ \boldsymbol{y}^{\mathrm{noise}} \triangleq \boldsymbol{y}^{\mathrm{data}} - \bar{\boldsymbol{y}}, \\ \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} \triangleq \mathrm{Re}\{\boldsymbol{y}^{\mathrm{noise}} \mathrm{e}^{-\mathrm{i}\bar{\boldsymbol{\phi}}}\}, \\ \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} \triangleq \mathrm{Im}\{\boldsymbol{y}^{\mathrm{noise}} \mathrm{e}^{-\mathrm{i}\bar{\boldsymbol{\phi}}}\}, \\ \bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}} \triangleq \begin{bmatrix} \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} \\ \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} \end{bmatrix}, \end{cases} \tag{E5}$$

we gather

$$\begin{cases} \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} = [\bar{\boldsymbol{a}} + \boldsymbol{a}^{\mathrm{noise}}] \cos \boldsymbol{\phi}^{\mathrm{noise}} - \bar{\boldsymbol{a}}, \\ \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} = [\bar{\boldsymbol{a}} + \boldsymbol{a}^{\mathrm{noise}}] \sin \boldsymbol{\phi}^{\mathrm{noise}}. \end{cases} \tag{E6}$$

A complex vector is Gaussian if and only if each of its components is Gaussian. A complex is Gaussian if and only if, in any Cartesian basis, its two components are Gaussian. So $\boldsymbol{y}$ is Gaussian if and only if $\bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}}$ is Gaussian, which is not the case [20]. In what follows, we show how to optimally approximate the distribution of $\bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}}$ by a Gaussian distribution.

2. Gaussian Approximation

We characterize our Cartesian additive Gaussian approximation, *i.e.,* its mean $\langle \bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}} \rangle$ and covariance $\mathbf{R}_{\bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}}}$, by minimizing the Kullback–Leibler distance between the two noise distributions, which gives [20]

$$\begin{cases} \langle \bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}} \rangle = E\left\{ \begin{bmatrix} \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} \\ \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} \end{bmatrix} \right\} = \begin{bmatrix} \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{rad}} \\ \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{tan}} \end{bmatrix}, \\ \boldsymbol{R}_{\bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}}} = E\left\{ \begin{bmatrix} \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{rad}} - \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} \\ \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{tan}} - \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{rad}} - \boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}} \\ \bar{\boldsymbol{y}}^{\mathrm{n}}_{\mathrm{tan}} - \boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}} \end{bmatrix}^{\mathrm{T}} \right\}, \end{cases} \tag{E7}$$

and we define

$$\boldsymbol{R}_{\bar{\bar{\boldsymbol{y}}}^{\mathrm{noise}}} \triangleq \begin{bmatrix} \boldsymbol{R}_{\mathrm{rad,rad}} & \boldsymbol{R}_{\mathrm{rad,tan}} \\ \boldsymbol{R}^{\mathrm{T}}_{\mathrm{rad,tan}} & \boldsymbol{R}_{\mathrm{tan,tan}} \end{bmatrix}.$$

For a zero-mean Gaussian vector $\boldsymbol{\phi}^{\mathrm{noise}}$ of covariance matrix $\boldsymbol{R}_{\boldsymbol{\phi}}$,

$$E\{\sin \phi_i^{\mathrm{noise}}\} = 0,$$

$$E\{\cos \phi_i^{\mathrm{noise}}\} = \exp\left(-\frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}}}{2} \right),$$

$$E\{\sin \phi_i^{\mathrm{noise}} \sin \phi_j^{\mathrm{noise}}\} = \sinh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \exp\left(-\frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}}}{2} \right),$$

$$E\{\cos \phi_i^{\mathrm{noise}} \cos \phi_j^{\mathrm{noise}}\} = \cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \exp\left(-\frac{\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}}}{2} \right),$$

$$E\{\cos \phi_i^{\mathrm{noise}} \sin \phi_j^{\mathrm{noise}}\} = 0. \tag{E8}$$

By combining Eq. (E7), (E5), (E6), and (E8), we obtain

$$E\{\boldsymbol{y}^{\mathrm{n}}_{\mathrm{rad}_i}\} = \bar{a}_i[\mathrm{e}^{-\boldsymbol{R}_{\boldsymbol{\phi}_{ii}}/2} - 1],$$

$$E\{\boldsymbol{y}^{\mathrm{n}}_{\mathrm{tan}_i}\} = 0,$$

$$[\boldsymbol{R}_{\mathrm{rad,rad}}]_{ij} = [\bar{a}_i \bar{a}_j (\cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} - 1) + \boldsymbol{R}_{a_{ij}} \cosh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}}] \cdot \mathrm{e}^{-[(\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}})/2]},$$

$$[\boldsymbol{R}_{\mathrm{rad,tan}}]_{ij} = 0,$$

$$[\boldsymbol{R}_{\mathrm{tan,tan}}]_{ij} = (\bar{a}_i \bar{a}_j + \boldsymbol{R}_{a_{ij}}) \sinh \boldsymbol{R}_{\boldsymbol{\phi}_{ij}} \cdot \mathrm{e}^{-[(\boldsymbol{R}_{\boldsymbol{\phi}_{ii}} + \boldsymbol{R}_{\boldsymbol{\phi}_{jj}})/2]}. \tag{E9}$$

3. Scalar Case

Now we make the additional assumption that both $\boldsymbol{\phi}^{\mathrm{noise}}$ and $\boldsymbol{a}^{\mathrm{noise}}$ are decorrelated, i.e.,

$$\begin{cases} \boldsymbol{R}_{\boldsymbol{a}} = \mathrm{Diag}\{\sigma^2_{a,i}\}, \\ \boldsymbol{R}_{\boldsymbol{\phi}} = \mathrm{Diag}\{\sigma^2_{\phi,i}\}. \end{cases}$$

We obtain

$$\begin{cases} \boldsymbol{R}_{\mathrm{rad,rad}} = \mathrm{Diag}\{\sigma^2_{rad,i}\}, \\ \boldsymbol{R}_{\mathrm{tan,tan}} = \mathrm{Diag}\{\sigma^2_{tan,i}\}, \\ \boldsymbol{R}_{\mathrm{rad,tan}} = 0, \end{cases}$$

with

$$\sigma^2_{rad,i} = \frac{\bar{a}_i^2}{2}(1 - \mathrm{e}^{-\sigma^2_{\phi,i}})^2 + \frac{\sigma^2_{a,i}}{2}(1 + \mathrm{e}^{-2\sigma^2_{\phi,i}}),$$

$$\sigma^2_{tan,i} = \frac{\bar{a}_i^2}{2}(1 - \mathrm{e}^{-2\sigma^2_{\phi,i}}) + \frac{\sigma^2_{a,i}}{2}(1 - \mathrm{e}^{-2\sigma^2_{\phi,i}}). \tag{E10}$$

In this case, we can plot for one complex visibility the true noise distribution—i.e., a Gaussian noise in phase and modulus—and our Gaussian approximation (see Fig. 1).

ACKNOWLEDGMENTS

The authors want to express their special thanks to Eric Thiébaut for his support and for letting them use his minimization software. Serge Meimon is very grateful to Guy Perrin and Sylvestre Lacour, who allowed him to participate in two IOTA observing campaigns. We also thank all the people who contributed to the existence and success of the IOTA interferometer, in particular John Monnier, Wes Traub, Jean-Philippe Berger, and Marc Lacasse. Serge Meimon also thanks Vincent Bix Josso for his help with Appendix B. Serge Meimon and Laurent Mugnier acknowledge support from PHASE, the space- and ground-based high-angular-resolution partnership among ONERA, Observatoire de Paris, CNRS, and University Denis Diderot Paris 7.

REFERENCES

1. J. D. Monnier, "An introduction to closure phases," in *Principles of Long Baseline Stellar Interferometry*, P. R. Lawson, ed. (Jet Propulsion Laboratory, 1999), Chap. 13, pp. 203–239.
2. A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio-Astronomy* (Wiley Interscience, 1986).
3. A. Lannes, E. Anterrieu, and P. Maréchal, "Clean and wipe," Astron. Astrophys. Suppl. Ser. **123**, 183–198 (1997).
4. J. Hogbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," Astron. Astrophys. Suppl. Ser. **15**, 417–426 (1974).
5. T. J. Cornwell and P. N. Wilkinson, "A new method for making maps with unstable radio interferometers," Mon. Not. R. Astron. Soc. **196**, 1067–1086 (1981).
6. A. Lannes, "Weak-phase imaging in optical interferometry," J. Opt. Soc. Am. A **15**, 811–824 (1998).
7. J. W. Goodman, *Statistical Optics* (Wiley-Interscience, 1985).
8. D. L. Fried, "Statistics of a geometric representation of wavefront distortion," J. Opt. Soc. Am. **55**, 1427–1435 (1965).
9. A. Quirrenbach, "Phase referencing," in *Principles of Long Baseline Stellar Interferometry*, P. R. Lawson, ed. (Jet Propulsion Laboratory, 1999), Chap. 9, pp. 143–160.
10. R. C. Jennison, "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distribution of small angular extent," Mon. Not. R. Astron. Soc. **118**, 276–284 (1958).
11. T. A. Pauls, J. S. Young, W. D. Cotton, and J. D. Monnier, "A data exchange standard for optical (visible/IR) interferometry," Publ. Astron. Soc. Pac. **117**, 1255–1262 (2005).
12. G. Perrin, "The calibration of interferometric visibilities obtained with single-mode optical interferometers. Computation of error bars and correlations," Astron. Astrophys. **400**, 1173–1181 (2003).
13. C. Haniff, "Least-squares Fourier phase estimation from the modulo 2π bispectrum phase," J. Opt. Soc. Am. A **8**, 134–140 (1991).
14. G. Le Besnerais, S. Lacour, L. M. Mugnier, E. Thiébaut, G. Perrin, and S. Meimon, "Advanced imaging methods for long-baseline optical interferometry," IEEE J. Sel. Top. Signal Process. (to be published).
15. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024–2036 (1989).
16. R. Nityananda and R. Narayan, "Maximum entropy image reconstruction—a practical non-information-theoretic approach," J. Astrophys. Astron. **3**, 419–450 (1982).
17. E. Thiébaut, P. J. V. Garcia, and R. Foy, "Imaging with Amber/VLTI: the case of microjets," Astrophys. Space Sci. **286**, 171–176 (2003).
18. J. Navaza, "Accurate solutions of the maximum entropy equations. Their impact on the foundations of direct methods," in *Crystallographic Computing 5: From Chemistry to Biology*, Oxford Internaturial Union of Crystallography Book Series (Oxford University Press, 1991), pp. 317–323.
19. Y. Bar-Shalom and X.-R. Li, *Multitarget–Multisensor Tracking: Principles and Techniques* (YBS Publishing, 1995).
20. S. Meimon, L. M. Mugnier, and G. Le Besnerais, "Convex approximation of the likelihood in optical interferometry," J. Opt. Soc. Am. A **22**, 2348–2356 (2005).
21. L. M. Mugnier, G. Le Besnerais, and S. Meimon, "Inversion in optical imaging through atmospheric turbulence," in *Bayesian Approach to Inverse Problems*, J. Idier, ed., Digital Signal and Image Processing Series (ISTE/Wiley, 2008), Chap. 10, pp. 243–283.
22. A. Lannes, "Integer ambiguity resolution in phase closure imaging," J. Opt. Soc. Am. A **18**, 1046–1055 (2001).
23. P. R. Lawson, W. D. Cotton, C. A. Hummel, J. D. Monnier, M. Zhao, J. S. Young, H. Thorsteinsson, S. C. Meimon, L. Mugnier, G. Le Besnerais, E. Thiébaut, and P. G. Tuthill, "An interferometry imaging beauty contest," Proc. SPIE **5491**, 886–899 (2004).
24. J. A. Benson, C. A. Hummel, and D. Mozurkewich, "Simultaneous 6-station observations with the NPOI," Proc. SPIE **4838**, 358–368 (2003).
25. F. P. Schloerb, J.-P. Berger, N. P. Carleton, P. Hagenauer, P. Y. Kern, P. R. Labeye, M. G. Lacasse, F. Malbet, R. Millan-Gabet, J. D. Monnier, M. R. Pearlman, E. Pedretti, K. Rousselet-Perraut, S. D. Ragland, P. A. Schuller, W. A. Traub, and G. Wallace, "IOTA: recent science and technology," Proc. SPIE **6268**, 62680I (2006).
26. G. Perrin, S. Ridgway, B. Mennesson, W. Cotton, J. Woillez, T. Verhoelst, P. Schuller, V. Coudé du Foresto, W. Traub, R. Millan-Galbet, and M. Lacasse, "Unveiling Mira stars behind the molecules. Confirmation of the molecular layer model with narrow band near-infrared interferometry," Astron. Astrophys. **426**, 279–296 (2004).
27. S. Lacour, "Imagerie des étoiles évoluées par interférométrie. Réarrangement de pupille," Ph.D. thesis (University of Paris VI, 2007).

Annexe G

Article Mugnier et coll., 2008 [A15]

On-Line Long-Exposure Phase Diversity: a Powerful Tool for Sensing Quasi-Static Aberrations of Extreme Adaptive Optics Imaging Systems.

L. M. Mugnier, J.-F. Sauvage, T. Fusco, A. Cornia and S. Dandy

Office National d'Études et de Recherches Aérospatiales, Optics Department, BP 72, F-92322 Châtillon cedex, France

mugnier@onera.fr

Abstract: The phase diversity technique is a useful tool to measure and pre-compensate for quasi-static aberrations, in particular non-common path aberrations, in an adaptive optics corrected imaging system. In this paper, we propose and validate by simulations an extension of the phase diversity technique that uses long exposure adaptive optics corrected images for sensing quasi-static aberrations during the scientific observation, in particular for high-contrast imaging. The principle of the method is that, for a sufficiently long exposure time, the residual turbulence is averaged into a convolutive component of the image and that phase diversity estimates the sole static aberrations of interest.

The advantages of such a procedure, compared to the processing of short-exposure image pairs, are that the separation between static aberrations and turbulence-induced ones is performed by the long-exposure itself and not numerically, that only one image pair must be processed, that the estimation benefits from the high SNR of long-exposure images, and that only the static aberrations of interest are to be estimated. Long-exposure phase diversity can also be used as a phasing sensor for a segmented aperture telescope. Thus, it may be particularly useful for future planet finder projects such as EPICS on the European ELT.

OCIS codes: wavefront sensing (010.7350); adaptive optics (010.1080); phase retrieval (100.5070); atmospheric turbulence (010.1330); inverse problems (100.3190); telescopes (110.6770).

References and links

1. F. Roddier, ed., *Adaptive Optics in Astronomy* (Cambridge University Press, Cambridge, 1999).
2. A. Blanc, T. Fusco, M. Hartung, L. M. Mugnier, and G. Rousset, "Calibration of NAOS and CONICA static aberrations. Application of the phase diversity technique," Astron. Astrophys. **399**, 373–383 (2003).
3. M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset, and R. Lenzen, "Calibration of NAOS and CONICA static aberrations. Experimental results," Astron. Astrophys. **399**, 385–394 (2003).
4. J.-F. Sauvage, T. Fusco, G. Rousset, and C. Petit, "Calibration and Pre-Compensation of Non-Common Path Aberrations for eXtreme Adaptive Optics," J. Opt. Soc. Am. A **24**, 2334–2346 (2007).
5. G. Rousset, "Wave-front sensors," in Roddier [1], chap. 5, pp. 91–130.
6. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829–832 (1982).
7. R. G. Paxman and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," J. Opt. Soc. Am. A **5**, 914–923 (1988).

8. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072–1085 (1992).
9. M. G. Löfdahl and G. B. Scharmer, "Wavefront sensing and image restoration from focused and defocused solar images," Astron. Astrophys. **107**, 243–264 (1994).
10. D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing," Appl. Opt. **36**, 9186–9197 (1997).
11. B. J. Thelen, R. G. Paxman, D. A. Carrara, and J. H. Seldin, "Maximum a posteriori estimation of fixed aberrations, dynamic aberrations, and the object from phase-diverse speckle data," J. Opt. Soc. Am. A **16**, 1016–1025 (1999).
12. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035–1045 (2003).
13. L. M. Mugnier, A. Blanc, and J. Idier, "Phase Diversity: a Technique for Wave-Front Sensing and for Diffraction-Limited Imaging," in *Advances in Imaging and Electron Physics*, P. Hawkes, ed., vol. 141, chap. 1, pp. 1–76 (Elsevier, 2006).
14. D. S. Acton, D. Soltau, and W. Schmidt, "Full-field wavefront measurements with phase diversity," Astron. Astrophys. **309**, 661–672 (1996).
15. N. Baba and K. Mutoh, "Measurement of telescope aberrations through atmospheric turbulence by use of phase diversity," Appl. Opt. **40**, 544–552 (2001).
16. D. J. Lee, B. M. Welsh, and M. C. Roggemann, "Diagnosing unknown aberrations in an adaptive optics system by use of phase diversity," Opt. Lett. **22**, 952–954 (1997).
17. F. Roddier, "The effects of atmospherical turbulence in optical astronomy," in *Progress in Optics*, E. Wolf, ed., vol. XIX, pp. 281–376 (North Holland, Amsterdam, 1981).
18. G. Demoment, "Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems," IEEE Trans. Acoust. Speech Signal Process. **37**, 2024–2036 (1989).
19. A. Blanc, "Identification de réponse impulsionnelle et restauration d'images : apports de la diversité de phase," Ph.D. thesis, Université Paris XI Orsay (2002).
20. J. Idier, L. Mugnier, and A. Blanc, "Statistical behavior of joint least square estimation in the phase diversity context," IEEE Trans. Image Processing **14**, 2107–2116 (2005).
21. L. Meynadier, V. Michau, M.-T. Velluet, J.-M. Conan, L. M. Mugnier, and G. Rousset, "Noise propagation in wave-front sensing with phase diversity," Appl. Opt. **38**, 4967–4979 (1999).
22. T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper, and P. Puget, "High order Adaptive Optics requirements for direct detection of Extra-solar planets. Application to the SPHERE instrument." Opt. Express **14**, 7515–7534 (2006).
23. J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, J. Antici, P. Baudoz, A. Boccaletti, M. Carbillet, J. Charton, R. Claudi, T. Fusco, R. Gratton, T. Henning, N. Hubin, F. Joos, M. Kasper, M. Langlois, C. Moutou, J. Pragt, P. Rabou, M. Saisse, H. M. Schmid, M. Turatto, S. Udry, F. Vakili, R. Waters, and F. Wildi, "SPHERE: A Planet Finder Instrument for the VLT," in *Proceedings of the conference In the Spirit of Bernard Lyot: The Direct Detection of Planets and Circumstellar Disks in the 21st Century.*, P. Kalas, ed. (University of California, Berkeley, CA, USA, 2007).
24. L. Jolissaint, J.-P. Véran, and R. Conan, "Analytical modeling of adaptive optics: foundations of the phase spatial power spectrum approach," J. Opt. Soc. Am. A **23**, 382–394 (2006).
25. R. Conan, T. Fusco, G. Rousset, D. Mouillet, J.-L. Beuzit, M. Nicolle, and C. Petit, "Modeling and analysis of XAO systems. Application to VLT-Planet Finder," in *Advancements in Adaptive Optics*, vol. 5490 (Proc. Soc. Photo-Opt. Instrum. Eng., 2004). Conference date: Jun. 2004, Glasgow, UK.
26. M. Kasper, C. Verinaud, J.-L. Beuzit, N. Yaitskova, N. Hubin, A. Boccaletti, K. Dohlen, T. Fusco, A. Glindemann, R. Gratton, and N. Thatte, "EPICS: A Planet Hunter for the European ELT," in *Proceedings of the conference In the Spirit of Bernard Lyot: The Direct Detection of Planets and Circumstellar Disks in the 21st Century.*, P. Kalas, ed. (University of California, Berkeley, CA, USA, 2007). Conference date: Jun. 2007,.
27. J.-F. Sauvage, L. Mugnier, T. Fusco, and G. Rousset, "Post processing of differential images for direct extrasolar planet detection from the ground," in *Advances in Adaptive Optics II*, L. Ellerbroek, B. and D. Bonaccini Calia, eds., vol. 6272 (Proc. Soc. Photo-Opt. Instrum. Eng., 2006).
28. J.-F. Sauvage, "Calibration et méthodes d'inversion en imagerie haute dynamique pour la détection directe d'exoplanètes." Ph.D. thesis, Université Paris VII (2007).

1. Introduction

Calibrating the quasi-static aberrations of a ground-based Adaptive Optics (AO) [1] corrected imaging system is an important issue, especially for extreme AO high contrast instruments such as the proposed planet finder instruments for the ESO and Gemini 8-meter telescopes.

Measuring these aberrations off-line, *i.e.*, during a day-time calibration on an internal refer-

ence source, has been successfully applied to existing systems such as NAOS [2, 3] or Keck, and recently refined for the SPHERE project in order to achieve a nanometric accuracy [4]. The main limitations of such a procedure directly stem from its off-line nature: the aberrations located before the internal reference source are not sensed, and the aberrations may evolve between the day-time calibration and the night-time observation. These two limitations can be circumvented by an appropriate calibration performed on-line, *i.e.*, during night-time observations, as proposed in the following.

Two problems must be addressed in order to calibrate on-line the quasi-static aberrations of the optical system made of the telescope, its AO system and the camera: the first one is to distinguish between the turbulence-induced component of the wavefront and the static one, which is the only wavefront of interest for the problem at hand, and the second one is to sense all the aberrations from the (potentially segmented) primary down to the focal plane of the camera.

A wave-front sensor (WFS) is able to measure the aberrations seen by the telescope on-line, but these consist of the sum of a turbulence-induced component, which is partially compensated for by the AO, and a static component. Because the turbulence, whether corrected or not, evolves quickly, the WFS measurements are generally performed with integration times that freeze the turbulence evolution, typically a few milli-seconds.

There is today a large number of WFSs, which are thoroughly reviewed in Ref. [5] and can be classified into two families: pupil-plane sensors, such as the Hartmann-Shack and the curvature sensors, and focal-plane sensors. A focal-plane WFS is the only way to be sensitive to *all* aberrations down to the focal plane, and in particular to the so-called non-common path aberrations of an AO system, which motivates our choice for a focal-plane WFS in the following. Estimating aberrations from a single focal-plane image of a point source is a difficult problem known as phase retrieval. Phase-retrieval has two major limitations. Firstly, it only works with a point source. Secondly, there is generally a sign ambiguity in the recovered phase, i.e., the solution is not unique. Gonsalves [6] showed that by using a second image with an additional known phase variation with respect to the first image such as defocus, it is possible to estimate the unknown phase even when the object is extended and unknown. The presence of this second image additionally removes the above-mentioned sign ambiguity of the solution. This technique, referred to as phase diversity, has been significantly developed in the past twenty years, both for wavefront sensing and for imaging; see for instance Refs. [7, 8, 9, 10, 11, 12, 3], and Ref. [13] for a review.

The estimation of static aberrations from a series of short-exposure phase-diversity data has been performed using a series of image pairs of an astronomical object [14, 15]. Lee *et al.* [16] have performed such a calibration of static aberrations with a series of images instead of image pairs, and an original diversity: no additional defocused image was used, and successive changes to the adaptive optics introduced the required diversity.

In both approaches, the static aberrations are obtained as an empirical average of the phase estimates corresponding to each short-exposure data. This is notably suboptimal for at least three reasons:

- the images correspond to short integration times, and are consequently noisier than the corresponding long-exposure image pair, so each phase estimate suffers from this noise;
- the computational cost is high because many short-exposure images must be processed in order to estimate the sought static aberrations;
- the phase estimation accuracy may be penalized because the estimation must be performed on a number of phase parameters that is large enough to describe the short-exposure phase, whereas only a smaller number of these parameters may be of interest,

if the sought static aberrations are of lower order than turbulence-induced ones;

In this paper, we propose and validate an extension of the phase diversity technique that uses long-exposure AO-corrected images for sensing quasi-static aberrations. This way, (1) the separation between quasi-static aberrations and turbulence-induced ones is performed by the long-exposure itself and not numerically, (2) only one image pair must be processed, (3) the estimation benefits from the high SNR of long-exposure images, and (4) only the static aberrations of interest are to be estimated.

2. Principle of long-exposure phase diversity

We consider a ground-based telescope observing Space through the turbulent atmosphere. The long-exposure optical transfer function (OTF) of the atmosphere+instrument system is the product of the OTF of the sole instrument $\tilde{h}_\mathrm{s}$, called static OTF in the following, by the atmospheric transfer function (ATF) $\tilde{h}_\mathrm{a}$ [17]:

$$\langle \tilde{h}_\mathrm{o} \rangle = \tilde{h}_\mathrm{s}\, \tilde{h}_\mathrm{a}. \tag{1}$$

The static OTF is a function of the unknown static aberrations, which are coded in the phase function φ in the aperture; let P be the indicator function of the aperture, *i.e.*, 1 in the aperture and 0 outside, $\tilde{h}_\mathrm{s}$ is given by:

$$\tilde{h}_\mathrm{s}(\varphi) = Pe^{i\varphi} \otimes Pe^{i\varphi} \tag{2}$$

where $\otimes$ denotes auto-correlation. The phase function $\varphi(u,v)$ is expanded on a basis $\{b_k\}$, which is typically either Zernike polynomials or the pixel indicator functions in the aperture: $\varphi(u,v) = \sum_k \phi_k\, b_k(u,v)$, where the summation is, in practice, limited to the number of coefficients considered sufficient to correctly describe the static aberrations to be estimated. We shall denote by ϕ the vector concatenating the set of unknown aberration coefficients ϕ_k.

Assuming phase perturbations with Gaussian statistics, the ATF at any spatial frequency $\mathbf{f}$ is given by [17]:

$$\tilde{h}_\mathrm{a}(\mathbf{f}) = e^{-\frac{1}{2} D_\phi(\lambda \mathbf{f})} \tag{3}$$

where λ is the imaging wavelength and D_ϕ is the phase structure function. If the turbulence is partially compensated by an AO system, Equations (1) and (3) remain valid [1], although slightly approximate because the residual phase after AO correction is not stationary.

With these equations in hand, and assuming that the image is not larger than the isoplanatic patch, we can now model the long-exposure image. The image is recorded by a detector such as a CCD camera, which integrates the flux on a grid of pixels. This can be conveniently modeled as the convolution by a detector PSF h_d, assumed to be known in the sequel, followed by a sampling operation. Using Eq. (1), the global long-exposure PSF of the instrument is thus:

$$h_\mathrm{le} = h_\mathrm{d} \star \langle h_\mathrm{o} \rangle = h_\mathrm{d} \star h_\mathrm{s} \star h_\mathrm{a}, \tag{4}$$

where h_s is the PSF due to static aberrations, given by the inverse Fourier transform of Eq (2), h_a is the atmospheric PSF given by the inverse Fourier transform of Eq. (3), and $\star$ denotes convolution.

Due to the inevitable sampling and noise of the detection processes, the image $\boldsymbol{i}_\mathrm{f}$ recorded in the focal plane is the noisy sampled convolution of the long-exposure point-spread function (PSF) h_le with the observed object o. This model is generally approximated by a noisy discrete convolution with the sampled version $\boldsymbol{o}$ of the object o:

$$\boldsymbol{i}_{\mathrm{f}} = \boldsymbol{h}_{\mathrm{le}} \star \boldsymbol{o} + \boldsymbol{n}, \tag{5}$$

where $\boldsymbol{h}_{\mathrm{le}}$ is the sampled version of h_{le}, and $\boldsymbol{n}$ is a corruptive noise process. If the noise is not additive and independent from the noiseless image, for instance if it is predominantly photon noise, then Eq. (5) should read $\boldsymbol{i} = \boldsymbol{h}_{\mathrm{le}} \star \boldsymbol{o} \diamond \boldsymbol{n}$, with the symbol $\diamond$ representing a pixel-by-pixel operation [18]. For legibility we shall keep the additive notation.

Combining Eq. (5) with the discrete counterpart of Eq. (4) yields the following discrete image model for the focused and defocused images respectively:

$$\boldsymbol{i}_{\mathrm{f}} = \boldsymbol{h}_{\mathrm{d}} \star \boldsymbol{h}_{\mathrm{s}}^{(\phi)} \star (\boldsymbol{h}_{\mathrm{a}} \star \boldsymbol{o}) + \boldsymbol{n} \tag{6}$$

$$\boldsymbol{i}_{\mathrm{d}} = \boldsymbol{h}_{\mathrm{d}} \star \boldsymbol{h}_{\mathrm{s}}^{(\phi+\phi_d)} \star (\boldsymbol{h}_{\mathrm{a}} \star \boldsymbol{o}) + \boldsymbol{n}', \tag{7}$$

where $\boldsymbol{h}_{\mathrm{d}}$, $\boldsymbol{h}_{\mathrm{s}}$ and $\boldsymbol{h}_{\mathrm{a}}$ are the sampled versions of h_{d}, h_{s} and h_{a}, ϕ_d is the known additional phase introduced in image $\boldsymbol{i}_{\mathrm{d}}$, and the superscripts on $\boldsymbol{h}_{\mathrm{s}}$ are reminders of the aberrations that enter the static PSF of each image.

Let $\boldsymbol{o}'$ be the convolution of the atmospheric PSF with the observed object:

$$\boldsymbol{o}' = \boldsymbol{h}_{\mathrm{a}} \star \boldsymbol{o} \tag{8}$$

The phase-diversity data model of Eqs. (6) and (7) is strictly identical to the one that would be obtained by imaging the pseudo-object $\boldsymbol{o}'$ in the absence of turbulence and with the same static aberrations. Thus, all the phase estimation methods developed for short-exposure images, in which the OTF of the system is completely described by a phase function, can be applied here to estimate the sole static aberrations.

Additionally, it is well-known that, for a given noise level, the estimation quality of the aberrations in phase diversity depends on the spectral content of the observed scene. Thus, in the method proposed here, the estimation quality of the aberrations will depend both on the spectral content of the observed object and on the ATF, *i.e.*, on the turbulence correction quality provided by the AO.

The appropriate implementation of this long-exposure phase diversity technique depends on the type of instrument. For non coronagraphic instruments, one should use images provided by the science sensor; the defocused image can be either obtained simultaneously with the focused science image by means, *e.g.*, of a beamsplitter, or alternately. In the latter case, the deformable mirror itself can be used to provide the defocus [2, 3]. In both cases the fraction of the incoming flux allotted to the defocused image may be notably less than 50%, in order to maximize the flux on the scientific data. Indeed, if the quasi-static aberrations are measured at intervals of, *e.g.*, a half-hour, defocused images must only be available with that kind of rate.

For coronagraphic systems, for which the aberrations to be minimized are the ones located before the coronagraph, one may use a beam-splitter and an auxiliary image sensor located just before the coronagraph. For the SPHERE instrument such a sensor actually already exists in the design for centering the star image on the coronagraph: it is the so-called differential tip-tilt sensor. This sensor could easily be adapted and used for the long-exposure phase diversity measurements. As in the non coronagraphic case, the defocused image may be obtained simultaneously with the focused image by use of a beamsplitter, or alternately by means of a longitudinal displacement of the sensor by a few millimeters.

3. Chosen phase estimation method

Among all the possible estimation methods (see, *e.g.*, Ref. [13] for a review) in this paper we choose, for simplicity, the conventional least-squares joint estimation of the phase ϕ and the object $\boldsymbol{o}'$, with a regularization on both quantities: $(\hat{\boldsymbol{o}}', \hat{\phi}) = \arg\min J(\boldsymbol{o}', \phi)$ with

$$J(\boldsymbol{o}',\phi)=\frac{1}{2\sigma_n^2}\|\boldsymbol{i}_{\mathrm{f}}-\boldsymbol{h}_{\mathrm{d}}\star\boldsymbol{h}_{\mathrm{s}}^{(\phi)}\star\boldsymbol{o}'\|+\frac{1}{2\sigma_{n'}^2}\|\boldsymbol{i}_{\mathrm{d}}-\boldsymbol{h}_{\mathrm{d}}\star\boldsymbol{h}_{\mathrm{s}}^{(\phi+\phi_d)}\star\boldsymbol{o}'\|+R_o(\boldsymbol{o}')+R_\phi(\phi), \qquad (9)$$

where σ_n^2 and $\sigma_{n'}^2$ are the noise variances of the two images, estimated beforehand. The object regularization is chosen as a quadratic function, so that the whole criterion J is quadratic with respect to $\boldsymbol{o}'$ and thus has an unique, closed-form solution $\hat{\boldsymbol{o}}'(\phi)$ for a given phase ϕ. This allows one to replace the optimization of $J(\boldsymbol{o}',\phi)$ with that of criterion $J'(\phi)\triangleq J(\hat{\boldsymbol{o}}'(\phi),\phi)$, as commonly done in the unregularized case [6, 8].

Following the findings of Blanc [19, 12], we under-regularize the object in order to best estimate the phase. This strategy is supported by the fact that it yields a phase estimation with satisfactory asymptotic properties, as shown in Ref. [20], and that these properties hold even if the noise is not Gaussian.

Concerning the phase, we choose the basis of the pixel indicator functions rather than, *e.g.*, a truncated basis of Zernike polynomials, in order to model and reconstruct phases with a high spatial frequency content. Because of the potentially large number of phase unknowns we are lead to regularize the phase estimation. To this aim, we use a functional proposed specifically for such a phase basis in Refs. [19, chap. 7] and [13, Sect. 8], which is recalled below:

$$R_\phi(\phi)=\sum_{(l,m)\in S}\left[\left|e^{j(\phi_{l-1,m}-\phi_{l,m})}-e^{j(\phi_{l,m}-\phi_{l+1,m})}\right|^2+\left|e^{j(\phi_{l,m-1}-\phi_{l,m})}-e^{j(\phi_{l,m}-\phi_{l,m+1})}\right|^2\right],$$

where the summation is done on all the pixels within the pupil (S is the pupil support). Furthermore, we impose a strict support constraint *i.e.*, all terms $|...|^2$ that contain, at least, a pixel out of the pupil support, are suppressed.

This regularization function has been constructed in such a way that it is insensitive, as the data is, to a global piston, to tip-tilt, and to any 2π variation of the phase on any pixel. This way, no local minimum is introduced by the regularization into the minimized criterion.

4. Validation by simulations

We shall now validate the proposed method by simulations. We shall essentially study the influence of the exposure time. Indeed, the main specific assumption of the proposed method lies in Eq. (1), because the factorization of the OTF in a static OTF and an ATF is strictly valid only if the turbulence is perfectly averaged *i.e.*, for an infinite exposure time. Note that by exposure time we mean the (finite) number of independent turbulence realizations, not the noise level. In all the simulations presented here, we have considered noiseless images. We have checked that the behavior of the phase estimation in the proposed method with respect to the noise level is not specific and is the same as conventional phase diversity with short-exposure images: the estimation error is usually proportional to the average standard deviation of the noise in the images [21, 12]. Then we shall briefly study the influence of the AO correction quality.

4.1. Conditions of simulation

We consider here a point-like source, observed with an 8 m ground-based telescope, equipped with AO. The simulations take into account both the AO-corrected turbulence and static aberrations. The baseline adaptive optics system considered is the high-performance AO system SAXO [22] of the SPHERE [23] instrument. We use a Fourier-based simulation method that describes the AO via the spatial power spectrum of the residual phase [24] and is presented in Ref. [25]. The simulation takes the following realistic set of parameters: a 41×41 subaperture

Shack-Hartmann, a 1.2 kHz sampling frequency, a guide star of magnitude 8, and a Paranal-like turbulence profile, with a seeing of 0.8 arcsec at 0.5 µm. Static aberrations are randomly generated according to a f^{-2} spectrum, f being the spatial frequency in the pupil, with a total wavefront error of 0.23 radian RMS at 1.6 µm, *i.e.*, 60 nm RMS.

For some of the simulations, the number of actuators on a pupil diameter will be decreased from 40×40 to 14×14 and 7×7 in order to study the influence of the AO correction quality. The above simulation conditions lead to a phase variance of the residual turbulence which is respectively 1, 7 and 21 time(s) the variance of the static aberrations.

Two simultaneous long-exposure images are simulated, with a phase diversity between these images consisting of a 1.814 radian RMS defocus. The corresponding defocus distance is proportional to the square of the f-number of the system and, at 1.6 µm, is 2.9 mm for an $f/15$ system such as NAOS and 2 cm for an $f/40$ system such as SPHERE. These images are simulated following two schemes:

- Finite exposure time images are made of the summation of N short exposures. The long-exposure OTF is then the sum of N short-exposure OTFs, each of which being computed through Eq. (2), with a phase φ composed of the sum of the static aberrations and of the instantaneous AO-corrected turbulent wavefront. The turbulent wavefront is randomly generated from a PSD that takes into account both the turbulence profile and the AO correction [25].

- Infinite exposure time images are not computed as an empirical average. Instead, the residual phase structure function D_φ is computed from the PSD of the AO-corrected turbulent wavefront, then the ATF $\tilde{h}_\mathrm{a}$ is computed via Eq. (3), and the images $\boldsymbol{i}_\mathrm{f}$ and $\boldsymbol{i}_\mathrm{d}$ are computed according to Equations (6) and (7).

4.2. Influence of exposure time

We first consider the case of a high performance, SAXO-like correction. Figure 1 shows the evolution of the reconstruction error with the number of exposures used for the simulation. The first points (from 10 to 1000 exposures) are simulated with a finite exposure time, whereas the last point (noted infinity on the X-axis) is simulated with an infinite exposure time. Because the correlation time of corrected turbulence is typically 10 to 100 milliseconds depending on turbulence parameters and on the AO correction quality, the simulations with 1000 exposures correspond to an integration time between 10 and 100 seconds.

The reconstruction error decreases with the number of exposures, down to very weak values (less than 0.01 radian) for an infinite exposure time. As the number of exposures increases, Eq. (1) becomes more valid, turbulence residuals become better fitted by a modification of the estimated object only (into an object $\boldsymbol{o}' = \boldsymbol{h}_\mathrm{a} \star \boldsymbol{o}$, *cf.* Eq. 8), and the estimated phase is eventually only the set of static aberrations. Incidentally, we have checked that for a single turbulent exposure, the phase estimated is the sum of the turbulent wavefront and the static aberrations.

This simulation study shows that the estimation of static aberrations from a single pair of turbulence-degraded images is possible. The quality of aberration reconstruction is directly linked to the convergence of the images towards long-time exposures. For the AO system and the level of static aberrations considered here, an integration time corresponding to a thousand independent turbulence realizations yields a phase estimation error of about 0.012 radian, close to that obtained with an infinite exposure time. At 1.6 µm, this number translates into a 3 nm RMS optical path difference. This precision is compatible with the $5-10$ nm RMS static aberration residual that is needed for the detection of warm Jupiters on an 8-meter telescope. Incidentally, we see that with less than 1000 exposures the required precision would not be

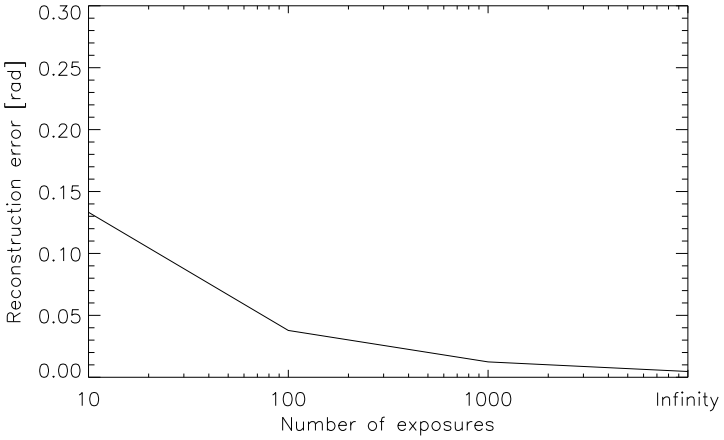


Fig. 1. Evolution of the reconstruction error with the exposure time (in number of independent turbulence realizations). Static aberrations are randomly generated according to conditions described in Subsection 4.1. AO correction is assumed to be performed by a SAXO-like system.

obtained for such a mission.

4.3. Influence of AO correction quality

Figure 2 shows the evolution of the estimation error, for different correction qualities, obtained here simply by varying the number of actuators of the AO system, all other parameters being equal. One can see that as the correction quality degrades, the turbulence residuals are more important and thus the estimation of the static aberrations needs more exposures for the same error level. The estimation errors for an infinite exposure time are equivalent for all three correction qualities.

In the case of a NAOS correction with 14×14 actuators and 1000 exposures, the phase estimation error is about 0.13 radian, which at 2.2 µm translates into a 45 nm RMS optical path difference. This is almost three times smaller than the residual static aberrations of 120 nm RMS measured on NAOS-CONICA *after* off-line phase diversity measurement and correction [3]. The on-line long-exposure phase diversity technique could thus be an attractive way to calibrate quasi-static aberrations on non-extreme AO systems too.

We now detail the spectral analysis of the estimated aberrations. In an AO system, the number of actuators determines the highest spatial frequency of the turbulence to be corrected. This parameter therefore directly impacts on the spectral content of the turbulence residuals. On the following figures we plot the circularly averaged spectra of the estimated aberations with respect to the spatial frequency, for different AO correction levels.

Figure 3 shows the error spatial spectrum in the case of a NAOS correction (7×7 actuators), for different exposure time. One can see that the spectrum of estimated aberrations get closer to that of the true ones as the number of exposures increases, and that the convergence is slower for the uncorrected frequencies of the turbulence (above 3.5 cycles/pupil for this 7×7 actuator system).

In the case of a higher correction (respectively 14×14 and 40×40 actuators, on Figures 4 and 5), the turbulence residuals are reduced and the convergence of the spectrum to true

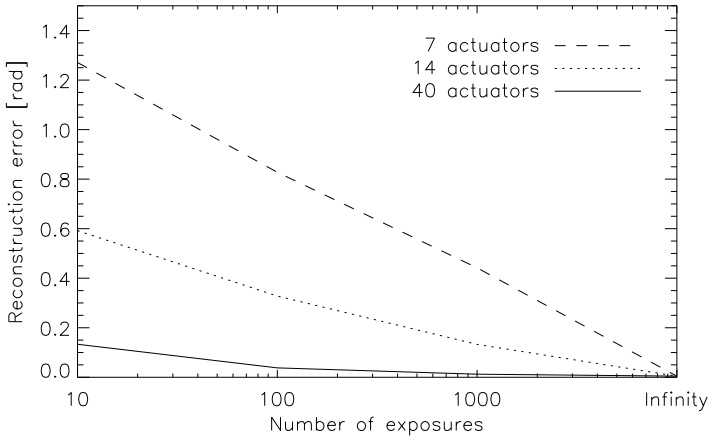


Fig. 2. Evolution of the reconstruction error with the exposure time, for several levels of AO correction. Static aberrations are randomly generated according to conditions described in Subsection 4.1.

spectrum is consequently faster. Moreover, as in the NAOS-7 case, convergence is notably faster in the corrected part of the aberration spectrum, which correspond to a limit of 7 and 20 cycles/pupil respectively.

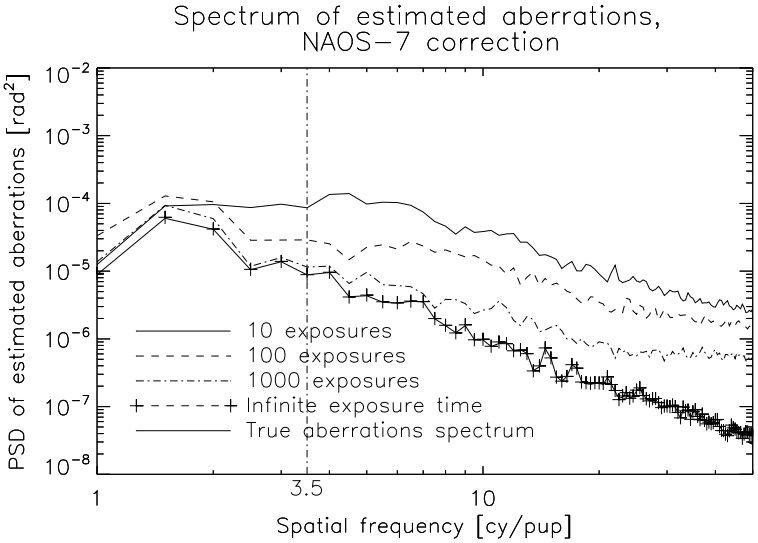


Fig. 3. Spatial spectrum of the estimated aberrations, in the case of a NAOS-7 correction. The vertical line represents the maximum spatial frequency that is corrected by the AO system.

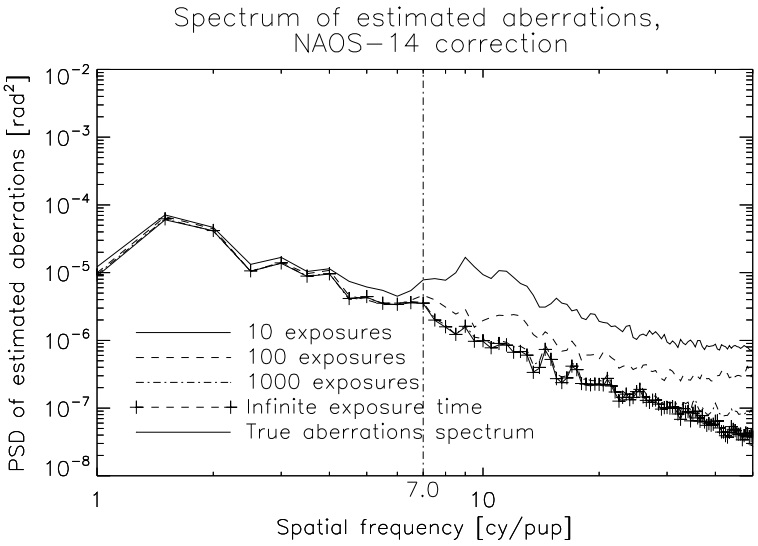


Fig. 4. Spatial spectrum of the estimated aberrations, in the case of a NAOS-14 correction. The vertical line represents the maximum spatial frequency that is corrected by the AO system.

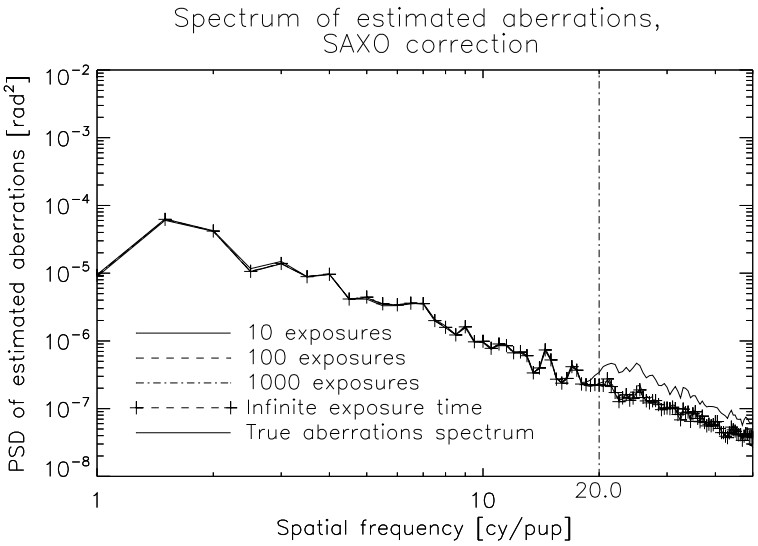


Fig. 5. Spatial spectrum of the estimated aberrations, in the case of a SAXO/SPHERE correction. The vertical line represents the maximum spatial frequency that is corrected by the AO system.

5. Conclusion and perspectives

The phase diversity technique is a powerful tool to measure and pre-compensate for quasi-static aberrations, in particular non-common path aberrations, in an AO-corrected imaging system.

So far it has, to the best of our knowledge, either been used off-line on an internal calibration

source, or on-line from short-exposure turbulence-degraded data. Each approach has its own limitations, discussed in this paper. We have proposed and validated by simulations an extension of the phase diversity technique that uses long exposure AO corrected images for sensing quasi-static aberrations.

The principle of the method is that, for a sufficiently long exposure time, the residual turbulence is averaged into a convolutive component of the image and that phase diversity estimates the sole static aberrations of interest.

Technically, the advantages of such a procedure, compared to the processing of short-exposure image pairs, are that the separation between static aberrations and turbulence-induced ones is performed by the long-exposure itself and not numerically, that only one image pair must be processed, that the estimation benefits from the high SNR of long-exposure images, and that only the static aberrations of interest are to be estimated. Compared to pupil-plane wavefront sensing techniques, phase diversity has the advantages that it senses all aberrations down to the focal plane and that the hardware is extremely simple.

From a system point of view, on-line long-exposure phase diversity opens a new area of applications, in particular it will allow one to correct in real time (meaning during the scientific exposure) for any evolution of instrumental defects. This may be considered for improvements to the SPHERE instrument and should significantly improve the overall system detectivity.

This technique can also be used as a phasing sensor for a segmented aperture telescope. Indeed, phase diversity can be applied to the peculiar aberrations constituted by the differential tip-tilts and pistons of such a telescope.

Thus, long-exposure phase diversity may be particularly useful for the future Planet Finder project on the E-ELT called EPICS [26]. Indeed, on the one hand, for this project, the detectivity requirements are by far more stringent than for SPHERE and the on-line correction of non-common path aberrations is mandatory. And on the other hand, without any opto-mechanical modification of the sensor, long-exposure phase diversity should enable very accurate measurements of the segments' phasing.

Lastly, for some applications it may be useful to estimate not only the static aberrations but also the ATF, which is part of the estimated quantities. The method proposed herein can be refined to estimate the ATF precisely using (a) a parametrization of the ATF through the phase structure function and (b) a specific space-varying regularization criterion for the ATF along the lines of [27, 28].

Short-term perspectives include a more thorough study of the performance of the long-exposure phase diversity technique, coupled with a global system analysis.

Annexe H

Article Chenegros et coll., 2007 [A18]

Chenegros *et al.* Vol. 24, No. 5/May 2007/J. Opt. Soc. Am. A 1349

3D phase diversity: a myopic deconvolution method for short-exposure images: application to retinal imaging

Guillaume Chenegros and Laurent M. Mugnier

Department of Optics, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon Cedex, France

François Lacombe

Mauna Kea Technologies, 9 rue d'Enghien, 75010 Paris, France

Marie Glanc

Laboratoire d'Études Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris-Meudon, 5 place Jules Janssen, 92195 Meudon Cedex, France

Received July 31, 2006; revised October 2, 2006; accepted October 3, 2006; posted October 10, 2006 (Doc. ID 73587); published April 11, 2007

3D deconvolution is an established technique in microscopy that may be useful for low-cost high-resolution imaging of the retina. We report on a myopic 3D deconvolution method developed in a Bayesian framework. This method uses a 3D imaging model, a noise model that accounts for both photon and detector noises, a regularization term that is appropriate for objects that are a mix of sharp edges and smooth areas, a positivity constraint, and a smart parameterization of the point-spread function (PSF) by the pupil phase. It estimates the object and the PSF jointly. The PSF parameterization through the pupil phase constrains the inversion by dramatically reducing the number of unknowns. The joint deconvolution is further constrained by an additional longitudinal support constraint derived from a 3D interpretation of the phase-diversity technique. This method is validated by simulated retinal images.

OCIS codes: 100.1830, 100.3020, 100.5070, 100.6890, 170.6900, 010.1080.

1. INTRODUCTION

Early detection of pathologies of the human retina calls for an *in vivo* exploration of the retina at the cell scale. Direct observation from the outside suffers from the poor optical quality of the eye. The time-varying aberrations of the eye can be compensated *a posteriori* if measured simultaneously with the image acquisition; this technique is known as deconvolution from wavefront sensing[1,2] and has been successfully applied to the human retina.[3] These aberrations can also be compensated for in real time by use of adaptive optics[4] (AO). Yet, the correction is always partial.[5–7] Additionally, the object under examination (the retina) is three dimensional (3D), and each recorded image contains contributions from the whole object's volume.

In two-dimensional (2D) deconvolution, each image is deconvolved separately; i.e., only one object plane is assumed to contribute to each image. This is an appropriate image model in astronomy, for instance, but is a somewhat crude approximation in microscopy, as it does not properly account for the halo in each image that comes from the parts of the observed object that are out of focus.

Three-dimensional deconvolution is an established technique in microscopy and, in particular, in conventional fluorescence microscopy.[8] The combination of a conventional microscope with deconvolution is often referred to as deconvolution microscopy or even "digital confocal," because the use of 3D deconvolution can notably improve the resolution of the recorded conventional images, especially in the longitudinal dimension, while remaining simpler and cheaper than a confocal microscope.

Yet, to the best of our knowledge, deconvolution of retinal images has so far been performed with 2D deconvolution techniques, both in deconvolution from wavefront sensing[3] and in deconvolution of AO-corrected images.[9]

Besides, because deconvolution is an ill-posed inverse problem,[10–12] most modern deconvolution methods use regularization in order to avoid an uncontrolled amplification of the noise. The regularization that is commonly used in 3D deconvolution is the classical Tikhonov regularization, which is quadratic (see Subsection 2.B) and thus tends to oversmooth edges. In Section 3 we present a regularized edge-preserving 3D deconvolution method.

Furthermore, a deconvolution method needs a precise estimate of the point-spread function (PSF), which is not always available. This is particularly true for 3D imaging. We thus propose a myopic deconvolution method that estimates simultaneously the PSF and the object in Section 3.

To better constrain the problem, we propose the use of

1084-7529/07/051349-9/$15.00

an additional constraint in Section 5. The efficiency of this proposed constraint is shown on realistic simulated retinal images.

2. 3D DECONVOLUTION METHOD WITH KNOWN PSF

A. Imaging Model

The image formation is modeled as a 3D convolution:

$$\mathbf{i} = \mathbf{h} * \mathbf{o} + \mathbf{n},$$

where $\mathbf{i}$ is the (3D) pile of (2D) recorded images, $\mathbf{o}$ is the 3D unknown observed object that concatenates each object slice (which slices are regularly spaced out of δz), $\mathbf{h}$ is the 3D PSF, $\mathbf{n}$ is the noise, and * denotes the 3D convolution operator. For a system with N images of N object planes, this 3D convolution can be rewritten as

$$\mathbf{i}_k = \left(\sum_{l=0}^{N-1} \mathbf{h}_{k-1} * \mathbf{o}_l \right) + \mathbf{n}_k, \tag{1}$$

where $\mathbf{o}_j$ is the object in plane j, $\mathbf{i}_k$ is the kth recorded image, and $\mathbf{h}_{k-j}$ is the 2D PSF corresponding to a defocus of $(k-j)\delta z$. The PSF is that of the system composed of the eye, the imaging system (including the AO); and the detector. We assume that the whole recording process is fast enough that the different 2D PSFs differ only by a defocus (see Section 3). Figure 1 illustrates the imaging process in the case of three object and image planes.

B. 3D Deconvolution Method

Most deconvolution techniques boil down to the minimization (or maximization) of a criterion. An important task is the definition of a suitable criterion for the given inverse problem.

Following the Bayesian[12] maximum *a posteriori* (MAP) approach, the deconvolution problem can be stated as follows: we look for the most likely object $\hat{\mathbf{o}}$, given the observed image $\mathbf{i}$ and our prior information on $\mathbf{o}$, which is summarized by a probability density $p(\mathbf{o})$. This reads as

$$\hat{\mathbf{o}} = \arg\max_{\mathbf{o}} p(\mathbf{o}|\mathbf{i}) = \arg\max_{\mathbf{o}} p(\mathbf{i}|\mathbf{o}) \times p(\mathbf{o}).$$

Equivalently, $\hat{\mathbf{o}}$ can be defined as the object that minimizes a compound criterion $J(\mathbf{o})$ defined as follows:

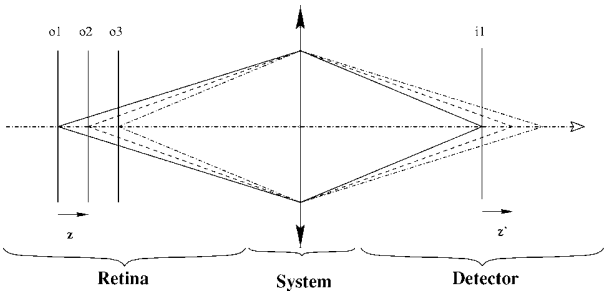


Fig. 1. Illustration of the 3D image formation for three object planes. The object is on the left, and the image is on the right. The system is composed of the eye and the optical system (including the AO). In image i_1, object o_1 is focused; o_2 and o_3 are defocused. Images i_2 and i_3 are not represented here.

$$J(\mathbf{o}) = J_i(\mathbf{o}) + J_o(\mathbf{o}),$$

where the negative log likelihood $J_i = -\ln p(\mathbf{i}|\mathbf{o})$ is a measure of fidelity to the data and $J_o = -\ln p(\mathbf{o})$ is a regularization or penalty term, so the MAP solution can equivalently be called a penalized-likelihood solution.

The noise is a mixture of nonstationary, Poisson-distributed photon noise and detector noise, which can be reasonably modeled as nonstationary white Gaussian as soon as the flux level is a few tens of photoelectrons per pixel.[13] If the noise statistics are additive, nonstationary white Gaussian, then the data fidelity term is a simple weighted least-squares difference between the actual data $\mathbf{i}$ and our model of the data for a given object, $\mathbf{h}*\mathbf{o}$:

$$J_i(\mathbf{o}) = \frac{1}{2} \sum_{k=0}^{N-1} \sum_{p,q=0}^{N_{pix}-1} \left(\frac{1}{\sigma_k^2(p,q)} \left| \mathbf{i}_k(p,q) - \sum_{t=0}^{N-1} (\mathbf{h}_{k-l}(p,q) * \mathbf{o}_l(p,q)) \right|^2 \right), \tag{2}$$

where $\sigma_k(p,q)$ is the noise variance in the layer k for the pixel (p,q).

C. Object Prior

The choice of a Gaussian prior probability distribution for the object can be justified from an information theory standpoint as being the least informative, given the first two moments of the distribution. In this case, a reasonable model of the object's power spectral density (PSD) can be found[14] and used to derive the regularization criterion J_o, which is then quadratic (or $L2$ for short). The chosen PSD model is

$$\mathrm{PSD}(f) = E[|\mathbf{o}(f)|^2] - |\mathbf{o}_m(f)|^2 = k/[1 + (f/f_0)^p] - |\mathbf{o}_m(f)|^2,$$

where f is the spatial frequency, $\mathbf{o}_m$ is the *a priori* object (it is typically a constant), p characterizes the regularity of the object, and f_0 is a cutoff frequency introduced to avoid the divergence at the origin and is typically the inverse of the characteristic size of the image. Additionally, the parameters of the object's PSD can be estimated automatically (i.e., in an unsupervised way) from the data by a maximum-likelihood method[15] derived from the method developed by Blanc *et al.*[16] in the phase-diversity context.

The disadvantage of a Gaussian prior (or, equivalently, of a quadratic regularization term), especially for objects with sharp edges such as photoreceptors or vessels, is that it tends to oversmooth edges. A possible remedy is to use an edge-preserving prior that is quadratic for small gradients and linear for large ones.[17] The quadratic part ensures a good smoothing of the small gradients (i.e., of noise), and the linear behavior cancels the penalization of large gradients (i.e., of edges), as explained by Bouman and Sauer.[18] Such priors are called quadratic–linear, or $L2$–$L1$ for short.[19] Here we use a function that is an isotropic version of the expression suggested by Rey[20] in the context of robust estimation, used by Brette and Idier[21] for image restoration, and recently applied to imaging through turbulence.[2,13] The choice of the crossover point from $L2$ to $L1$ is currently supervised and performed as

explained by Mugnier *et al.*[13] It is typically of the order of the mean difference between adjacent pixels in the image.

The functional J_o is strictly convex, and J_i of Eq. (2) is convex because it is quadratic, so that the global criterion $J=J_i+J_o$ is strictly convex. This ensures uniqueness and stability of the solution with respect to noise and also justifies the use of a gradient-based method for the minimization.

3. MYOPIC 3D DECONVOLUTION

In this section we address the case where the PSF **h** is not known precisely. An approach that has proven effective for 2D imaging is myopic deconvolution, i.e., performing a joint estimation of the object **o** and the PSF **h**. Unfortunately, for an N-plane 3D object and 3D image, the 3D PSF is composed of $2N-1$ layers. The problem is more underdetermined than in two dimensions. Furthermore, this method does not make use of the strong relationship between PSF planes: the different 2D PSFs differ only by a defocus. Because we have short-exposure images, we can parameterize the whole 3D PSF by a common pupil phase φ plus a known defocus phase that depends on the considered PSF plane. This has been already used for short-exposure 2D imaging through atmospheric turbulence.[2,22,23] This dramatically reduces the number of unknowns (we assume that we know the distance between two layers). Additionally, the pupil phase is expanded on Zernike polynomials (as defined by Noll[24]) so that at most a few tens of coefficients are required to describe the 3D PSF:

$$h_k(\varphi) = |\mathrm{FT}^{-1}\{P(x,y)\exp(j(\varphi(x,y) + \varphi_d^k(x,y)))\}|^2,$$

$$\varphi(x,y) = \sum_{m=5}^{M} a_m Z_m(x,y), \tag{3}$$

where P is the pupil function and φ_d^k corresponds to the defocus phase of layer k. φ is the unknown pupil phase, and $j^2=-1$. This defocus phase is calculated with

$$\varphi_d^k(x,y,\delta z) = a_4^d(\delta z)\,.\,Z_4(x,y),$$

where

$$a_4^d(\delta z) = \pi \cdot \frac{\delta z}{8\,.\,\sqrt{3}\,.\,\frac{\lambda}{n}\,.\left(\frac{f\,.\,n}{D}\right)^2},$$

where λ is the imaging wavelength in the air, n is the refractive index, f is the focal distance of the eye in the air, and D is the pupil diameter.

We jointly estimate the 3D object and the pupil phase in the same MAP framework. This joint MAP estimator is

$$[\hat{\mathbf{o}},\hat{\varphi}] = \arg\max_{\mathbf{o},\varphi} p(\mathbf{o},\varphi|\mathbf{i}) = \arg\max_{\mathbf{o},\varphi} p(\mathbf{i}|\mathbf{o},\varphi)\times p(\mathbf{o})\times p(\varphi).$$

Equivalently, $\hat{\mathbf{o}}$ and $\hat{\varphi}$ can be defined as the object and the phase that minimize a compound criterion $J(\mathbf{o},\varphi)$ defined as follows:

$$J(\mathbf{o},\varphi) = J_i(\mathbf{o},\varphi) + J_{\mathbf{o}}(\mathbf{o}) + J_\varphi(\varphi), \tag{4}$$

where $J_i=-\ln p(\mathbf{i}|\mathbf{o},\varphi)$ is the negative log likelihood and is given by Eq. (2), except that now it is considered a function of **o** and φ. $J_{\mathbf{o}}(\mathbf{o})=-\ln p(\mathbf{o})$ is a $L2$ or $L2-L1$ regularization criterion (see Subsection 2.C). We assume a Gaussian probability density function for φ, so $J_\varphi(\varphi)=-\ln p(\varphi)$ is a regularization criterion on the phase defined by

$$J_\varphi(\varphi) = \tfrac{1}{2}(\varphi-\overline{\varphi})^t C_\varphi^{-1}(\varphi-\overline{\varphi}),$$

where $\overline{\varphi}$ is the *a priori* phase mean (usually zero) and C_φ is the *a priori* phase covariance matrix. Additionally, because the images considered here are illuminated rather uniformly (due to all the out-of-focus object planes contributing to each image), stationary white Gaussian statistics, with a constant variance equal to the mean number of photoelectrons per pixel, is a reasonable approximation for the noise model, so that J_i simplifies to

$$J_i(\mathbf{o},\varphi) = \frac{1}{2\sigma_{\mathbf{n}}^2}\sum_{k=0}^{N-1}\left\|\mathbf{i}_k - \sum_{l=0}^{N1}(\mathbf{h}_{k-l}(\varphi) * \mathbf{o}_l)\right\|^2.$$

The criterion $J(\mathbf{o},\varphi)$ of Eq. (4) is minimized numerically on **o** and φ. The minimization is performed by the optimpack variable metric with limited memory and bounds (OP-VMLMB) method, designed by Thiébaut.[25] This method is faster than the conjugate-gradient method.

The simplest way to organize the unknowns for the minimization is to stack the object and the phase together into a vector and to run the OP-VMLMB routine on this variable. Yet, this can be slow as the gradients of the criterion with respect to the object and to the phase may have different orders of magnitude. We have found that the minimization is speeded up by splitting it into two blocks and alternating between minimizations on the object for the current phase estimate and minimizations on the phase for the current object estimate.

The minimization starts by estimating the object for a fixed (zero) phase. The initial guess for the object is, for instance, the image itself. The minimization is not stopped by hand but, rather, when the estimated object and phase no longer evolve (i.e., when their evolution from one iteration to the next is close to machine precision).

4. VALIDATION BY SIMULATIONS AND LIMITATIONS

A. Simulations

To validate our deconvolution method by simulations, we created a simulated object that complies with the overall structure of a retina. Figure 2 represents the original simulated object, composed of vessels, ganglion cells, and photoreceptors. The vessels are simulated by moving a ring in a random walk, the ganglion cells are simulated by empty globes, and photoreceptors are represented by two empty half-spheres joined by an empty tube. The cube's height on Fig. 2 is approximative 52 μm, and the depth and the width of this cube are 300 pixels.

In the simulations presented here, we use a five-slice object obtained by averaging the data from Fig. 2 into five

13 μm thick slices from which we select a 128×128 region of interest (the depth of focus is approximatively 18 μm). The five slices obtained are presented on Fig. 3.

The PSFs used to compute the 3D image **i** are currently purely diffractive (no multiple scattering). They are generated with a set of aberrations expanded on the Zernike basis (the Z_i coefficients are normalized); we use 0.2 rd root-mean-square (RMS) on the aperture of astigmatism (Z_5), −0.1 rd RMS on the aperture of astigmatism (Z_6), and −0.5 rd RMS on the aperture of spherical aberration (Z_{11}). These PSFs are oversampled (with respect to the Nyquist frequency) by a factor of 1.5. With the object and the PSF, we simulate the image by means of Eq. (1). The noise added is white Gaussian and stationary; its standard deviation is 3% of the maximum intensity in the object **o** (corresponding roughly to 1000 photoelectrons per pixel (ph/pix) for photon-limited data). The five image layers are presented on Fig. 4. From these images, it is clear that all object slices contribute to all images. With the relatively small chosen separation between planes (13 μm), the first two images are visually identical, whereas the corresponding object slices are very different. The deconvolution aims at disentangling the contribution of each object slice and improving the resolution within each plane.

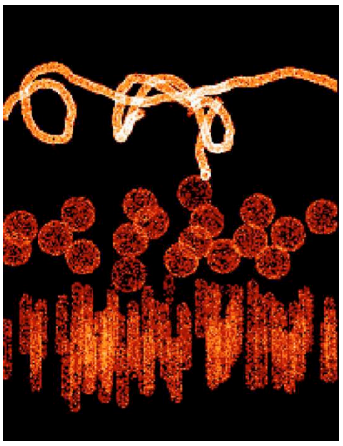


Fig. 2. (Color online) Perspective view of the 3D object used for the simulations.

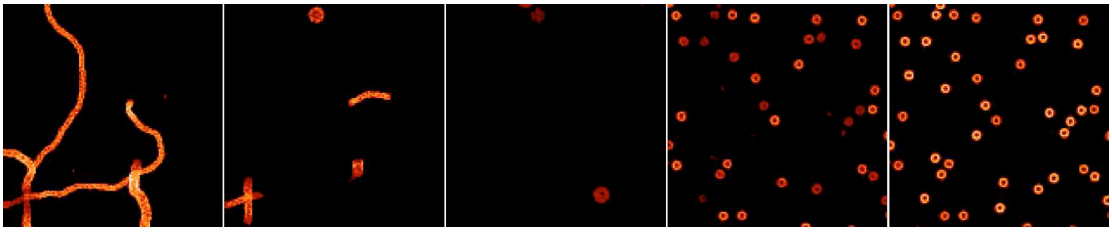


Fig. 3. (Color online) Five object layers [black corresponds to 0 photoelectrons per pixel (ph/pix)].

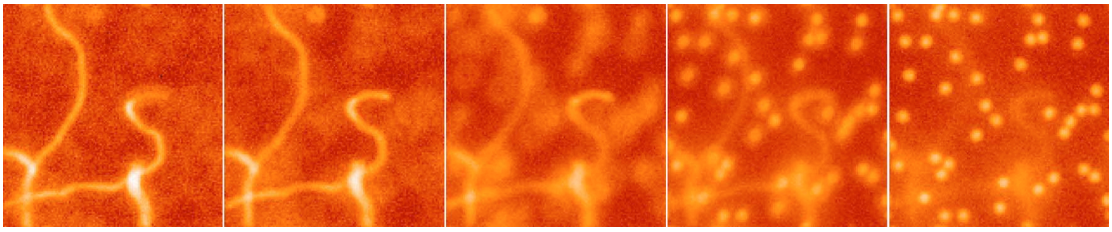


Fig. 4. (Color online) Five image layers.

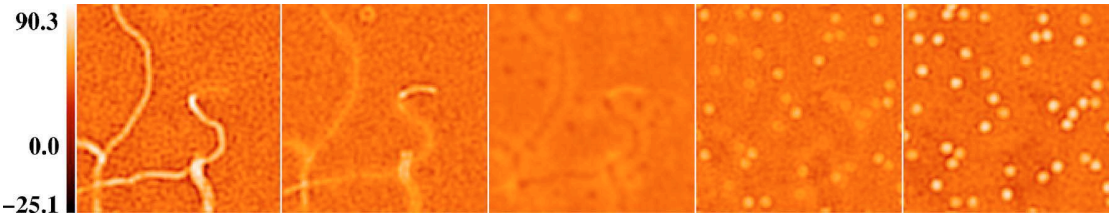


Fig. 5. (Color online) Five estimated object layers with $L2$ regularization without the positivity constraint and using the true PSF.

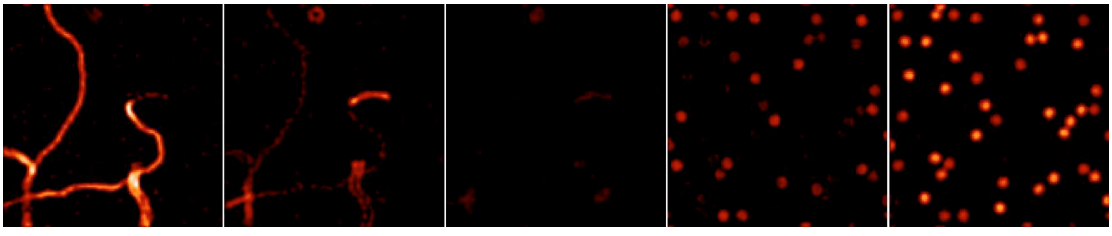


Fig. 6. (Color online) Five estimated object layers with $L2$ regularization under the positivity constraint and using the true PSF (black corresponds to 0 ph/pix).

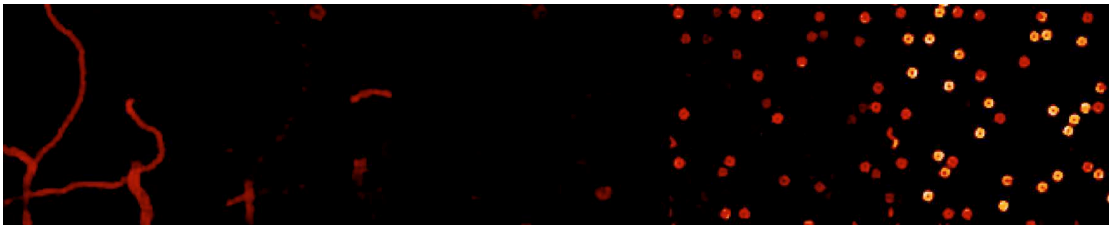


Fig. 7. (Color online) Five estimated object layers with $L2-L1$ regularization under the positivity constraint and using the true PSF (black corresponds to 0 ph/pix).

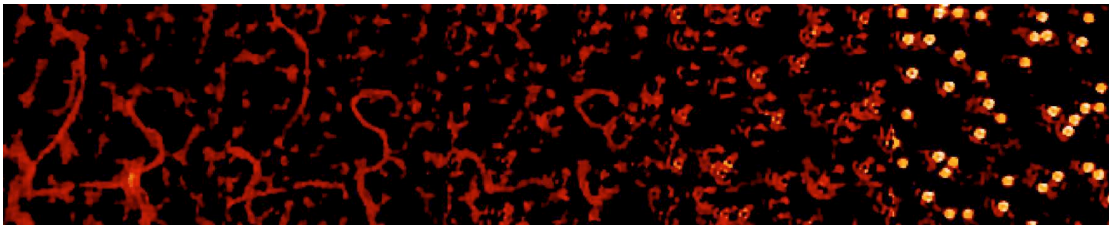


Fig. 8. (Color online) Deconvolution with a wrong (unaberrated) PSF. The different object planes are not correctly disentangled because of the mismatch between true PSF and assumed PSF (black corresponds to 0 ph/pix).

B. Deconvolution with Known PSF

In this subsection, we present three results obtained with our deconvolution method and the two priors mentioned in Subsection 2.C. The first simulation, presented on Figs. 5 and 6, shows the deconvolution results obtained with $L2$ regularization without and with the positivity constraint, respectively. We can see ghosts of vessels (in the middle plane, for example) on Fig. 5 and a residual blur: the missing cone of 3D frequencies makes it difficult for the restoration procedure to correctly disentangle the contribution of all planes. Edges are not preserved ($L2$ regularization and no positivity constraint prevent spectral extrapolation). The positivity constraint used in Fig. 6 helps the algorithm disentangle the differents planes and visibly reduces ghosts of vessels in the middle plane. More quantitatively, the RMS restoration error is 8.34 ph/pix with the positivity constraint and 10.31 ph/pix without (the object average level is 15.34 ph/pix). On Fig. 7 we present a deconvolution performed with $L2-L1$ regularization under the positivity constraint. The edges are much better preserved, and the separation between the different planes is also slightly better on the second restored image plane. The RMS restoration error is 6.33 ph/pix. To evaluate the need for precision in the PSF knowledge, we performed a deconvolution with a wrong (unaberrated) PSF shown in Fig. 8. The regularization used is $L2-L1$ under the positivity constraint, and the RMS restoration error is 11.28 ph/pix. On both Figs. 7 and 8, the lateral resolution is improved with respect to that of the images (Fig. 4). But only on Fig. 7 are the object planes correctly disentangled. In other words, the longitudinal resolution is very poor in Fig. 8 due to the mismatch between the true PSF and the one assumed for the deconvolution.

C. Results with the Myopic Method

We present here the estimated aberrations (see Fig. 9) with the myopic method [joint estimation of $\mathbf{o}$ and φ by minimization of criterion $J(\mathbf{o},\varphi)$ of Eq. (4)]. The true pupil phase standard deviation is $\sigma_\varphi = 0.53$ rd, and the RMS error with the positivity constraint is $\sigma_\varepsilon = 0.24$ rd. Without

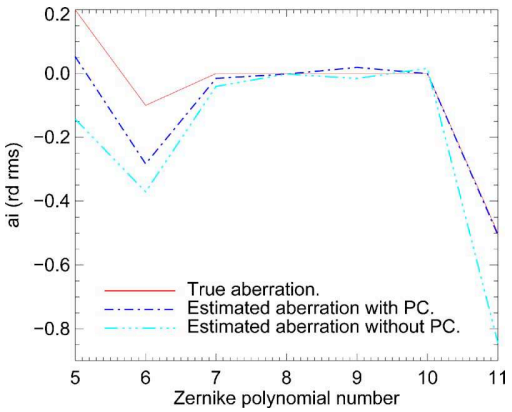


Fig. 9. (Color online) Estimated aberrations with and without the positivity constraint (PC).

the positivity constraint, the RMS error is $\sigma_\varepsilon = 0.56$ rd. The estimated phase without the positivity constraint cannot reasonably be used to deconvolve images. A likely explanation for the poor results of the method without the positivity constraint is that the criterion may have several minima. It has been shown in Subsection 2.C that the criterion $J(\mathbf{o},\varphi)$ is strictly convex for any given φ so there exists a unique object solution for a given set of aberrations φ, denoted by $\hat{\mathbf{o}}(\varphi)$: $\hat{\mathbf{o}}(\varphi) = \arg\min_{\mathbf{o}} J(\mathbf{o},\varphi)$. To validate the hypothesis of a nonconvex criterion, we define a partially optimized criterion as $J'(\varphi) \triangleq J(\hat{\mathbf{o}}(\varphi),\varphi)$, and we perform a plot of this criterion. If several minima are found on φ for J', then it is the unambiguous sign of the existence of several minima on $(\mathbf{o},\varphi)$ for the criterion $J(\mathbf{o},\varphi)$: $\hat{\varphi} = \arg\min_{\varphi} J'(\varphi) \Leftrightarrow (\hat{\mathbf{o}}(\hat{\varphi}),\hat{\varphi}) = \arg\min_{\mathbf{o},\varphi} J(\mathbf{o},\varphi)$. The plots of the values of $J'(\varphi)$ (computed for a grid of a_5 and a_6 values taken between -1 rd and 1 rd and for the true values of the other aberrations) are presented in Figs. 10 and 11. The criterion plotted on Fig. 11 without the positivity constraint presents several minima and is obviously nonconvex, whereas the one obtained with the positivity constraint plotted on Fig. 10 shows a global minimum that is close to the true aberra-

tions. In this case at least, the positivity constraint restricts the solution space to a unique solution to the minimization problem.

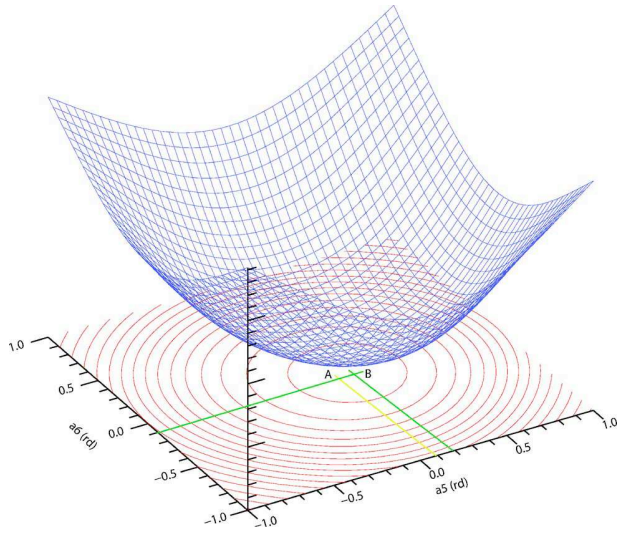


Fig. 10. (Color online) Criterion surface with the positivity constraint: the criterion is strictly convex. *A* is the position of the global minimum; it is close to the position of the true aberrations *B*.

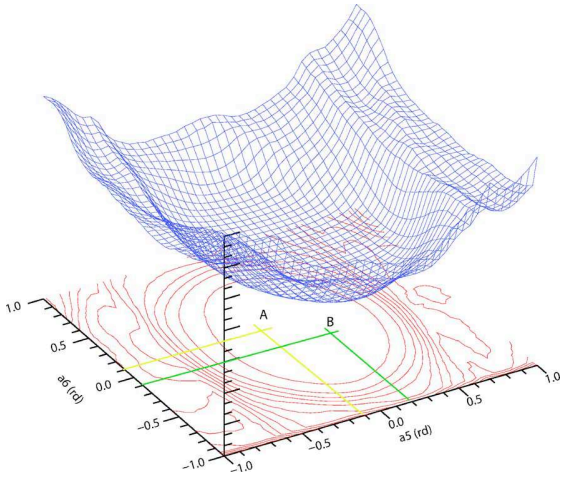


Fig. 11. (Color online) Criterion surface without the positivity constraint: the criterion is nonconvex. *A* is the position of the global minimum, and *B* is the position of the true aberrations.

For an object with a background, the positivity constraint becomes less and less effective as the background level increases. For a very high background, the deconvolution tends to the one obtained on Fig. 11 without the positivity constraint, as checked by earlier simulation.[26] Because the positivity constraint is not always effective, we wish to find another, more effective, constraint in order to improve the phase estimation.

5. PHASE DIVERSITY

We first briefly present the classical phase-diversity wavefront sensing technique and the case in which it is used. Then we introduce our 3D extension of it, and we validate it with some simulations.

A. Conventional Phase Diversity

Phase diversity is a focal-plane wavefront sensing technique proposed by Gonsalves[27] (see Ref. 28 for a review), which uses two (or more) images close to a focal plane to estimate the aberrations of an optical instrument. These two images (as shown on Fig. 12) differ by a known aberration (for instance, defocus) in order to estimate the pupil phase via a criterion minimization.

The two images recorded on the imaging camera are the convolution of the object by the PSF plus photon and detector noises. As shown in Eqs. (5), there is a nonlinear relation between the PSF and the parameter of interest φ:

$$i_f = h(\varphi) * o + n,$$

$$i_d = h(\varphi + \varphi_d) * o + n, \tag{5}$$

where h is defined in Eq. (3), φ is the phase, φ_d is the known aberration, o is the observed object, n is the noise, and * stands for the convolution process.

B. 3D Phase Diversity

Despite the fact that the myopic method described in Section 3 uses a 3D imaging model, that it uses a PSF model parameterized by the pupil phase (only a few tens of Zernike coefficients are required to describe the 3D PSF), and that it uses the positivity constraint, the precision obtained on estimated aberrations is modest (see Fig. 9). Furthermore, the estimation of aberrations without the positivity constraint is unacceptable. A reinterpretation of 2D phase diversity, classically used with opaque objects,

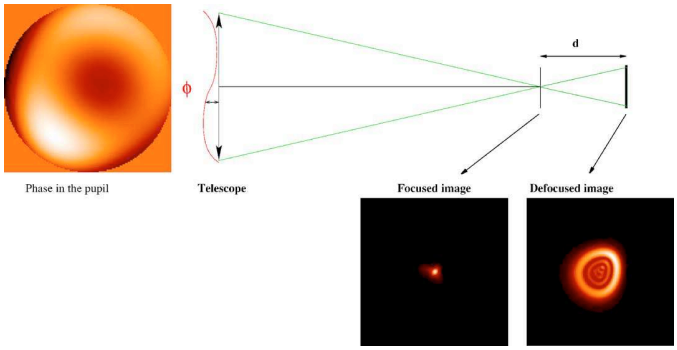


Fig. 12. (Color online) Principle of phase diversity: two images differing by a known aberration (here defocus) are used to estimate the pupil phase.

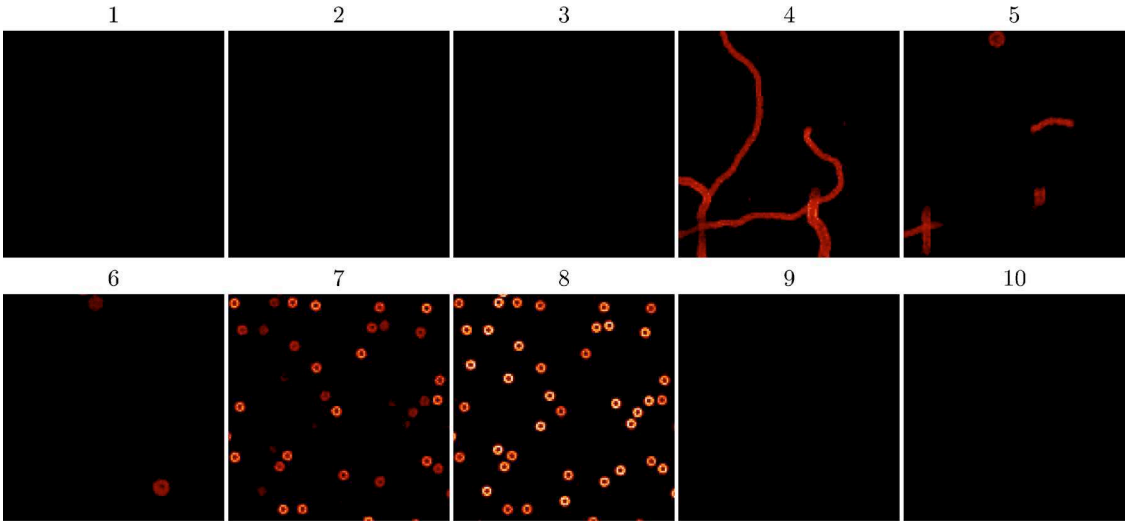


Fig. 13. (Color online) Ten object layers where five layers are empty (black corresponds to 0 ph/pix).

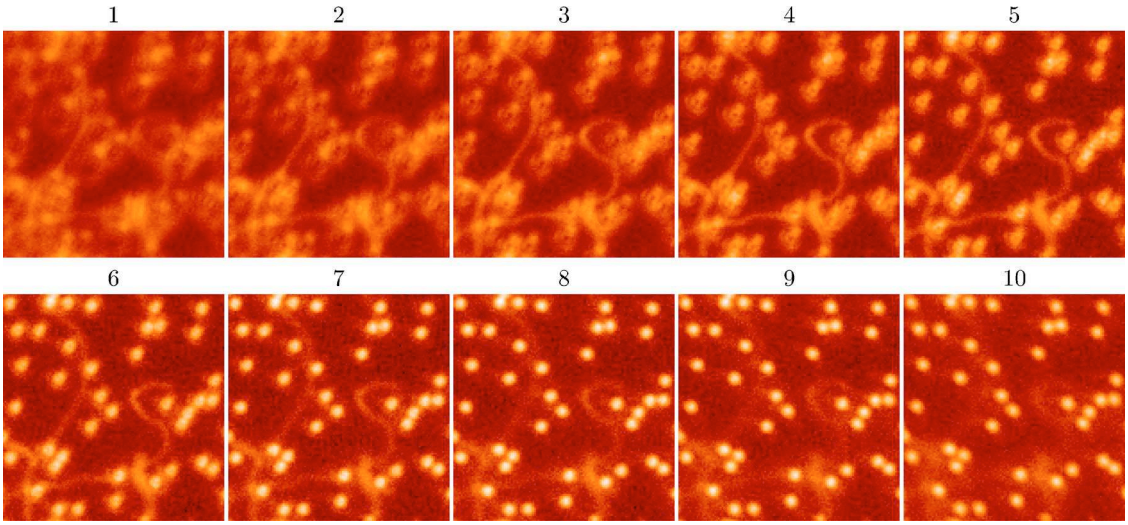


Fig. 14. (Color online) Ten simulated image layers.

is that only one out of the two image planes contains an object. In contrast, the myopic deconvolution used so far in this paper uses as many object planes as there are images. This 3D interpretation of conventional phase diversity prompts us to use a few additional images focused before (and possibly after) the object of interest. Furthermore, in the eye, we can indeed easily record images with no object (images focused in the vitreous, for instance). We assume and impose that some $\mathbf{o}$ planes are empty (see Fig. 13) and call this the Z support constraint. This additional prior knowledge is a strong constraint for the phase inversion, which makes the positivity constraint unnecessary at least in the conditions of the simulations presented below. The criterion of fidelity to the data $J_i(\mathbf{o},\varphi)$ (see Section 3) becomes

$$J_i(\mathbf{o},\varphi)=\frac{1}{2\sigma_{\mathbf{n}}^2}\sum_{k=0}^{N-1}\left\|\mathbf{i}_k-\sum_{\substack{l=0\\ l\in S_{\mathbf{o}}}}^{N-1}(\mathbf{h}_{k-l}(\varphi)*\mathbf{o}_l)\right\|^2,$$

where $S_{\mathbf{o}}$ is the list of nonempty object plane numbers. Typically, $S_{\mathbf{o}}=[l_{min},l_{max}]$ with $l_{min}\geq 0$ and $l_{max}\leq N-1$.

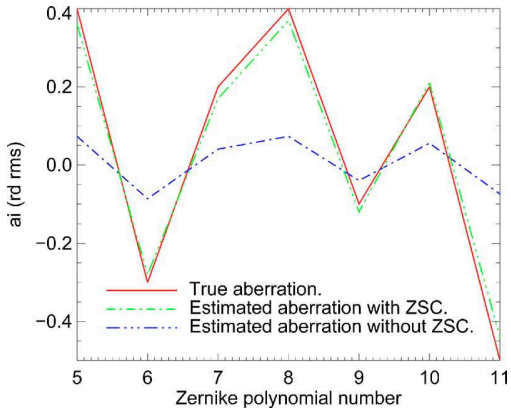


Fig. 15. (Color online) Estimated aberrations with and without the Z support constraint (ZSC).

The 3D imaging performed on an opaque 3D object can use a hard constraint: for any (x,y), at most one object voxel $[x,y,z(x,y)]$ re-emits the light because of the object opacity.[29,30] In our case, which is 3D imaging performed

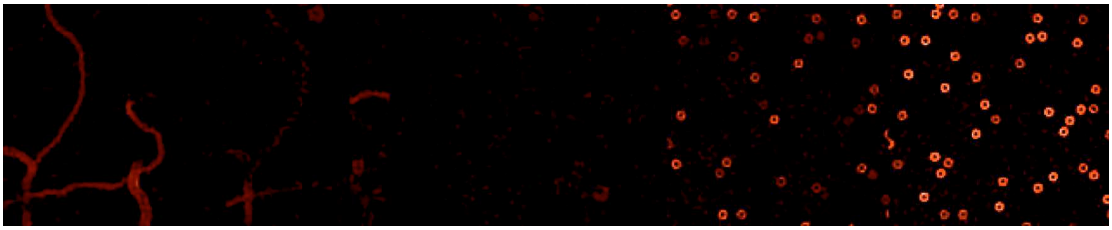


Fig. 16. (Color online) Five estimated object layers with $L2-L1$ regularization under the positivity constraint and ZSC (black corresponds to 0 ph/pix).

on a translucent object (the retina), such an opacity constraint is inappropriate. That is why we propose the Z support constraint, which can be expressed as $\mathbf{o}(x,y,z)=0\,\forall z \notin [z_{min},z_{max}]$.

C. Validation by Simulations

To validate the efficiency of the Z support constraint, we performed numerous simulations, of which one is presented here.

The simulation conditions (noise, distance between two planes, PSF oversampling) are the same as in Subsection 4.A except that the aberrations are stronger to test the 3D phase-diversity method. The object $\mathbf{o}$ is composed of five layers with the object (the same as in Subsection 4.A) and five without (see Fig. 13). The ten image layers are presented on Fig. 14. The true pupil phase standard deviation is $\sigma_\varphi=0.87$ rd. The results are presented in Fig. 15: the RMS error with the Z support constraint is only $\sigma_\varepsilon=0.088$ rd. Without the Z support constraint, the RMS error is $\sigma_\varepsilon=0.70$ rd, which is unacceptable. The phase estimation with the Z support constraint is precise enough to correctly deconvolve images: the deconvolution result (with this estimated phase) is given in Fig. 16. The RMS restoration error is 6.68 ph/pix with $L2-L1$ regularization under the positivity constraint and Z support constraint. We have checked that if we use the true PSF and $L2-L1$ regularization under the positivity constraint (the RMS restoration error is 6.33 ph/pix, see Subsection 4.B) instead of the one estimated with $L2-L1$ regularization under the positivity constraint and Z support constraint, we obtain a restored object that is visually identical to that of Fig. 16.

6. CONCLUSION AND PERSPECTIVES

A myopic 3D deconvolution method has been developed in a Bayesian framework. It uses a 3D imaging model, a fine noise model, an appropriate regularization term, and a parameterization of the point-spread function via the pupil phase. To improve the deconvolution performance, in particular for cases when the positivity constraint on the object is not effective (object with background), we have proposed the use of 3D phase diversity. This consists of recording additional images focused before (and possibly after) the object of interest and adding the corresponding longitudinal (Z) support constraint to the deconvolution. This can be very appropriate in particular for retinal images. We have demonstrated the effectiveness of the method and, in particular, of the Z support constraint on realistic simulated data.

To check the robustness of our myopic deconvolution method with respect to imperfections in the imaging model, a definitive validation should be performed on experimental data; this will constitute the next step of our work.

The corresponding author's e-mail address is Guillaume.Chenegros@onera.fr.

REFERENCES

1. J. Primot, G. Rousset, and J.-C. Fontanella, "Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images," J. Opt. Soc. Am. A **7**, 1598–1608 (1990).
2. L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem, "Myopic deconvolution from wave-front sensing," J. Opt. Soc. Am. A **18**, 862–872 (2001).
3. D. Catlin and C. Dainty, "High-resolution imaging of the human retina with a Fourier deconvolution technique," J. Opt. Soc. Am. A **19**, 1515–1523 (2002).
4. G. Rousset, J.-C. Fontanella, P. Kern, P. Gigan, F. Rigaut, P. Léna, C. Boyer, P. Jagourel, J.-P. Gaffard, and F. Merkle, "First diffraction-limited astronomical images with adaptive optics," Astron. Astrophys. **230**, 29–32 (1990).
5. M. C. Roggemann, "Limited degree-of-freedom adaptive optics and image reconstruction," Appl. Opt. **30**, 4227–4233 (1991).
6. J. M. Conan, P. Y. Madec, and G. Rousset, "Image formation in adaptive optics partial correction," in *Active and Adaptive Optics*, F. Merkle, ed., Vol. 48 of ESO Conference and Workshop Proceeding (European Southern Observatory/International Commission of Optics, 1994).
7. J.-M. Conan, "Étude de la correction partielle en optique adaptative," Ph.D. thesis (Université Paris XI, 1994).
8. J. G. McNally, T. Karpova, J. Cooper, and J. A. Conchello, "Three-dimensional imaging by deconvolution microscopy," Methods **19**, 373–385 (1999).
9. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," J. Opt. Soc. Am. A **21**, 1393–1401 (2004).
10. A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems* (Winston, 1977).
11. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024–2036 (1989).
12. J. Idier, ed., *Approche Bayésienne pour les Problèmes Inverses* (Hermès, 2001).
13. L. M. Mugnier, T. Fusco, and J.-M. Conan, "MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive optics-corrected long-exposure images," J. Opt. Soc. Am. A **21**, 1841–1854 (2004).
14. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
15. D. Gratadour, D. Rouan, L. M. Mugnier, T. Fusco, Y. Clénet, E. Gendron, and F. Lacombe, "Near-infrared

adaptive optics dissection of the core of NGC 1068 with NaCo," Astron. Astrophys. **446**, 813–825 (2006).
16. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035–1045 (2003).
17. P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," IEEE Trans. Med. Imaging **9**, 84–93 (1990).
18. C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," IEEE Trans. Image Process. **2**, 296–310 (1993).
19. J. Idier and L. Blanc-Féraud, "Deconvolution en imagerie," in *Approche Bayésienne pour les Problèmes Inverses*, J. Idier, ed. (Hermès, 2001), Chap. 6.
20. W. J. J. Rey, *Introduction to Robust and Quasi-Robust Statistical Methods* (Springer-Verlag, 1983).
21. S. Brette and J. Idier, "Optimized single site update algorithms for image deblurring," in *Proceedings of the International Conference on Image Processing* (IEEE Computer Society, 1996), pp. 65–68.
22. T. J. Schulz, "Multiframe blind deconvolution of astronomical images," J. Opt. Soc. Am. A **10**, 1064–1073 (1993).
23. E. Thiébaut and J.-M. Conan, "Strict *a priori* constraints for maximum-likelihood blind deconvolution," J. Opt. Soc. Am. A **12**, 485–492 (1995).
24. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. A **66**, 207–211 (1976).
25. E. Thiébaut, "Optimization issues in blind deconvolution algorithms," in *Astronomical Data Analysis. II*, J.-L. Starck and F. D. Murtagh, eds., Proc. SPIE **4847**, 174–183 (2002).
26. G. Chenegros, L. M. Mugnier, and F. Lacombe, "3D deconvolution of adaptive-optics corrected retinal images," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIII*, J.-A. Conchello, C. J. Cogswell, and T. Wilson, eds., Proc. SPIE **6090**, 60900P (2006).
27. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829–832 (1982).
28. L. M. Mugnier, A. Blanc, and J. Idier, "Phase diversity: a technique for wave-front sensing and for diffraction-limited imaging," in *Advances in Imaging and Electron Physics*, P. Hawkes, ed. (Elsevier, 2006), Vol. 141, Chap. 1.
29. M. F. Reiley, R. G. Paxman, J. R. Fienup, K. W. Gleichman, and J. C. Marron, "3D reconstruction of opaque objects from Fourier intensity data," in *Image Reconstruction and Restoration II*, T. J. Schulz, ed., Proc. SPIE **3170**, 76–87 (1997).
30. R. G. Paxman, J. H. Seldin, J. R. Fienup, and J. C. Marron, "Use of an opacity constraint in three-dimensional imaging," in *Inverse Optics III*, M. A. Fiddy, ed., Proc. SPIE **2241**, 116–126 (1994).

Annexe I

Article Gratadour et coll., 2005 [A23]

A&A 443, 357–365 (2005)
DOI: 10.1051/0004-6361:20042188

Astronomy & Astrophysics

Sub-pixel image registration with a maximum likelihood estimator

Application to the first adaptive optics observations of Arp 220 in the *L′* band⋆

D. Gratadour[1,2], L. M. Mugnier[1], and D. Rouan[2]

[1] DOTA – ONERA, Av. de la division Leclerc, Chatillon, France
[2] LESIA, Observatoire de Paris, 5 place Jules Janssen, 92195, Meudon, France
e-mail: damien.gratadour@obspm.fr

Received 15 October 2004 / Accepted 26 July 2005

ABSTRACT

We present a new method based on a maximum likelihood (ML) estimation of the sub-pixel shift between images of a given object observed with a single instrument. We first study the case of two noisy images and give the ML approach of the registration problem. By means of simulations, we show the gain obtained with this ML solution compared to a classical registration method with an academic noise model (stationary white Gaussian), and then demonstrate the relevance of this ML estimation with a more realistic noise model. We then address the problem of a sequence of low signal frames of the same object. We develop a joint ML approach in which we simultaneously estimate the reference (i.e. the noiseless) image and the shift parameters. The registration accuracy is increased at low photon levels as the number of frames grows, reaching the sub-pixel domain at very low *SNR* (about 1), when considering 100 frames. When applied to experimental data (thermal IR images of a faint galaxy), both ML methods show their efficiency to recover the resolution in averaged frames and totally outperform the classical cross-correlation.

Key words. techniques: image processing – instrumentation: adaptive optics – galaxies: individual: Arp 220

1. Introduction

Image registration at the level of a pixel or less is a common problem in many domains, and the number of developed methods makes it a full-fledged research field (see Brown 1992; and Zitová & Flusser 2003). Nevertheless, there is no universal solution as the method to use strongly depends on the nature of the transformation between the images to be registered. Registration for medical imaging is probably one of the most diversified fields, with many different types of problems (multi-modal, 3D or stereoscopic imaging for instance), for which different methods have been developed in each case (see Maurer & Fitzpatrick 1993; and Maintz & Viergever 1998, for a review). In this paper we focus on the problem of shifted noisy Nyquist-sampled images of the same object observed with the same instrument. Hence, following the general classification (Zitová & Flusser 2003; Roche et al. 2000) of registration methods, we propose a new area-based (or intensity-based) technique for image registration. The most popular method used in this case is the cross-correlation (CC) between the images (see e.g. Brown 1992; and Zitová & Flusser 2003, and references therein for a review). If the sub-pixel accuracy is required, interpolation of the CC function around its maximum is needed. Many interpolation methods have been tested which have different robustness and accuracy depending on the noise nature and intensity (see Roche et al. 2000; Zitová & Flusser 2003). For instance, in the case of undersampled frames, Carfantan & Rougé (2001) have studied different methods to find an unbiased estimation of the sub-pixel shift between two images, and conclude that only the interpolation of the CC function by a cardinal sine does not give a biased estimation. A similar approach is to compute the CC of images after oversampling them (with a cardinal sine or a cubic interpolation for instance). The accuracy of the method depends on the interpolation function. The latter approach is the most used method in astronomy when the images are at least Nyquist-sampled (see e.g. the registration method included in the IDL astro-Lib package from NASA). A classical method also used for astronomical images is the Drizzle algorithm (Fruchter & Hook 2002), developed for under-sampled HST images, which corrects for misregistration as well as for rotation and distortions. The correction for translation is also based on the CC of the interpolated images and thus suffers from the same limitations: explicit re-sampling and no precise noise model taken into account.

Some authors have proposed registration as a Maximum Likelihood (hereafter ML) problem (Mort & Srinath 1988; Costa et al. 1993), and Roche et al. (2000) have developed a

⋆ Based on observations collected at the ESO/Paranal YEPUN telescope, Proposal 70.B-0307(A).

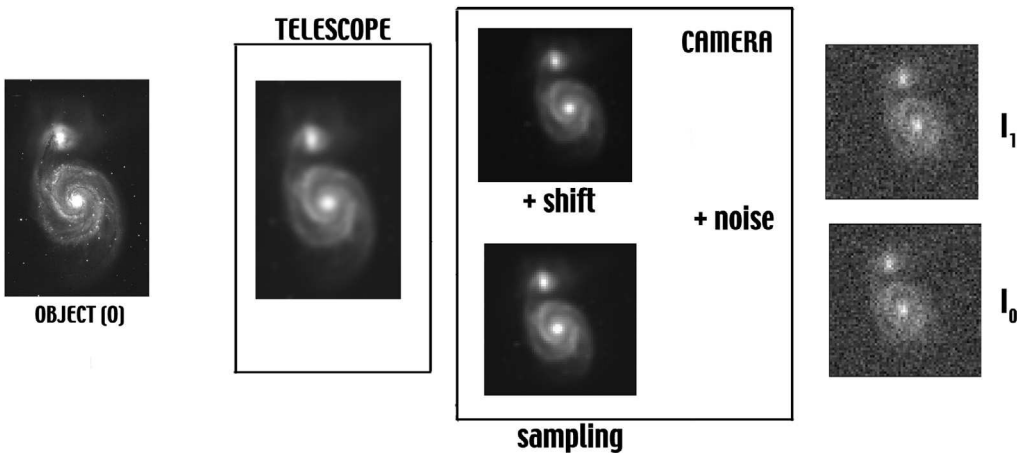


Fig. 1. Scheme describing the formation of an image. The sampling is assumed to verify the Nyquist-Shannon theorem.

general approach for registration problems using a ML formalism. The particular case of two images obtained with the same instrument is not studied because the main issue of their paper is the registration of images acquired under different observing conditions. Closer to our study, in the case of a sequence of translated Poisson limited Nyquist sampled images of the same object, a joint ML method has been developed to restore the reference image and the shift parameters of the sequence simultaneously (Guillaume et al. 1998).

The registration problem also occurs when collecting several short exposure images under weak turbulent conditions. Indeed, under such conditions, the effect of turbulence is mainly a tip-tilt on the wavefront and thus a translation of the images. This problem is addressed in the seminal paper of Snyder & Schulz (1990), in which they describe the EM algorithm they used to restore a high resolution image from a sequence of translated short exposures. This work and the work of Guillaume et al. (1998) are dedicated to extremely low photon levels (a few photons per image) whereas we consider images with a few photons per pixel corrupted by background and detector noise. With the greater photon level we consider, we aim at sub-pixel accuracy even when the object flux is comparable to the noise variance (detector+background noise). Moreover, when the problem allows such an approach, conjugate gradient algorithms are usually recognised to be faster than EM.

Some authors have also considered registration as part of a more global restoration problem, and propose ML approaches, depending on the images properties and the nature of the transformation, and including the registration problem. Recently, Girémus & Carfantan (2003) have developed a multiframe deconvolution algorithm, for which sub-pixel accuracy is required, and propose a joint ML estimator of the images and the shift parameters, but only in the case of stationary white Gaussian noise. They note that in the case of low *SNR*, the ML method for the estimation of the shift parameters gives more accurate results than classical empirical estimators. However, if we consider a PSF almost constant in time, as in the case of adaptive optics (hereafter AO) observations, no gain is expected by jointly registering and restoring as demonstrated by Girémus & Carfantan (2003). It is obviously less costly to register first all the frames and then deconvolved the average (long exposure) image.

This paper is organized as follows. We first describe our image model and introduce the ML formalism in the simple case of registering two images. We test our method with simulated images, in the case of stationary white Gaussian noise as well as with a more realistic mixture of Poissonian and stationary Gaussian noise. In Sect. 3, we address the problem of registering a sequence of images and develop a joint maximum likelihood approach to estimate simultaneously the shift parameters and the reference image. We then demonstrate the efficiency of this method with simulated images. In Sect. 4, we use various methods to register very noisy images of a distant galaxy acquired with an AO system in the thermal infrared and we compare the results. We conclude in the last section.

2. Description of the method

2.1. Data model

In the following, we use a one dimensional development for clarity. This section aims at presenting the context from which we naturally introduce the ML approach. The simplest statistical model of the data is obtained when one of the two images is considered to be noiseless, and the other one is a shifted and noisy version of the former. An illustration of this scheme is presented in Fig. 1.

We consider an object $O(x)$ seen through an instrument of known PSF $H(x)$. The latter includes the effect of the propagating medium, the telescope and the detector. Assuming no detection noise, a reference image is then $R(k) = [R(x)]_{\text{Ш}}(k) = [O(x) * H(x)]_{\text{Ш}}(k)$, where Ш is the sampling operator, $[R(x)]_{\text{Ш}}(k)$ is the kth sample of the function $R(x)$ and $*$ the convolution operator. We will assume that the sampling process respects the Shannon-Nyquist criterion, meaning that the highest spatial frequency in the image is at most half the sampling frequency. This implies that a continuous version of the image can be reconstructed via the Shannon reconstruction theorem. A shifted and noisy version of the reference

(acquired with a finite exposure time after jittering the telescope), is then written:

$$I_1(k) = [R(x) * \delta(x - x_1)]_{\text{III}}(k) + N_1(k)$$
$$= [R(x - x_1) \times \text{III}(x)]\,(k) + N_1(k)$$

where x_1 is the shift parameter and N_1 an additive noise.

2.2. Gaussian noise – formulation of the pairwise ML approach

In this section, we present the standard ML formulation of the registration problem in the image formation scheme described above. We will name this approach pairwise ML, as we process pairs of images, to distinguish this method from the one developed in Sect. 3. If we assume a Gaussian additive noise, the likelihood of observing an image $I_1(k)$ for the reference (noiseless image) $R(x)$ and for the hypothesis x_1 is given by:

$$\mathcal{L}\Big(I_1(k); R(x), x_1\Big) \propto \prod_k \mathrm{e}^{\left(-\frac{1}{2\sigma_1^2(k)}\left|I_1(k)-[R(x-x_1)]_{\text{III}}(k)\right|^2\right)}$$

where σ_1^2 is a map of the noise variance (see for instance Sect. 3 in Van Trees 1968). The negative log-likelihood is then:

$$\mathcal{J}(x_1) = \sum_k \frac{1}{2\sigma_1^2(k)}\left|I_1(k) - [R(x - x_1)]_{\text{III}}(k)\right|^2 \quad (1)$$

and the ML estimate $\hat{x}_{1_{\text{ML}}}$ of the shift between the two images is given by minimizing numerically $\mathcal{J}(x_1)$ by, e.g., a conjugate gradient method.

2.2.1. Stationary white Gaussian noise: the cross-correlation

If the noise is additionally stationary, which is for instance approximatively the case for background dominated noise if the background is uniform (thermal IR observations for instance) or for low photon level images dominated by detector noise, the pairwise ML solution of the registration problem is known to be the maximum of the cross-correlation between the two images, assuming their periodicity (see Appendix A).

The formulation of the registration in a ML framework suggests an implementation of the solution that can be arbitrarily sub-pixel without any explicit resampling of the images, and which turns out to be more precise than readily available (IDL astro library, for instance) code, as shown in next section.

2.2.2. Generalization to non-stationary Gaussian noise

We now consider a more realistic noise model describing astronomical observations in the infrared. In this framework, the global noise on the images can be decomposed in two components: a Poisson distributed noise due to the detection process (object + background) and a Gaussian distributed noise due to the CCD electronics. This case has been previously studied in Snyder et al. (1995), where a sophisticated noise model is considered. In their paper, the distribution of the mixture of Gaussian and Poissonian noises is either approximated by a Poissonian distribution or computed numerically using the saddle point method. The performance of the image restoration ML criteria deduced from this two distributions are very close which demonstrates the robustness of the Poissonian approximation in their study (image reconstruction using one image of the object).

In our study, we use the alternative approach of approximating our mixed noise as non-stationary white Gaussian. Actually, in the case of astronomical infrared imaging, the level of the IR background is comparable to the photon flux of the observed object and is typically over a thousand photons per pixel. We are thus working at high background level and the IR background noise distribution is well approximated by a Gaussian. After background subtraction the noise variance is doubled, and the photon level on the object is of the order of a few photons per pixel (more than the variance of the detector noise, i.e. typically more than ten for nowaday's detectors).

In the case described above, the global noise variance can be estimated directly from the images as in Mugnier et al. (2004). The variance of the detector component is estimated beforehand, on a so-called *dark* image, and the photon noise variance is estimated as:

$$\sigma^2_{\text{ph}}(k) = \max[I(k), 0].$$

We obtain the pairwise ML solution of the shift by minimizing numerically the criterion of Eq. (1) (and the periodic assumption is no longer needed). With a map of the noise variance, estimated as just described, the ML estimation of the shifts is easier than in the study of Snyder & Schulz (1990) in which the noise variance and the shifts are estimated simultaneously.

2.2.3. Generalization to an unknown reference

In practice, the noiseless reference R is not available. Instead, let us now consider a noisy frame I_0:

$$I_0(k) = [R(x)]_{\text{III}} + N_0(k)$$

assumed to be a noisy version of the reference. A shifted noisy image I_1 can be written:

$$I_1(k) = [R(x) * \delta(x - x_1)]_{\text{III}} + N_1(k)$$
$$= [I_0(x) * \delta(x - x_1)]_{\text{III}} + N(k)$$

where: $N(k) = N_1(k) - [N_0(x) * \delta(x - x_1)]_{\text{III}}$. If we assume N_0 as a Gaussian distribution of variance $\sigma_0^2(x)$ and N_1 the noise of the same detector, then we have: $\sigma_0^2(x) * \delta(x - x_1) = \sigma_1^2(x)$.

As the noise variance of I_1 is estimated on the image, and that $\sigma_0^2(x)$ is deduced from this estimation, the noise variance is estimated before the minimization of $\mathcal{J}(x_1)$ and thus no more depends on the shift parameters. The neg-log-likelihood to be minimized has so the same expression as in Eq. (1) changing $R(x)$ into $I_0(x)$ and $\sigma_1^2(k)$ into $\sigma^2(k) = \sigma_1^2(k) + \left[\sigma_0^2(x) * \delta(x - x_1)\right]_{\text{III}} = 2\sigma_1^2(k)$:

$$\mathcal{J}(x_1) = \sum_k \frac{1}{4\sigma_1^2(k)}\left|I_1(k) - [I_0(x - x_1)]_{\text{III}}\right|^2. \quad (2)$$

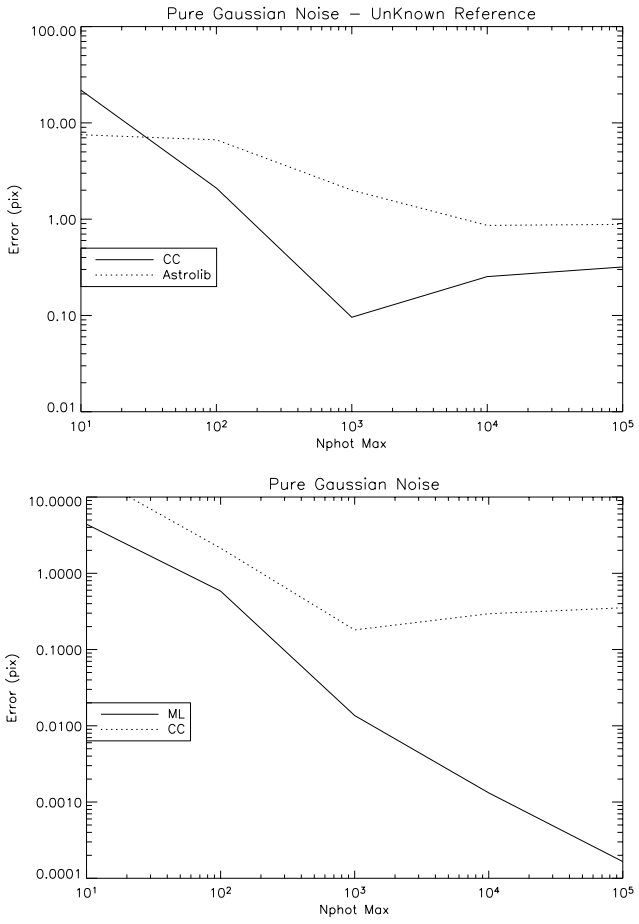


Fig. 2. *Top*: performance of the IDL astro-Lib registration method in the case of purely Gaussian additive noise compared to the CC we implemented. *Bottom*: respective error of each method in the same conditions, solid line, the pairwise ML method, dotted line, the classical cross-correlation method we implemented. The variance of the Gaussian noise is kept a constant at 100, while the number of photons max per pixel in the initial images ranged from 10 to 10^5.

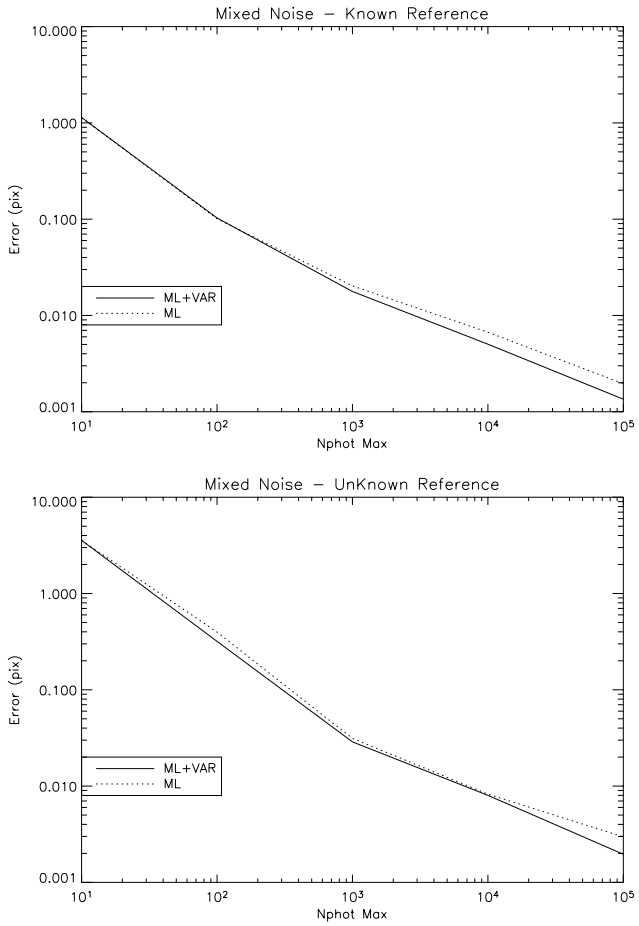


Fig. 3. Performance of the pairwise ML method in the case of mixed (stationary Gaussian + Poissonian) noise. The variance of the Gaussian noise is kept constant to 100, while the max number of photons per pixel in each image ranges from 10 to 10^5. Solid line: the ML method considering the noise variance map, dotted line: the ML method considering a constant variance. *Top*, the academic case of a known reference and *bottom*, the realistic case of an unknown one.

We see that even if the 2 images are noisy, one of them can still serve as a reference for the other and that the ML solution just derived is still valid (see Fig. 3).

2.3. Implementation and validation with simulated images

In order to compare our method to commonly used ones in this kind of registration problems, we implemented (in IDL) the minimization of the criterion of Eq. (1). We compared it to an *easy-to-implement* cross-correlation method which consists of a discrete computation of the CC function followed by a fit of the latter around its maximum by a continuous function. We chose a Gaussian over a second order polynomial to fit the CC function, in a 2×2 pixel box around its maximum. When compared to a classical image interpolation-based method (the one included in the IDL astro-library for instance) this CC method is obviously less time consuming and has better performance (see Fig. 2 top). This CC method thus seems a robust and fast member of the classical CC methods family and so is a relevant test competitor to compare to our pairwise ML method.

If the images are Nyquist sampled, the Fourier transforms (FT) of the sampled images coincide with the FT of the continuous images within $[-\frac{\nu_s}{2}, \frac{\nu_s}{2}]$ where ν_s is the sampling frequency. Therefore, we are able to reconstruct a sub-pixel shifted version of any Nyquist sampled image, by multiplying the phase of the FT of this image by a tilt whose slope is a non integer number. In practice, we used the Fast FT (FFT) algorithm to make the computation faster. This algorithm computes a Discrete FT (DFT) which represents a small approximation. In order to remove noise and to reduce the effect of this approximation at high frequencies, we have found it useful to low-pass filter the noisy reference before applying the above-mentioned tilt to its phase. The cut-off is chosen to equal the diffraction cut-off frequency of the telescope, so that only noise is removed from this image. The criterion is then implemented as:

$$\mathcal{J}(x_1) = \sum_k \frac{1}{4\sigma_1^2(k)} \left| I_1(k) - \mathcal{F}^{-1}\left[\widetilde{I_0}(u) \times \Pi_{u_c}(u) \times \mathrm{e}^{-2i\pi u x_1}\right]\right|^2 \qquad (3)$$

where $\mathcal{F}^{-1}$ denotes the inverse FFT, $\widetilde{I}$ denotes the FT of an image, x_1 a real number and $\Pi_{u_c}(u)$ is the low-pass mask in the Fourier domain.

We consider an object that we convolve by a perfect PSF (i.e. the PSF of a perfect telescope without aberrations of turbulence). We then build series of 100 identical images randomly shifted, with a uniform distribution of shift parameters, in each direction. The maximum shift value is 1 since we want to test the sub-pixel accuracy. Moreover, the integer part of the shift parameters can usually be retrieved from the files headers when the shift is induced by a jitter of the instrument. These 100 images can then be corrupted by noise.

2.3.1. Pure stationary Gaussian noise

In this case we consider only detector noise: Gaussian stationary noise is added to each image with a constant variance of 100, which is typical of current AO equipped cameras (e.g. NAOS/CONICA). We then build 5 series of 100 images, increasing the maximum number of photons per pixel in each series, i.e. increasing the *SNR*. Each series is then registered with the three methods (astro-lib, CC and pairwise ML). The reference is taken as one image in each series, so we are in the realistic case of an unknown reference.

The results presented in Fig. 2 bottom demonstrate that, more than being intrinsically sub-pixel, our method exhibits an increase in accuracy with *SNR* in the images. The performance of the interpolated cross-correlation obviously depends on the interpolation method, and the robust one we used saturates close to 0.3 pixel accuracy. We see here that the interpolation method we used, in the sub-pixel regime appears equivalent to oversampling the images by a factor of 3 or more, as usually done (in IDL astro-Lib registration code, for instance). Nevertheless, none of the two interpolation-based CC method is able to do better than 0.1 pixel whereas the proposed ML method's accuracy is approximatively inversely proportional to the maximum number of photons per pixel.

We have hence demonstrated the efficiency of our sub-pixel method to register very noisy frames, in the classical case of stationary white Gaussian noise, and considering a noisy image as a reference. We reach the pixel accuracy as the maximum number of photon per pixel is comparable to the variance of the Gaussian noise and can go down to the thousandth of pixel at very high photon levels (10^5), which completely outperforms the classical cross-correlation method and its refinements.

2.3.2. Mixture of Gaussian and Poissonian noise

The method has also been tested with a more realistic noise model. The same shifted series of a reference object convolved by a PSF is first corrupted by Poissonian noise. We then add Gaussian stationary noise with a constant variance of 100. The total variance of the noise is then the shifted reference image (object convolved by the PSF and shifted) plus a constant for the Gaussian noise. We tested the method on 5 series of images with a maximum number of photons per pixel ranging from 10 to 10^5. We present the results in Fig. 3, in the case of known (i.e. a non noisy image of the sequence) and unknown reference (i.e. a noisy image of the sequence) of the ML method. Two methods are tested in each case, one considering a noise variance map, and one considering a constant variance. The aim of this more realistic modeling of the noise is, firstly, to validate the method in the case of a more sophisticated noise model, and secondly to prove the efficiency of the exact knowledge of the noise distribution. We see that in both cases (known and unknown reference), introducing a map of the noise variance slightly increases the accuracy as the max number of photon per pixel ($N_{\rm phot}$) reaches a reasonable value. Indeed, until the max number of photon per pixel is significantly greater than the variance of the detector noise, the total noise variance can be considered as constant, and the accuracy of the two methods is the same. Moreover, in the realistic case of an unknown reference, the advantage of using a true variance is lessened. Finally, we note that the accuracy depends on the *SNR*. Two regimes appear: at low photon level (detector dominated noise, $SNR \propto N_{\rm phot}$), the accuracy is proportional to the inverse of $N_{\rm phot}$; at high photon level (photon limited noise, $SNR \propto N_{\rm phot}^{1/2}$ the accuracy is proportional to the inverse of the square of $N_{\rm phot}$. As expected, the transition between the two regimes is smoother in the academic case of the known reference.

3. Sequences of images: the Joint ML approach

In this section, we address the problem of registering several (more than 2) images. Actually, it is very rare to have only a couple of noisy images of an object, and the general case is to have a sequence of many images. Intuitively, one should so be able to obtain a better estimation of the shift parameters between images if considering all the sequence. This approach is similar to what was developed by Guillaume et al. (1998) in the case of pure photon noise. We propose here a joint version of our ML registration, in order to estimate simultaneously the non noisy reference image and the shift parameters between the images.

3.1. Joint estimation of the reference and the shift parameters

We consider a series of images $\{I_j(k)\}$ randomly shifted. If we try to find simultaneously the shift parameters $\{x_j\}$ and the reference image $R(x)$, then the joint likelihood can be written as:

$$\mathcal{L}\Big(\{I_j(k)\}; R(x), \{x_j\}\Big) \propto \prod_m \prod_k \exp\left(-\frac{1}{2\sigma_m^2(k)}\Big|I_m(k) - \Big[R * \delta(x - x_m)\Big]_{\text{Ш}}(k)\Big|^2\right). \quad (4)$$

And, the neg-log-likelihood to be minimized is then:

$$\mathcal{J}\left(\{I_j(k)\}; R(x), \{\mu_j\}\right) = \sum_m \sum_k \frac{1}{2\sigma_m^2(k)}\Big|I_m(k) - \Big[R * \delta(x - x_m)\Big]_{\text{Ш}}(k)\Big|^2. \quad (5)$$

3.2. The joint ML estimate of the reference image

Following the approach of Guillaume et al. (1998), canceling this joint criterion with respect to the reference gives an analytical expression of a reference estimate that minimizes the criterion for a given set of $\{x_j\}$ (see Appendix A):

$$\hat{R}_{\mathrm{ML}}(k, \{(x_m)\}) = \frac{1}{N_{\mathrm{images}}} \sum_m [I_m(x) * \delta(x + x_m)]_{\mathrm{III}}(k). \quad (6)$$

This is a quite intuitive result, as it is nothing but the average of the shifted-back images. This result is the same as the one obtained by Guillaume et al. (1998) in the case of photon limited data. The criterion obtained by substituting R by its ML estimation $\hat{R}_{\mathrm{ML}}$ depends only on the shifts parameters. Therefore, the criterion to be minimized is given by:

$$\mathcal{J}\left(\{I_j(k)\}; \hat{R}_{\mathrm{ML}}(k), \{x_j\}\right) = \sum_m \sum_k \frac{\left|I_m(k) - \left[\hat{R}_{\mathrm{ML}}(x) * \delta(x - x_m)\right]_{\mathrm{III}}(k)\right|^2}{2\sigma_m^2(k)}. \quad (7)$$

Additionally, as shown by Blanc et al. (2003), assuming a Gaussian distribution for the object and the noise, we have:

$$\mathcal{J}\left(\{I_j(k)\}; \hat{R}_{\mathrm{ML}}(k), \{x_j\}\right) = \sum_m \sum_k \frac{\left|I_m(k) - \left[\hat{R}_{\mathrm{ML}}(x) * \delta(x - x_m)\right]_{\mathrm{III}}(k)\right|^2}{2\sigma_m^2(k)} + \mathcal{K} \quad (8)$$

where $\mathcal{K}$ is a corrective term depending on the square modulus of the transfer function H and the object and noise power spectral density. This term is constant in our problem, as we consider that the PSF is the same for all images. The solution of the joint estimator proposed can thus be interpreted as true ML solution of the sole shifts provided the prior distribution of the object is Gaussian.

3.3. Implementation and validation with simulated images

The criterion of Eq. (8) is implemented and minimized as in the previous case:

$$\mathcal{J}\left(\{I_j(k)\}; \hat{R}_{\mathrm{ML}}(k), \{_j\}\right) = \sum_m \sum_k \frac{1}{2\sigma_m^2(k)} \left|I_m(k) - \mathcal{F}^{-1}\left[\widetilde{\hat{R}_{\mathrm{ML}}}(u) \times \Pi_{u_c}(u) \mathrm{e}^{-2\mathrm{i}\pi(ux_1)}\right]\right|^2 \quad (9)$$

where $\hat{R}_{\mathrm{ML}}$ is given in Eq. (7) and $\Pi_{u_c}(u)$ is the low-pass mask in the Fourier domain cutting at the telescope cutoff frequency u_c, the same as the one used in the pairwise ML case. We performed the same kind of simulations as described previously. We compared the pairwise ML algorithm (ML × 2) described in Sect. 2 and the joint estimate with 10 groups of 10 images (ML × 10) and 1 group of 100 images (ML × 100). We plotted the average root-mean-square error in each direction for the whole sequence against the max number of photons per pixel in an image. At low photon levels, we see that increasing the number of frames increases the accuracy of the shift estimation. Actually, we reach a sub-pixel accuracy when the max number of photon per pixel in the image is comparable to the detector noise in the case of a group of 100 images. This has great interest in infrared astronomical imaging where the number of images is usually huge and the *SNR* is very poor. Note that in the case of a group of 100 images, as expected, the performance are very close to the case of a known reference since the large number of images allows us to retrieve an almost noiseless estimate of the reference.

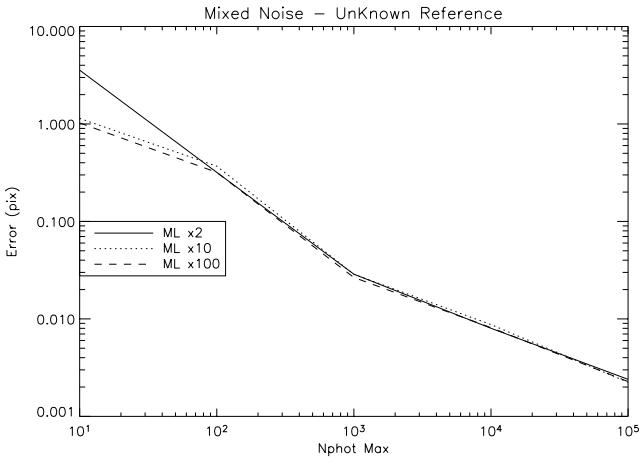


Fig. 4. Performance of the joint estimator compared to the pairwise ML estimator, in the case of mixed (additive Gaussian+Poissonian) noise. The variance of the Gaussian noise is constant: $\sigma^2_{\mathrm{gauss}} = 100$ photons. The number of photons max per pixel ranges from 10 to 10^5.

4. Application to infrared astronomical images: first images of Arp 220 in the *L*-band with adaptive optics

In order to evaluate the performance of our method in a real case, we recorded a sequence of images of a very faint galaxy. Arp 220 is a typical Ultra-Luminous Infrared Galaxy, caracterised by a very powerful emission in the infrared bands but a very faint counterpart at visible wavelengths. Observation of such a galaxy is a challenge as its distance to earth impose high angular resolution imaging in the infrared. NACO (NAOS+CONICA) at the VLT is the only AO system that can servo on infrared sources and thus achieve diffraction limited images of such galaxies in the infrared with a large ground-based telescope (Rousset et al. 2003; Lagrange et al. 2003). A series of 85 images of this galaxy has been acquired in the L' band (3.8 μm) with this instrument in March 2003. While the background dynamics of each image is around 80 000 photons per pixel before sky subtraction, the maximum level on the source is around 100 photons per pixel which is comparable to the total variance of the background after sky subtraction. We are here in the case where the CC was not efficient enough to ensure good registration (accuracy poorer than a pixel). One elementary frame is presented in Fig. 5 as well as the 85 registered and averages frames using the three methods (interpolated cross-correlation, pairwise ML and joint ML. The

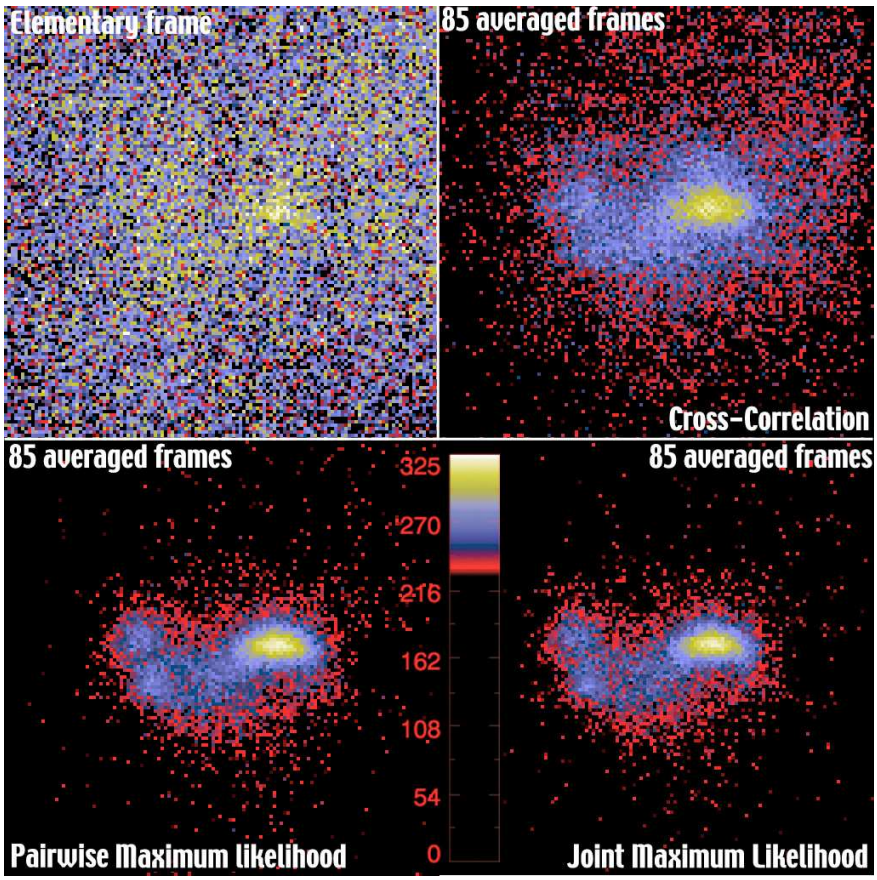


Fig. 5. Adaptive optics image of ARP 220 in the *L*-band with NACO at VLT. *Top left*: elementary frame, *top right*: 85 frames registered with a classical cross-correlation method and averaged; *bottom left*: 85 frames registered with the pairwise ML method and averaged; *bottom right*: 85 frames registered with the joint ML method and averaged. We chose a log-scale representation to ensure a good contrast for all structures in the image. The color distribution is kept constant for all averaged images to ensure comparison.

obtained Strehl ratio ranged from 0.1 to 0.15 at 2.2 μm during the observations meaning 0.25 at L'.

As obvious in Fig. 5, the images registered with ML methods are better defined than the one obtained with CC. The brightest source intensity is enhanced and it exhibits a clear East-West elongation. Moreover, in the case of the joint ML, it appears slightly bent, similar to the crescent nucleus observed with HST at 2.2 μm (Scoville et al. 1998). The surrounding structures show also more details with a clear *double-arm* pattern linking the two secondary sources to the brightest elongated source which was not obvious on the CC registered image, and globally, the joint ML method gives slightly sharper results than the pairwise method.

The comparison of the circular mean of the Fourier transforms of each image is displayed in the left part of Fig. 6 and shows more quantitatively the improvement obtained with both ML methods. High spatial frequencies are enhanced especially in the range 0.1 to 0.4 (in units of the cutting frequency of the telescope) compared to the CC method. The improvement obtained with the joint ML compared to the pairwise is also visible just below and around frequency 0.1.

The previous *L*-band images, obtained without adaptive optics on a 10-m class telescope (Soifer et al. 1999) are similar to the result obtain with CC registration. The super-imposition of the contour of their study on our image is presented on the right side of Fig. 6 (inset A). In their image, neither the two secondary sources nor the *double-arm* pattern were clearly defined. Clearly, the details we now obtain with the ML methods at 3.8 μm allow a more precise study of the secondary structures and their link to the brightest source. Precise photometry of these sources are now possible and can be compared to the same pattern found by HST at 2.2 μm and by Keck 12 μm as shown in insets B and C of Fig. 6. Very preliminary morphological interpretations were presented in Gratadour et al. (2003) and a precise photometric study including AO observations at 2.2 μm as well as more a refined interpretation will be presented in a future paper in preparation.

5. Conclusion

In this paper, we have formulated and tested by means of simulations a ML approach for the registration of shifted images of the same object observed with a single instrument. We have studied the performance of this method as a function of the *SNR* of the images. We have then demonstrated the efficiency of our ML method to register two noisy images at a sub-pixel accuracy, even when the signal is very low.

In order to obtain a higher accuracy when registering a large amount of very noisy elementary frames, we developed and

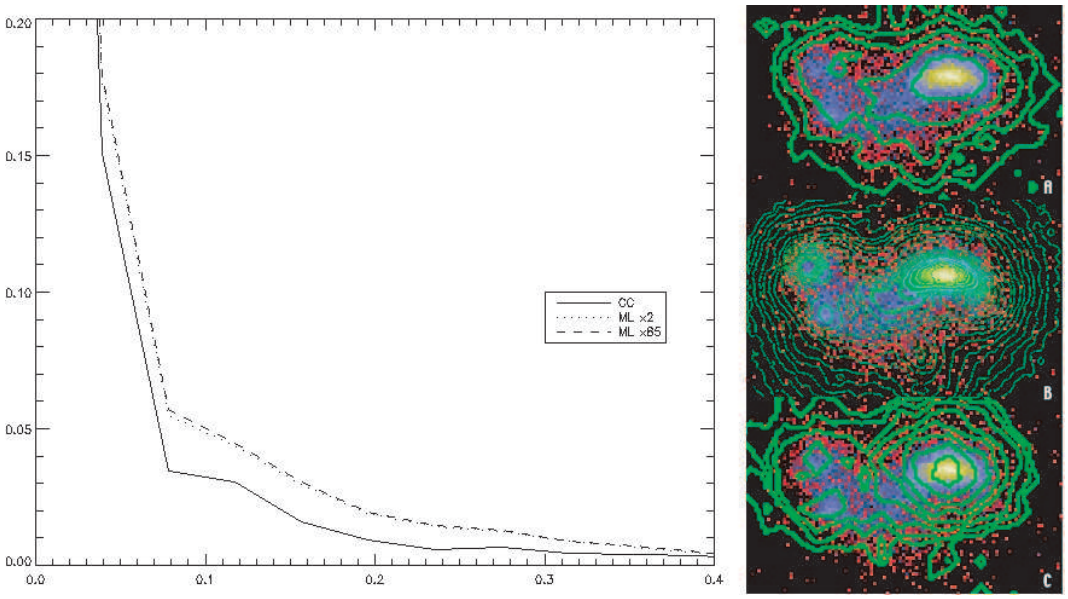


Fig. 6. *Left*: circular means of the Fourier transforms of the final images obtained with the cross-correlation (CC), the pairwise ML (ML × 2) and the joint ML (ML×85) methods. *Right*: super-imposition of our joint ML image with contours at other wavelengths: A, 3.8 μm with Keck, B, 2.2 μm with HST and C, 12 μm with Keck.

tested a joint ML approach which estimates simultaneously the reference image and the shift parameters between the frames. By deriving analytically the ML solution for the reference image, we obtain a criterion depending on the sole shift parameters which allows a fast and precise minimization. We demonstrated the efficiency of this joint ML estimate at low *SNR*s, and noted the similarity between this joint ML method and that of the pairwise ML at high *SNR*s. Finally, we registered experimental astronomical images of a faint galaxy and showed the ability of our ML approach to preserve the resolution obtained in each elementary images thanks to the AO after averaging all the frames.

Appendix A: The ML solution in the case of stationary white Gaussian noise

In this appendix, we recall how, with a mild assumption (periodicity of the images), in the case of stationary white Gaussian noise, the pairwise ML solution of the registration problem is the maximum of the cross-correlation. Indeed, an analytical expression of the gradient of the negative log-likelihood $\mathcal{J}$ with respect to x_1, assuming σ_1 is a constant, is given by:

$$\nabla_{x_1}\mathcal{J}(x_1) \propto \sum_k \left[I_1(k) - [R(x-x_1)]_{\mathrm{III}}(k)\right]\left[\frac{\partial}{\partial x_1}[R(x-x_1)]_{\mathrm{III}}(k)\right].$$

Under the assumption that the reference is periodic or not truncated (i.e. an object surrounded by a dark background larger than the largest possible shift): $\sum_k \left|[R]_{\mathrm{III}}(k)\right|^2 = \text{const.}$ we have:

$$\sum_k \left\{\frac{\partial}{\partial x_1}[R(x-x_1)]_{\mathrm{III}}(k)\right\}[R(x-x_1)]_{\mathrm{III}}(k) = 0$$

and, nulling $\nabla_{x_1}\mathcal{J}(x_1)$ is then equivalent to solving

$$\sum_k \left\{\frac{\partial}{\partial x_1}[R(x-x_1)]_{\mathrm{III}}(k)\right\} I_1(k) = 0$$

i.e., to find the maximum of:

$$C(I_1, R) = \sum_k I_1(k)[R(x-x_1)]_{\mathrm{III}}(k) \tag{A.1}$$

which is nothing but the linear cross-correlation function of the image and the reference.

Appendix B: The ML estimation of the reference in the joint ML case

In this appendix, we derive the analytical expression of the ML estimate of the reference in the joint ML approach by derivating the joint ML criterion with respect to the reference. We recall the expression of the joint ML criterion:

$$\mathcal{J}\left(\{I_j(k)\}; R(k), \{x_j\}\right) = \sum_m \sum_k \frac{\left|I_m(k) - \left[R(x) * \delta(x-x_m)\right]_{\mathrm{III}}(k)\right|^2}{2\sigma_m^2(k)}.$$

The maximum likelihood estimate $\hat{R}_{\mathrm{ML}}(x)$ of the reference image, is obtained by nulling $\forall l$:

$$\frac{\partial}{\partial R(l)}\mathcal{J}\left(\{I_j(k)\}; R(x), \{x_j\}\right) = \frac{\partial}{\partial R(l)} \sum_m \sum_k \frac{1}{2\sigma_m^2(k)}\left|I_m(k) - [R * \delta(x-x_m)]_{\mathrm{III}}(k)\right|^2.$$

If we consider Nyquist sampled images then we can write:

$$R(x) = \sum_{n\in\mathbb{Z}} \mathrm{sinc}\,(x-n)R(n) \tag{B.1}$$

and:

$$[R(x)]_{\text{Ш}}(k) = R(x)\times\delta(x-k)$$

and:

$$\Big[R(x)*\delta(x-x_m)\Big]_{\text{Ш}}(k) = \sum_{n\in\mathbb{Z}} \mathrm{sinc}\,(k-x_m-n)R(n)$$

so:

$$\frac{\partial}{\partial R(l)}\Big[R(x)*\delta(x-x_m)\Big]_{\text{Ш}}(k) = \mathrm{sinc}\,(k-x_m-l).$$

Then, we have:

$$\frac{\partial}{\partial R(l)}\mathcal{J}\left(\{I_j(k)\};R(x),\{x_j\}\right) =$$
$$\sum_m\sum_k \frac{-1}{\sigma_m^2(k)}\mathrm{sinc}\,(k-x_m-l)\Big[I_m(k)-\Big[R(x)*\delta(x-x_m)\Big]_{\text{Ш}}(k)\Big].$$

We see that:

$$\sum_k \frac{1}{\sigma_m^2(k)}\mathrm{sinc}\,(k-x_m-l)I_m(k) =$$
$$\sum_k \frac{1}{\sigma_m^2(k)}\mathrm{sinc}\,(l+x_m-k)I_m(k) = \left[\frac{I_m(x+x_m)}{\sigma_m^2(x+x_m)}\right]_{\text{Ш}}(l)$$

and, assuming that R is Nyquist sampled,

$$\sum_k \frac{[R(x-x_m)]_{\text{Ш}}(k)}{\sigma_m^2(k)}\mathrm{sinc}\,(k-x_m-l) = \left[\frac{R(x)}{\sigma_m^2(x+x_m)}\right]_{\text{Ш}}(l).$$

So we see that $\forall l$:

$$\frac{\partial}{\partial R(l)}\mathcal{J}\left(\{I_j(k)\};R(x),\{x_j\}\right) = 0 \Leftrightarrow$$
$$\sum_m\left[\frac{I_m(x+x_m)}{\sigma_m^2(x+x_m)}\right]_{\text{Ш}}(l) = \sum_m\left[\frac{R(x)}{\sigma_m^2(x+x_m)}\right]_{\text{Ш}}(l).$$

As argued in Sect. 2.2.3 $\sigma_m^2(x)*\delta(x-x_m) = \sigma_n^2(x)*\delta(x-x_n)$. Then, we can write:

$$\sum_m [I_m(x+x_m)]_{\text{Ш}} = \sum_m [R(x)]_{\text{Ш}}$$

and so:

$$\hat{R}_{\mathrm{ML}}(k) = \frac{1}{N_{\mathrm{images}}}\sum_m [I_m(x)*\delta(x+x_m)]_{\text{Ш}}(k). \tag{B.2}$$

Minimizing

$$\mathcal{J}\left(\{I_j(k)\};R(x),\{x_j\}\right)$$

on $R(x)$ and $\{x_j\}$ is equivalent to minimize:

$$\mathcal{J}\left(\{I_j(k)\},R(k)=\hat{R}_{\mathrm{ML}}(k);\{x_j\}\right)$$

on $\{x_j\}$ only as discussed in Sect. 3.2.

References

Blanc, A., Mugnier, L. M., & Idier, J. 2003, J. Opt. Soc. Am. A, 20, 1035

Brown, L. G. 1992, ACM Compting Surveys, 4, 325

Carfantan, H., & Rougé, B. 2001, in XVIII^{e} GRETSI, 849

Costa, W. L. S., Haynor, D. R., Haralick, R. M., Lwellen, T. K., & Graham, M. M. 1993, in IEEE Nuclear Science Symp. and Medical Imaging Conf., 974, 38

Fruchter, A. S., & Hook, R. N. 2002, PASP, 114, 144

Girémus, A., & Carfantan, H. 2003, in XIX^{e} GRETSI, 2

Gratadour, D., Rouan, D., Clenet, Y., Gendron, E., & Lacombe, F. 2003, in IAU Symp., 221

Guillaume, M., Melon, P., Refregier, P., & Llebaria, A. 1998, J. Opt. Soc. Am. A, 15, 2841

Lagrange, A., Chauvin, G., Fusco, T., et al. 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye, & A. F. M. Moorwood, Proc. SPIE, 4841, 860

Maintz, J. B. A., & Viergever, M. A. 1998, Medical Image Analysis, 2, 1

Maurer, C. R., & Fitzpatrick, J. M. 1993, Interactive image-guided neurosurgery, ed. R. J. Maciunas, American Association of Neurological Surgeons, Park Ridge, IL, 17

Mort, M. S., & Srinath, M. D. 1988, in Applications of digital image processing, IEEE Proc. of SPIE, 974, 38

Mugnier, L. Z., Fusco, T., & Conan, J. 2004, JOSAA, in press

Roche, A., Malandain, G., & Ayache, N. 2000, Int. J. Imaging Syst. Techn., 11, 71

Rousset, G., Lacombe, F., Puget, P., et al. 2003, in Adaptive Optical System Technologies II, ed. P. L. Wizinowich, & D. Bonaccini, Proc. SPIE, 4839, 140

Scoville, N. Z., Evans, A. S., Dinshaw, N., et al. 1998, ApJ, 492, L107

Snyder, D. L., Helstrom, C. W., Lanterman, A. D., Faisal, M., & White, R. L. 1995, OSAJ, 12, 272

Snyder, D. L., & Schulz, T. 1990, OSAJ, 7, 1251

Soifer, B. T., Neugebauer, G., Matthews, K., et al. 1999, ApJ, 513, 207

Van Trees, H. L. 1968, Detection, Estimation and Modulation theory Part I (John Wiley & Sons)

Zitová, B., & Flusser, A. 2003, Image and Vision Computing, 21, 977

Annexe J

Article Idier et coll., 2005 [A21]

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 14, NO. 12, DECEMBER 2005 2107

Statistical Behavior of Joint Least-Square Estimation in the Phase Diversity Context

Jérôme Idier, Laurent Mugnier, and Amandine Blanc

***Abstract*—The images recorded by optical telescopes are often degraded by aberrations that induce phase variations in the pupil plane. Several wavefront sensing techniques have been proposed to estimate aberrated phases. One of them is *phase diversity*, for which the joint least-square approach introduced by Gonsalves *et al.* is a reference method to estimate phase coefficients from the recorded images. In this paper, we rely on the asymptotic theory of Toeplitz matrices to show that Gonsalves' technique provides a consistent phase estimator as the size of the images grows. No comparable result is yielded by the classical *joint maximum likelihood* interpretation (*e.g.*, as found in the work by Paxman *et al.*). Finally, our theoretical analysis is illustrated through simulated problems.**

***Index Terms*—Error analysis, least-squares methods, optical image processing, parameter estimation, phase diversity, statistics, Toeplitz matrices.**

I. Introduction

THE images recorded by optical telescopes are often degraded by aberrations that induce phase variations in the pupil plane. In the case of ground telescopes, atmospheric turbulence is typically responsible for such phase aberrations. Imperfections of the optical system are another important source of errors, most of the latter being static while the former evolves with atmospheric turbulence.

Phase aberration is an ackowledged cause of degradation of the optical transfer function (OTF). The situation becomes far more favorable if the aberrated phases can be inferred and compensated. Several wavefront sensing techniques have been proposed to allow phase estimation. One of them is Gonsalves' *phase diversity* technique [1], [2]. It consists in the simultaneous acquisition of the usual focal plane image and of (at least) one additional image with a known defocus. Then the aberrations are numerically estimated using the information brought by the set of measured images.

Joint least-square (JLS) estimation of the aberrations and the observed object has been proposed by Gonsalves [1], [2], and it has since become the reference phase diversity technique. In [3], a statistical interpretation is given: The JLS estimate can be viewed as a joint maximum likelihood (ML) solution under the assumption of additive white Gaussian noise. In the present paper, our main objective is to examine its asymptotical statistical properties w.r.t. aberration estimation. By "asymptotical," we refer to a situation where the number of data points grows to infinity, and preferentially to the case where the size of the acquired images is arbitrarily large. Such a situation is clearly formal, *i.e.*, it is not aimed to be reproduced in practice.[1] However, the asymptotical behavior of the solution may bring meaningful information about its behavior in realistic situations.

As the number of data points grows to infinity, the optimality of ML estimation is granted in a wide theoretical framework. Unfortunately, the usual JLS solution to the phase estimation problem does not pertain to this framework, since the number of unknowns (*i.e.*, both the aberrated phase parameters *and the object*) increases with the number of observations. It rather corresponds to an approach studied by Little and Rubin [5]. According to their conclusions, this approach is not generally reliable from the statistical viewpoint, especially when the relative proportion of unknowns does not go to zero as the size of the data set increases.

In [6] and [7], a true ML estimate in the sense of [5] is proposed for the aberration parameters in the context of phase diversity: The unknown object is treated as a nuisance parameter, which means that it is integrated out to form the likelihood with respect to phase parameters. In contrast with the JLS solution, the theoretical asymptotical optimality of such a solution is granted.

Yet, it has been established by practical evidence that the behavior of the JLS solution to the phase parameter estimation is globally satisfactory. It is the aim of the present paper to examine the statistical properties of JLS type solutions more specifically. To our best knowledge, this is a fully open question, since the only few contributions devoted to statistical analysis of phase diversity imaging assume that the source object is known [8], [9].

Our main result is that the JLS solution possesses the essential features of a *minimum contrast estimator* [10, Section 3.2]. As such, it is a consistent estimator (*i.e.*, it converges toward the true value as the size of the data set increases).

II. Data Model

Let $\boldsymbol{i}_N = \{i(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\}$ correspond to a focused image measured on a square grid $\Omega_N = \{-N, \ldots, N\} \times$

Manuscript received August 24, 2004; revised November 2, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Robert P. Loce.

J. Idier is with the Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN), École Centrale de Nantes, 44321 Nantes Cedex 3, France (e-mail: jerome.idier@irccyn.ec-nantes.fr).

L. Mugnier is with the Office National d'Etudes et de Recherches Aérospatiales (ONERA), 92322 Chatillon Cedex, France (e-mail: laurent.mugnier@onera.fr).

A. Blanc is with the Office National d'Etudes et de Recherches Aérospatiales (ONERA), 92322 Chatillon Cedex, France, and also with the Laboratoire d'Astrophysique de l'Observatoire de Grenoble, 38041 Grenoble Cedex 9, France (e-mail: amandine.blanc@onera.fr).

Digital Object Identifier 10.1109/TIP.2005.859365

[1]Physical phenomena should then be taken into account, such as anisoplanatism in the case of extended objects observed through turbulence [4].

$\{-N,\ldots,N\}$ of size $Q_N \times Q_N$, with $Q_N = 2N+1$. In the isoplanatic patch of the telescope, it is obtained by noisy convolution of the object o with the focused point-spread function (PSF) h: $\forall\, \boldsymbol{x} \in \Omega_N$

$$i(\boldsymbol{x}) = (h \star o)(\boldsymbol{x}) + n(\boldsymbol{x}) = \int_{\mathbb{R}^2} h(\boldsymbol{y}) o(\boldsymbol{x}-\boldsymbol{y}) d\boldsymbol{y} + n(\boldsymbol{x}) \quad (1)$$

where $\star$ denotes convolution between functions of $\mathbb{R}^2$ and n corresponds to observation noise. The PSF is given by $h = |z|^2$, with

$$\forall\, \boldsymbol{x} \in \mathbb{R}^2, \quad z(\boldsymbol{x}) = \int_{\mathbb{R}^2} Z(\boldsymbol{\nu}) e^{2i\pi\langle x,\boldsymbol{\nu}\rangle} d\boldsymbol{\nu} \quad (2)$$

where $\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^2$ and $Z = Ae^{i\phi}$. The aperture function $A = |Z|$ is known and of limited spatial extent, and ϕ is the unknown aberrated phase function. Following [3] and others, we shall consider a finite linear decomposition for ϕ: $\phi = \sum_{j=1}^{J} \alpha_j \psi_j$. Typically, $\{\psi_j\}$ is a set of Zernike polynomials [11].

Let us remark that the OTF (*i.e.*, the Fourier transform of the PSF) has a finite support since A is of limited extent. In the sequel, it will be assumed that the OTF vanishes outside the square $\Gamma = [-1/2, 1/2] \times [-1/2, 1/2]$, so that no aliasing effect occurs. Then, it is possible to cast the observation model within a fully discrete framework

$$\forall\, \boldsymbol{x} \in \Omega_N, \quad i(\boldsymbol{x}) = (h * o)(\boldsymbol{x}) + n(\boldsymbol{x}) \quad (3)$$

where $*$ denotes convolution between functions of $\boldsymbol{Z}^2$, *i.e.*, $(h * o)(\boldsymbol{x}) = \sum_{\boldsymbol{y}\in\boldsymbol{Z}^2} h(\boldsymbol{y}) o(\boldsymbol{x}-\boldsymbol{y})$.

In order to allow practical computations, the convolution product in (3) must be restricted to finite arrays. Depending on the assumption made at the boundaries, several alternatives are possible, none of which being exact unless the object has a known finite support. In particular, the *periodic boundary* condition corresponds to *cyclic convolution*. For the sake of computational simplicity, it is the most commonly adopted approximation.

In all cases, the approximate observation model can then be described in a vector-matrix formulation using lexicographical orderings of the image and of the object [12]. Let $\mathrm{col}\{\boldsymbol{v}_N\}$ denote a column vector of length Q_N^2 corresponding to an array $\boldsymbol{v}_N = \{v(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\}$ scanned in lexicographical order. Then, all approximate observation models read

$$\mathrm{col}\{\boldsymbol{i}_N\} \simeq \mathbf{H}_{NM}\mathrm{col}\{\boldsymbol{o}_M\} + \mathrm{col}\{\boldsymbol{n}_N\} \quad (4)$$

where $\mathbf{H}_{NM}$ is a $Q_N^2 \times Q_M^2$ *Toeplitz-block-Toeplitz* (TBT) matrix with blocks of size $Q_N \times Q_M$, where M depends on the adopted approximation. In the case of cyclic convolution, $M = N$ and $\mathbf{H}_{NM} = \mathbf{H}_N$ is a square matrix with a *circulant-block-circulant* (CBC) structure: $\mathbf{H}_N\mathrm{col}\{\boldsymbol{o}_N\} = \mathrm{col}\{\boldsymbol{h}_N \circledast \boldsymbol{o}_N\}$, where the symbol $\circledast$ denotes cyclic convolution between two-dimensional (2-D) finite arrays. The most usual approximation corresponds to $\boldsymbol{h}_N = |\boldsymbol{z}_N|^2$ where $\boldsymbol{z}_N$ is obtained from Z by inverse 2-D discrete Fourier transform (DFT)

$$\boldsymbol{z}_N(\boldsymbol{x}) = Q_N^{-2} \sum_{y\in\Omega_N} Z\left(\frac{\boldsymbol{y}}{Q_N}\right) e^{2i\pi\langle x,y\rangle/Q_N}$$

which is practically implemented using fast Fourier transform techniques.

In defocused planes, the observation model (1)–(3) generalizes under the following form: $\forall\, k = 2,\ldots,K$

$$\begin{aligned} \forall\, \boldsymbol{x} \in \Omega_N, \quad i_k(\boldsymbol{x}) &= (h_k \star o)(\boldsymbol{x}) + n_k(\boldsymbol{x}), \\ &= (h_k * o)(\boldsymbol{x}) + n_k(\boldsymbol{x}) \quad (5) \\ \forall\, \boldsymbol{x} \in \mathbb{R}^2, \quad h_k(\boldsymbol{x}) &= |z_k(\boldsymbol{x})|^2 \quad (6) \\ z_k(\boldsymbol{x}) &= \int_\Gamma Z_k(\boldsymbol{\nu}) e^{2i\pi\langle x,\boldsymbol{\nu}\rangle} d\boldsymbol{\nu} \quad (7) \end{aligned}$$

with $Z_k = Ae^{i\phi_k}$, where $\phi_k - \phi = \theta_k$ are known phase increments. The usual approximation of cyclic convolution corresponds to[2]

$$\boldsymbol{i}_{kN} \simeq \mathbf{H}_{kN}\boldsymbol{o}_N + \boldsymbol{n}_{kN} = \boldsymbol{h}_{kN} \circledast \boldsymbol{o}_N + \boldsymbol{n}_{kN} \quad (8)$$

where $\forall\, \boldsymbol{x} \in \Omega_N$

$$\boldsymbol{h}_{kN}(\boldsymbol{x}) = |\boldsymbol{z}_{kN}(\boldsymbol{x})|^2 \quad (9)$$

with

$$\boldsymbol{z}_{kN}(\boldsymbol{x}) = Q_N^{-2} \sum_{\boldsymbol{y}\in\Omega_N} Z_k\left(\frac{\boldsymbol{y}}{Q_N}\right) e^{2i\pi\langle \boldsymbol{x},\boldsymbol{y}\rangle/Q_N}. \quad (10)$$

III. Joint Least-Square Approach

Let us introduce the following penalized least-square criterion:

$$\mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}_N) = \sum_{k=1}^{K} \|\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N\|^2 + \mu_N \|\mathbf{F}_N \boldsymbol{o}_N\|^2 \quad (11)$$

where conventionaly, the index value $k = 1$ refers to focused quantities (*e.g.*, $\boldsymbol{i}_1 \equiv \boldsymbol{i}$). In what follows, the default range of summation on image indices is $1,\ldots,K$.

Within the probabilistic framework, choosing the above penalizing term corresponds to assuming that the object $\boldsymbol{o}_N$ is a centered random vector with a covariance matrix proportional to $(\mathbf{F}_N^t\mathbf{F}_N)^{-1}$ (provided that $\mathbf{F}_N^t\mathbf{F}_N$ is actually invertible). Let us also remark that the original approach introduced by Gonsalves corresponds to $K = 2$ and $\mu_N = 0$. Choosing $\mathbf{F}_N$ as filtering low frequencies out and setting μ_N to some strictly positive value has a favorable regularizing effect on the quality of the restored object [7, Fig. 4(b)]. At low signal-to-noise ratio, it is also favorable with respect to phase estimation [7, Fig. 4(a)].

Finding a joint minimizer $(\hat{\boldsymbol{\alpha}}_N, \hat{\boldsymbol{o}}_N)$ of $\mathcal{J}_N$ can be partially simplified as follows [2], [3], [13]. For any $\boldsymbol{\alpha}$, minimizing $\mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}_N)$ as a function of $\boldsymbol{o}_N$ amounts to solving a quadratic programming problem. The set of solutions is characterized by the normal equation

$$\mathbf{R}_N\boldsymbol{o}_N = \sum_k \mathbf{H}_{kN}^t \boldsymbol{i}_{kN} \quad (12)$$

where

$$\mathbf{R}_N = \sum_k \mathbf{H}_{kN}^t\mathbf{H}_{kN} + \mu_N\mathbf{F}_N^t\mathbf{F}_N$$

[2]Thereafter, the operation of lexicographical reordering $\mathrm{col}\{\cdot\}$ is understood whenever unambiguous.

is the associated normal matrix. In particular, $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha}) = \mathbf{R}_N^+ \sum_k \mathbf{H}_{kN}^t \boldsymbol{i}_{kN}$ is the *generalized inverse* solution (matrix $\mathbf{R}_N^+$ denotes the pseudo inverse of $\mathbf{R}_N$) [14]. If $\mathbf{R}_N$ is full rank, then $\mathbf{R}_N^+ = \mathbf{R}_N^{-1}$ and $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})$ is the unique solution of (12). Let

$$\mathcal{K}_N(\boldsymbol{\alpha}) = \min_o \mathcal{J}_N(\boldsymbol{\alpha}, \boldsymbol{o}) = \mathcal{J}_N(\boldsymbol{\alpha}, \hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})).$$

Basic algebraic manipulations yield

$$\mathcal{K}_N(\boldsymbol{\alpha}) = \mathcal{K}'_N - \mathcal{K}''_N(\boldsymbol{\alpha}) \tag{13}$$

$$\mathcal{K}'_N = \sum_k \|\boldsymbol{i}_{kN}\|^2 \tag{14}$$

$$\mathcal{K}''_N(\boldsymbol{\alpha}) = \left(\sum_k \boldsymbol{i}_{kN}^t \mathbf{H}_{kN} \right) \mathbf{R}_N^+ \left(\sum_k \mathbf{H}_{kN}^t \boldsymbol{i}_{kN} \right). \tag{15}$$

Hence, in order to obtain $\hat{\boldsymbol{\alpha}}_N$, it suffices to maximize $\mathcal{K}''_N$, which only depends on the J unknown phase parameters $\boldsymbol{\alpha}$ through matrices $\mathbf{H}_{kN}$. Since no closed-form expression of the maximizer of $\mathcal{K}''_N$ is available, one must resort to some iterative optimization algorithm [3], [7], [13].

The case where matrices $\mathbf{H}_{kN}$ and $\mathbf{F}_N$ are CBC introduces further simplifications, since the eigenvalues of a CBC matrix correspond to the 2-D DFT of its first row [15]. When such approximations are adopted, it is numerically preferable (and formally equivalent according to Parseval identity), to maximize $\mathcal{K}''_N$ using quantities expressed in the Fourier domain [3], [7], [13].

IV. Asymptotic Behavior of the JLS Solution

In this section, "true" quantities are denoted using a tilde, *e.g.*, $\tilde{o}$ and $\tilde{h}_k$ denote the true object and the true kth PSF, respectively.

An asymptotical study of the behavior of the JLS solution $\hat{\boldsymbol{\alpha}}_N$ needs to refer to a statistical framework. Here, we shall assume that the noise signals $\left\{ n_k(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{Z}^2 \right\}$ are white, centered, identically distributed, of same finite variance σ^2, and uncorrelated: $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}^2$, $\forall\, k, \ell$

$$E\left[n_k(\boldsymbol{x})\right] = 0 \tag{16}$$

$$E\left[n_k(\boldsymbol{x}) n_\ell(\boldsymbol{y})\right] = \sigma^2 \delta_{xy} \delta_{k\ell} \tag{17}$$

where δ is the Kronecker delta symbol: $\delta_{xy} = 1$ if $x = y$, $\delta_{xy} = 0$ otherwise.

By "asymptotical," one could refer to at least three limiting situations

- high-SNR case : $\sigma \to 0$
- highly diverse case : $K \to +\infty$
- large size case : $N \to +\infty$.

This paper focuses on the last case, because it corresponds to a realistic situation (in the sense that N is usually much greater than J). Moreover, the other two cases can be studied in the usual ML framework, since they correspond to situations where the number of unknowns $Q_N^2 + J$ remains constant.

In the framework of *minimum contrast* estimation, one minimizes an objective function $\mathcal{L}_N(\boldsymbol{\alpha})$ that holds the following properties [10]:

C1) as $N \to +\infty$, $\mathcal{L}_N(\boldsymbol{\alpha})$ uniformly converges in probability toward a limiting function $\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$;

C2) $\mathcal{L}$ is a *contrast function relative to* $\tilde{\boldsymbol{\alpha}}$, *i.e.*, its minimum value as a function of $\boldsymbol{\alpha}$ is uniquely attained at $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$.

Under quite general regularity conditions, minimum contrast estimators are *weakly consistent* [10, Section 3.2.3], *i.e.*, the minimizer $\hat{\boldsymbol{\alpha}}_N$ of $\mathcal{L}_N(\boldsymbol{\alpha})$ converges in probability toward $\tilde{\boldsymbol{\alpha}}$

$$\forall\, \varepsilon > 0, \quad \lim_{N \to \infty} \Pr(\|\hat{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\| > \varepsilon) = 0$$

which will be noted $\operatorname{plim}_{N \to \infty} \hat{\boldsymbol{\alpha}}_N = \tilde{\boldsymbol{\alpha}}$. Under additional conditions, one can also establish that $\hat{\boldsymbol{\alpha}}_N$ is asymptotically normally distributed around $\tilde{\boldsymbol{\alpha}}$ with a standard deviation proportional to $1/Q_N$ [10, Section 3.3.4].

Least-square estimation constitutes a fundamental case of contrast estimation. Minimizing $\mathcal{K}_N$ actually falls within the *nonlinear generalized least-square* (NLGLS) approach: $\mathcal{K}_N$ is a quadratic objective function of the data $\boldsymbol{i}_{kN}$, which are nonlinear functions of the unknowns $\boldsymbol{\alpha}$. Moreover, $\mathcal{K}_N$ is not merely a sum of squared residuals, hence the mention *"generalized"*.

Both theory and practice of least-square estimation are well documented, particularly in the field of econometrics. For instance, [16] provides a detailed review of asymptotical statistical properties of least-square estimation. Some contributions address problems (such as estimation in an *errors in variables* model [17]) that are structurally close to phase diversity estimation using the JLS approach. Yet, we have been unable to find directly applicable results to the phase diversity problem. Nonetheless, a tailormade statistical study does seem achievable within the NLGLS framework. In the present paper, we only outline the main conditions that lead to establish consistency.

The most important step is to check that the limiting behavior of $\mathcal{L}_N = \mathcal{K}_N / Q_N^2$ meets Conditions C1 and C2 related to minimum contrast estimation. As the image size increases, two phenomena must be taken into account to establish the limiting expression of $\mathcal{L}_N$.

On the one hand, the effect of approximating the convolution on finite arrays vanishes. This phenomenon can be mathematically studied using Gray's theory of *asymptotically equivalent* matrices [18].

Definition 1: Two series of square matrices $\mathbf{A}_n, \mathbf{B}_n$ of size $n \times n$ are said asymptotically equivalent (which is denoted $\mathbf{A}_n \sim \mathbf{B}_n$) if

- $\mathbf{A}_n, \mathbf{B}_n$ are uniformly bounded in strong norm (*i.e.*, their maximal singular value is uniformly bounded);
- $\lim_{n \to \infty} \|\mathbf{A}_n - \mathbf{B}_n\|_F^2 / n = 0$, where $\|\cdot\|_F$ is the Frobenius norm: $\|\mathbf{A}\|_F^2 = \operatorname{Tr}\{\mathbf{A}^t \mathbf{A}\}$ ($\operatorname{Tr}\{\cdot\}$ is the trace of a square matrix, *i.e.*, the sum of its diagonal elements, which is also the sum of its eigenvalues).

Specifically, important results establish the asymptotical equivalence between Toeplitz and circulant matrices [18], and between TBT and CBC matrices [15], [19].

On the other hand, the random behavior of noise signals is averaged, according to a large numbers effect. Actually, we will also have to consider the true object $\tilde{o}$ from a statistical viewpoint, the latter being considered as a second-order stationary random process.

To simplify the derivations, we only establish the expression of $\mathcal{L} = \lim_{N\to\infty} E[\mathcal{L}_N]$, implicitly assuming that $\mathcal{L}_N$ uniformly converges in probability toward $\mathcal{L}$ under appropriate additional hypotheses (at least, $\tilde{o}$ should be a *correlation-ergodic* stationary random process). In Section V, it is checked by simulations that this conjectured behavior is in good agreement with practice.

Theorem 1: Let us assume the following.

1) The true object $\tilde{o}$ is a second-order stationary random process, centered, with a stable[3] correlation function $\tilde{r}$. Let $\tilde{P}$ denote the power spectral density function of $\tilde{o}$: $\tilde{P}(\boldsymbol{\nu}) = \sum_{x\in Z^2} \tilde{r}(\boldsymbol{x}) e^{-2i\pi\langle \boldsymbol{x},\boldsymbol{\nu}\rangle}$.
2) The noise signals n_k fulfill (16) and (17) and they are uncorrelated with $\tilde{o}$.
3) The cyclic convolution approximation (8) has been adopted: $\mathbf{H}_{kN}\boldsymbol{o}_N = \boldsymbol{h}_{kN} \circledast \boldsymbol{o}_N$.
4) $\forall\, N$, $\mathbf{F}_N$ is CBC and there exists an impulse response $f = \left\{f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{Z}^2\right\}$ such that $\mathbf{F}_N\boldsymbol{o}_N = \boldsymbol{f}_N \circledast \boldsymbol{o}_N$ with $\boldsymbol{f}_N = \{f(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\}$.
5) $\lim_{N\to\infty} \mu_N = \mu$ and $\forall\, \boldsymbol{\nu} \in \Gamma$, $\boldsymbol{\alpha}$

$$D(\boldsymbol{\nu}) = \mu|\rho(\boldsymbol{\nu})|^2 + \sum_k |\gamma_k(\boldsymbol{\nu})|^2 > d > 0$$

where γ_k and ρ are the 2-D discrete time Fourier transforms (DTFT) of h_k and f, respectively

$$\gamma_k(\boldsymbol{\nu}) = \sum_{x\in Z^2} h_k(\boldsymbol{x}) e^{-2i\pi\langle \boldsymbol{x},\boldsymbol{\nu}\rangle}$$
$$\rho(\boldsymbol{\nu}) = \sum_{x\in Z^2} f(\boldsymbol{x}) e^{-2i\pi\langle \boldsymbol{x},\boldsymbol{\nu}\rangle}.$$

Then, $E[\mathcal{L}_N(\boldsymbol{\alpha})] = E[\mathcal{K}_N(\boldsymbol{\alpha})]/Q_N^2$ converges toward $\mathcal{L}(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}}) = \mathcal{L}'(\tilde{\boldsymbol{\alpha}}) - \mathcal{L}''(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}})$, with

$$\mathcal{L}'(\tilde{\boldsymbol{\alpha}}) = K\sigma^2 + \int_\Gamma \tilde{P}(\boldsymbol{\nu}) \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu} \tag{18}$$

$$\mathcal{L}''(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}}) = \int_\Gamma \frac{\tilde{P}(\boldsymbol{\nu})}{D(\boldsymbol{\nu})} \left| \sum_k \gamma_k^*(\boldsymbol{\nu})\tilde{\gamma}_k(\boldsymbol{\nu}) \right|^2 d\boldsymbol{\nu} + \int_\Gamma \frac{\sigma^2}{D(\boldsymbol{\nu})} \sum_k |\gamma_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}. \tag{19}$$

Proof: See Appendix A.

Remark 1: In the first assumption of Theorem 1, it would be more realistic to suppose that the object has a strictly positive mean value $\tilde{m}$. Such a modification introduces the following additional term to $\mathcal{L}(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}})$

$$\mathcal{L}''' = \tilde{m}^2 \sum_k \tilde{\gamma}_k^2(0) + \frac{\tilde{m}^2}{D(0)} \left(\sum_k \gamma_k(0)\tilde{\gamma}_k(0) \right)^2$$

which is a constant term since

$$\gamma_k(0) = \tilde{\gamma}_k(0) = \int_\Gamma A^2(\nu) d\boldsymbol{\nu}$$

does not depend on $\boldsymbol{\alpha}$. For this reason, we maintain $\tilde{m} = 0$ in the rest of the paper.

[3]For example, absolutely summable: $\sum_{\boldsymbol{x}\in Z^2} |\tilde{r}(\boldsymbol{x})| < \infty$.

Remark 2: According to Assumption 3, the range of Theorem 1 is restricted to cyclic convolution, although generalization to alternate boundary conditions seems possible.

Theorem 2: If $\mu = 0$, then $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ minimizes $\mathcal{L}^0(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}}) = \mathcal{L}(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}};\mu = 0)$ and the minimum value is $\mathcal{L}^0(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}) = (K-1)\sigma^2$.

Proof: See Appendix B. ■

Several remarks can be made concerning Theorem 2.

Remark 3: To benefit from the statistical property of Theorem 2, it is required that the regularization term in (11) asymptotically vanish. Alternately, in strict conformity with a Bayesian approach, one could rather choose μ and ρ according to $\forall\, \boldsymbol{\nu}$, $\mu|\rho(\boldsymbol{\nu})|^2 = \sigma^2/\tilde{P}(\boldsymbol{\nu})$. Then, $\hat{\boldsymbol{o}}_N(\boldsymbol{\alpha})$ asymptotically corresponds to the Wiener solution. In this case, the identities of Appendix B yield $\mathcal{L}^W(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}) = K\sigma^2$ and

$$\mathcal{L}^W(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}}) \geq K\sigma^2 + \int_\Gamma \frac{\sigma^2}{D(\boldsymbol{\nu})} \sum_k \left(|\tilde{\gamma}_k(\boldsymbol{\nu})|^2 - |\gamma_k(\boldsymbol{\nu})|^2 \right) d\boldsymbol{\nu}.$$

Unfortunately, this does not allow to conclude that $\mathcal{L}^W(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}})$ is the minimum value of $\mathcal{L}^W(\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}})$.

Remark 4: According to Theorem 2, it seems preferable not to regularize the criterion when the dataset is large enough. This theoretical result meets the conclusions drawn from simulated experiments: In the most favorable situations (such as [7, Fig. 4(b)]), the empirical mean squared error (MSE) $\|\hat{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\|^2$ is an increasing function of μ_N. In such favorable cases, the estimation variance is small, so the MSE is mainly due to bias. In less favorable situations (such as [7, Fig. 4(a)]), penalization also creates bias, but, at the same time, it has a favorable effect on variance. This is a classical situation of bias/variance compromise. In Section V, the same phenomenon is reexamined as a function of the size of the dataset.

Remark 5: If (5) holds without aliasing, then the OTF $\gamma_k(\boldsymbol{\nu})$ necessarily vanishes on the boundaries of the square Γ. As a consequence, Assumption 5 of Theorem 1 holds only if $\mu > 0$, which contradicts the assumption of Theorem 2. Choosing strictly positive, possibly very small, values of μ is a satisfying option in practice. From a more theoretical viewpoint, a possibility to alleviate Assumption 5 is to modify the original least-square criterion. More precisely, let us replace each fidelity-to-data term $\|\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N\|^2$ by a generalized least-square term $\|\mathbf{G}_N(\boldsymbol{i}_{kN} - \mathbf{H}_{kN}\boldsymbol{o}_N)\|^2$, where $\mathbf{G}_N$ is a CBC matrix defined from an impulse response g that cancels high frequencies out (at least those that violate the condition $\sum_k |\gamma_k(\boldsymbol{\nu})|^2 > d > 0$). Incorporating $\mathbf{G}_N$ into our previous calculations, we are led to the following conclusions.

- Expression (13) of criterion $\mathcal{K}$ is still available provided that $\boldsymbol{i}_{kN}$ be replaced by $\mathbf{G}_N\boldsymbol{i}_{kN}$ in (14) and (15). As a particular case, $\mu_N = 0$ provides Löfdahl and Scharmer's solution based on *noise filtering* [20, Section 2.3]. A comparable data filtering procedure is proposed in [21].
- Theorem 2 still holds, provided that

$$\mathcal{L}^0(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}) = (K-1)\sigma^2\|g\|^2.$$

- Let us assume $\mu = 0$. Then Assumption 5 in Theorem 1 can be replaced by

5′). Let β be the DTFT of g. There exists $d > 0$ such that $\forall\, \boldsymbol{\nu} \in \Gamma, \forall\, \boldsymbol{\alpha}$

$$\sum_k |\gamma_k(\boldsymbol{\nu})|^2 \leq d \Longrightarrow \beta(\boldsymbol{\nu}) = 0.$$

Finally, let us seek conditions under which $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ is actually the *unique* minimizer of

$$\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) = (K-1)\sigma^2 + \int_\Gamma \tilde{P}(\boldsymbol{\nu}) \left(\sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 - \frac{\left| \sum_k \gamma_k^*(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu}) \right|^2}{\sum_k |\gamma_k(\boldsymbol{\nu})|^2} \right) d\boldsymbol{\nu}.$$

Any value of $\boldsymbol{\alpha}$ that cancels the integral part of $\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$ is obviously a minimizer, by necessary and sufficient condition. Then, equivalently, $(\gamma_1(\boldsymbol{\nu}), \ldots, \gamma_K(\boldsymbol{\nu}))$ and $(\tilde{\gamma}_1(\boldsymbol{\nu}), \ldots, \tilde{\gamma}_K(\boldsymbol{\nu}))$ are colinear for all $\boldsymbol{\nu}$, *i.e.*,

$$\begin{aligned} &\exists \lambda \text{ such that } \forall\, k, \boldsymbol{\nu}, \quad \gamma_k(\boldsymbol{\nu}) = \lambda(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu}) \\ &\Longleftrightarrow\ \exists g \text{ such that } \forall\, k, \quad h_k = g * \tilde{h}_k. \end{aligned} \tag{20}$$

Let us assume that there exists a filter g such that (20) holds for some $\boldsymbol{\alpha} \neq \tilde{\boldsymbol{\alpha}}$. This means that we are facing a strong identifiability problem: The two solutions $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ and $(\boldsymbol{\alpha}, g^{-1} * \tilde{o})$ are not distinguishable from each other on the basis of the data, whatever the size of measured images and whatever the adopted method of estimation. Such a situation happens if the phase diversity functions θ_k are not appropriately chosen, *e.g.*, [22, Appendix B].

- From only one measured image ($K = 1$), the sign of the symmetric component of ϕ (*i.e.*, ϕ_s with $\phi_s(\boldsymbol{\nu}) = (\phi(\boldsymbol{\nu}) + \phi(-\boldsymbol{\nu}))/2$) is not identifiable.
- The same undeterminacy holds if the phase diversity functions θ_k are chosen antisymmetric ($\theta_k(-\boldsymbol{\nu}) = -\theta_k(\boldsymbol{\nu})$, $\forall\, \boldsymbol{\nu}$). This does not occur in practice since defocus corresponds to $\theta_k(\boldsymbol{\nu}) \propto d_k \|\boldsymbol{\nu}\|^2$, where d_k is the kth defocus distance.
- In any case, the couple $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ is only identifiable up to an arbitrary spatial shift $g(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$ (here δ is the Dirac delta function), *i.e., tilt coefficients* are not identifiable [7].

- From only one measured image ($K = 1$), the sign of the symmetric component of ϕ (*i.e.*, ϕ_s with $\phi_s(\boldsymbol{\nu}) = (\phi(\boldsymbol{\nu}) + \phi(-\boldsymbol{\nu}))/2$) is not identifiable.
- The same undeterminacy holds if the phase diversity functions θ_k are chosen antisymmetric ($\theta_k(-\boldsymbol{\nu}) = -\theta_k(\boldsymbol{\nu})$, $\forall\, \boldsymbol{\nu}$). This does not occur in practice since defocus corresponds to $\theta_k(\boldsymbol{\nu}) \propto d_k \|\boldsymbol{\nu}\|^2$, where d_k is the kth defocus distance.
- In any case, the couple $(\tilde{\boldsymbol{\alpha}}, \tilde{o})$ is only identifiable up to an arbitrary spatial shift $g(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$ (here δ is the Dirac delta function), *i.e., tilt coefficients* are not identifiable [7].

V. Simulation Study

A. Conditions of Simulation

This section proposes an empirical study of the statistical behavior of estimated phase coefficients as a function of the size of the observed images.

TABLE I
VALUES OF ZERNIKE COEFFICIENTS (EXPRESSED IN RADIANS) USED TO SIMULATE THE FOCUSED PSF

j	1	2	3	4	5	6	7
α_j	0	0	0	-0.2	0.3	-0.45	0.4
j	8	9	10	11	12	13	14
α_j	0.3	-0.25	0.35	0.2	0.1	0.05	-0.05
j	15	16	17	18	19	20	21
α_j	0.05	0.02	0.01	-0.01	-0.02	0.01	0.01

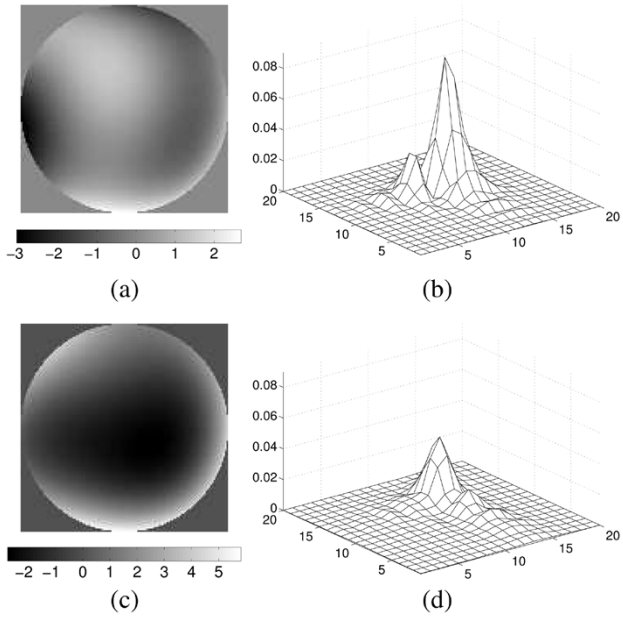


Fig. 1. (a) Aberrated phase $\phi_1 = \phi = \sum_{j=1}^{J} \alpha_j \psi_j$ in radians, where α_j are given in Table I. (b) Central 20×20 part of the resulting PSF h_1 on $\mathbb{Z}^2$. (c) Aberrated phase in radians after defocus $\phi_2 = \phi + \theta_2$. (d) central 20×20 part of the resulting PSF h_2. Both h_1 and h_2 almost vanish outside a central square of 20×20 pixels.

Following [7], we have simulated a focused PSF h_1 using the first 21 Zernike polynomials with coefficients given by Table I [see Fig. 1(a) and (b)]. We have also simulated one defocused PSF h_2 using $\theta_2 = \delta\alpha_4 \psi_4$, with $\delta\alpha_4 = 1.814$ radians [see Fig. 1(c) and (d)].

On the other hand, we have selected two different objects o.

- "`Noise`": The object is a Gaussian white noise sampled on a 512×512 grid.
- "`Earth`": The object is an Earth view sampled on the same 512×512 grid depicted on Fig. 2(a).

Couples of observed images of size 512×512 have been obtained using the approximate model (8), only the central 256×256 part of them being considered afterwards in order to get rid of the effect of cyclic convolution [see Fig. 2(b) and (c), respectively]. Finally, images have been corrupted by realizations of white Gaussian noise with a realistic signal to noise ratio of 100 dB.

In the sequel, $\boldsymbol{\alpha} = [\alpha_4, \ldots, \alpha_{21}]$, *i.e.*, the first three coefficients have not been estimated.

- The *piston* coefficient α_1 is a constant added to the phase and has no influence on the shape of the PSF.
- The *tilt* coefficients α_2, α_3 introduce a shift in the image that is of no importance for extended object.

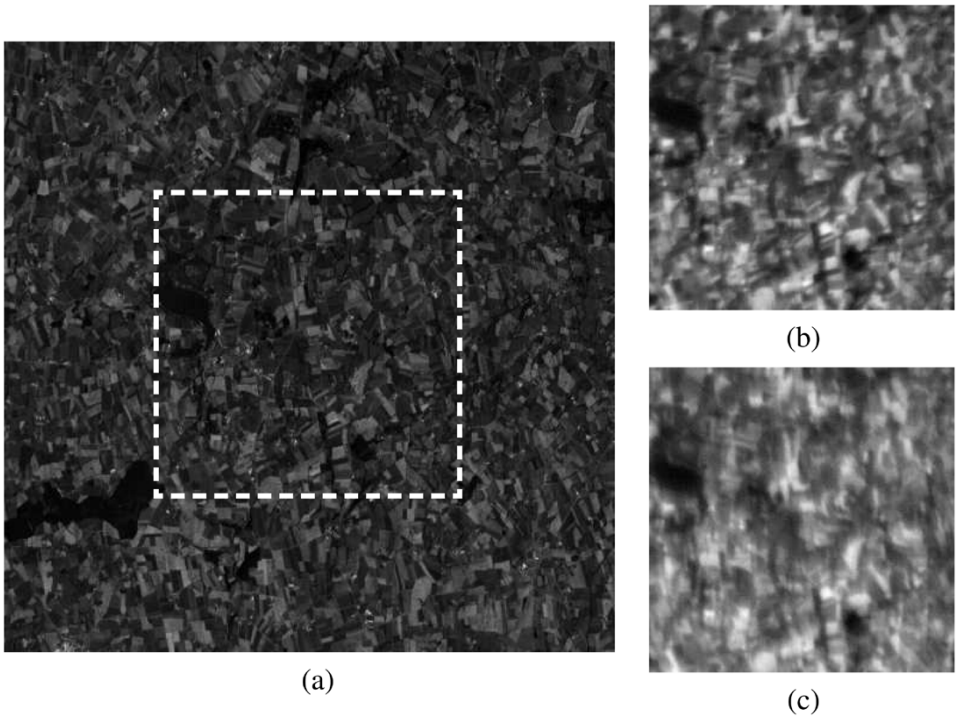


Fig. 2. (a) The 512 × 512 Earth view used to simulate the extended object called "Earth." The original source is an image taken by satellite SPOT3, which has been downsampled by a factor two in both directions—Copyright CNES/Distribution SPOT IMAGE. (b) Central 256 × 256 part of noiseless focused image. (c) Central 256 × 256 part of noiseless defocused image.

By mean of Monte-Carlo simulations using several independent couples of realizations of noise, it is possible to evaluate the statistical performance of $\hat{\boldsymbol{\alpha}}_N$ for different values of N. Here, we have evaluated and displayed the following quantities for $R = 50$ independent realizations of noise.

- The Euclidian norm of the empirical bias vector (henceforth referred as the bias of $\hat{\boldsymbol{\alpha}}_N$, for sake of brevity): $b_N \triangleq \|\check{\boldsymbol{\alpha}}_N - \tilde{\boldsymbol{\alpha}}\|$ where $\check{\boldsymbol{\alpha}}_N = \sum_{r=1}^{R} \hat{\boldsymbol{\alpha}}_N^{(r)}/R$.
- The Euclidian norm of the vector of empirical standard deviations (henceforth referred as the standard deviation of $\hat{\boldsymbol{\alpha}}_N$)

$$\sigma_N \triangleq \left(\frac{1}{R} \sum_{r=1}^{R} \|\hat{\boldsymbol{\alpha}}_N^{(r)} - \check{\boldsymbol{\alpha}}_N\|^2 \right)^{(1/2)}.$$

- The square root of the empirical MSE (henceforth referred as RMSE) $r_N \triangleq \sqrt{b_N^2 + \sigma_N^2}$.

B. "Noise" Object

Fig. 3(a)–(c) deal with the "Noise" object. They, respectively, depict the bias b_N, standard deviation σ_N and RMSE r_N of $\hat{\boldsymbol{\alpha}}_N$ as functions of the regularization parameter μ_N. In the "Noise" case, maximization of (15) has been considered with $\mathbf{F}_N$ equal to the identity matrix and under the usual CBC approximation for $\mathbf{H}_{1N}$ and $\mathbf{H}_{2N}$. Three nested images have been tested, of size $Q_N \times Q_N$, with $Q_N \in \{64, 128, 256\}$.

On the one hand, processing a larger image appears favorable in terms of bias [Fig. 3(a)]. The reason is that the effect of the CBC approximation becomes negligible for large size images. However, the relative improvement is more substantial for small values of μ_N. This empirical observation fully meets the conclusions of Section IV: As Q_N grows, a vanishing series of μ_N is required to get an asymptotically unbiased estimator $\hat{\boldsymbol{\alpha}}_N$.

On the other hand, the standard deviation of $\hat{\boldsymbol{\alpha}}_N$ is a decreasing function of μ_N [Fig. 3(b)]. It also decreases with Q_N, and it is important to notice that the corresponding decreasing rate is rather independent of μ_N. This is not surprising since random fluctuations are averaged whatever the value of μ_N.

As a global consequence in terms of bias/variance compromise, the minimizer of the RMSE r_N shifts leftward as Q_N grows [Fig. 3(c)]. The minimum value is $\min_{\mu_N} r_N = 0.2116$, 0.1052, and 0.0476 for $Q_N = 64$, 128, and 256, respectively: It roughly decreases proportionaly to $1/Q_N$, although other values of Q_N should be tested to assess the actual decrease rate.

C. "Earth" Object

Fig. 4(a)–(c) deal with the "Earth" object. Maximization of (15) has been considered in the same conditions as in Section V-B, provided that $\mathbf{F}_N$ has been deduced from a power spectral density model with parameters fitted using the true object (see [7, Eq. (13)]).

Results depicted on Figs. 3 and 4 are comparable, except that the bias reaches much larger values in the present case, even for the largest size of image. This is a consequence of *edge effects* due to the adopted cyclic convolution approximation in the presence of extended, structured objects: Nonrealistic sizes of images should be processed to get statistically meaningfull estimates of $\boldsymbol{\alpha}$.

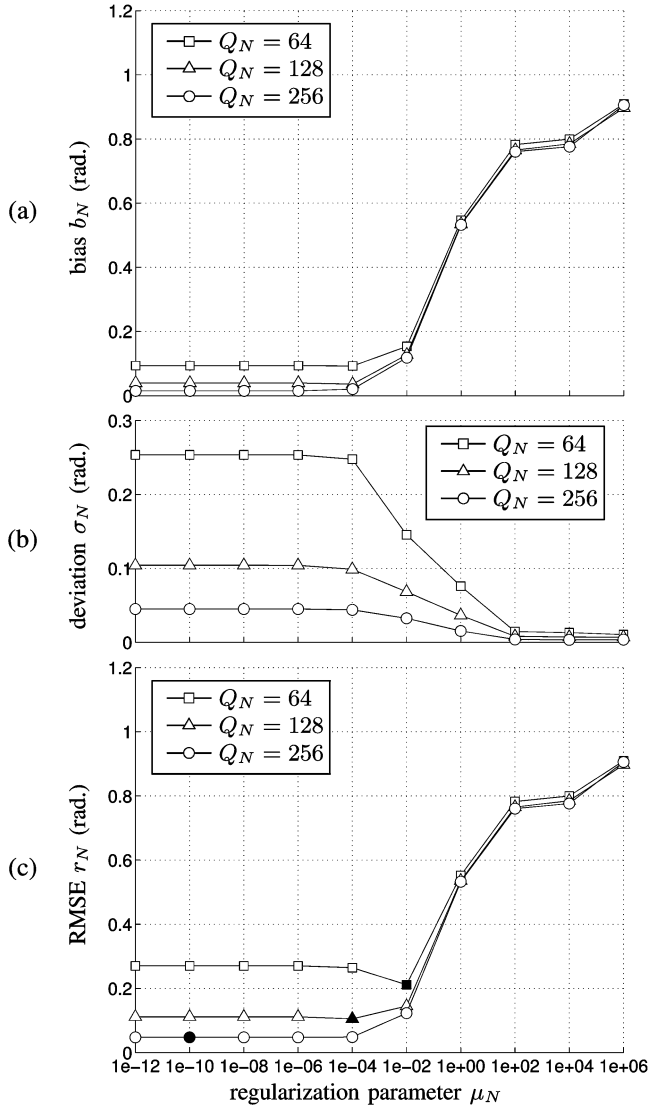


Fig. 3. (a) Bias b_N. (b) Standard deviation σ_N. (c) RMSE r_N of $\hat{\boldsymbol{\alpha}}_N$ in the "Noise" case, *i.e.*, the true object is the realization of a 2-D Gaussian white noise, of size $Q_N \times Q_N$. For the RMSE, the minimum value of each curve is indicated by a blackened symbol.

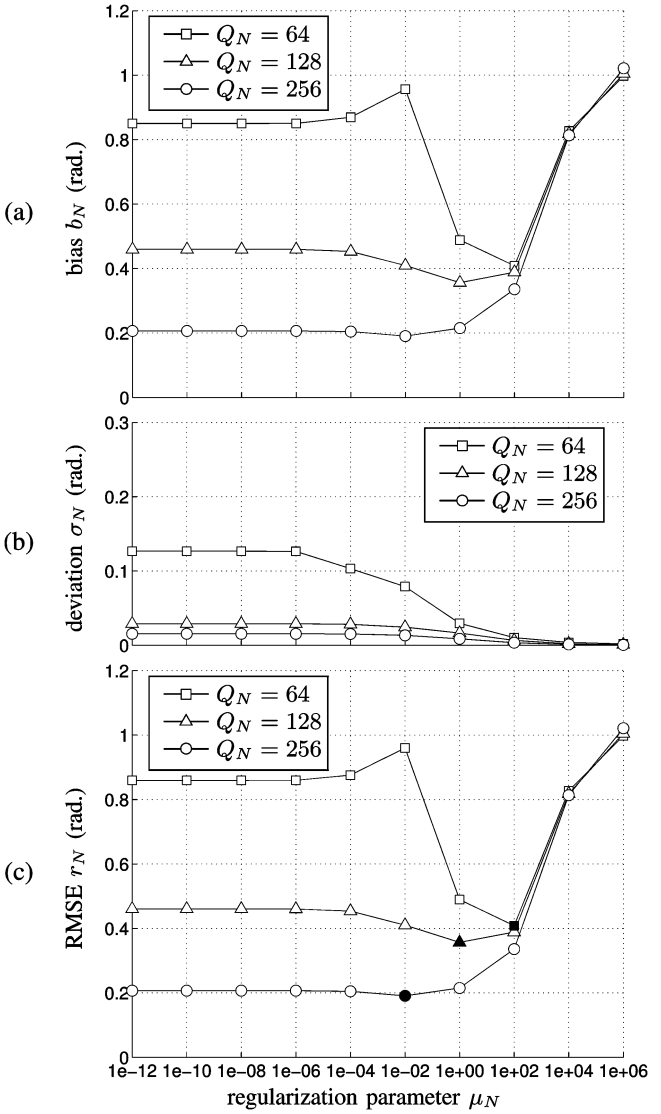


Fig. 4. Same as Fig. 3, except that the true object is the Earth view depicted at Fig. 2(a).

To overcome this difficulty, Löfdahl and Scharmer introduced a *tapering* technique [20],[4] where:

- the observed images are windowed in order to apodize the edges;
- in the least-square criterion (11), the fidelity-to-data term is modified: The norm is only considered over a central, nonapodized part of the images.

We have applied this technique to the data simulated in the "Earth" case, using a 2-D modified Hamming apodization window with a central plateau of size $Q_N/2 \times Q_N/2$. Fig. 5(a) [compared to Fig. 4(a)] shows an impressive effect on bias. Even if the variance slightly increases, at least for large size images [Fig. 5(b) compared to Fig. 4(b)], the overall effect is largely favorable in term of RMSE [Fig. 5(c) compared to Fig. 4(c)].

Our theoretical study has been derived in the case where the images are not apodized by Löfdahl and Scharmer's tapering technique. However, given the results depicted on Fig. 5, it seems reasonable to expect that consistency results still hold, provided that the size of the apodized edges vanishes as Q_N grows.

[4]See also [23] for an alternative technique based on the use of a guard-band.

VI. Conclusion

In this paper, we have studied some important statistical properties of the phase diversity technique introduced by Gonsalves [1], [2]. In particular, it has been shown that Gonsalves' technique is a *minimum contrast method*, with respect to phase estimation. As a consequence, it provided a consistent phase estimator as the size of the processed images grows (putting aside practical and physical limitations). No comparable result is yielded by the classical *joint ML* interpretation (*e.g.*, as found in [3]). In particular, the Gaussian character of the noise is not a prerequisite in our convergence study.

By simulation, we have checked that the JLS method behaves as predicted by theory in the case of extended objects. We have also observed that the edge effects due to cyclic convolution introduce a strong bias on phase estimation, that only slowly diminishes as the image size grows. Modified versions of the

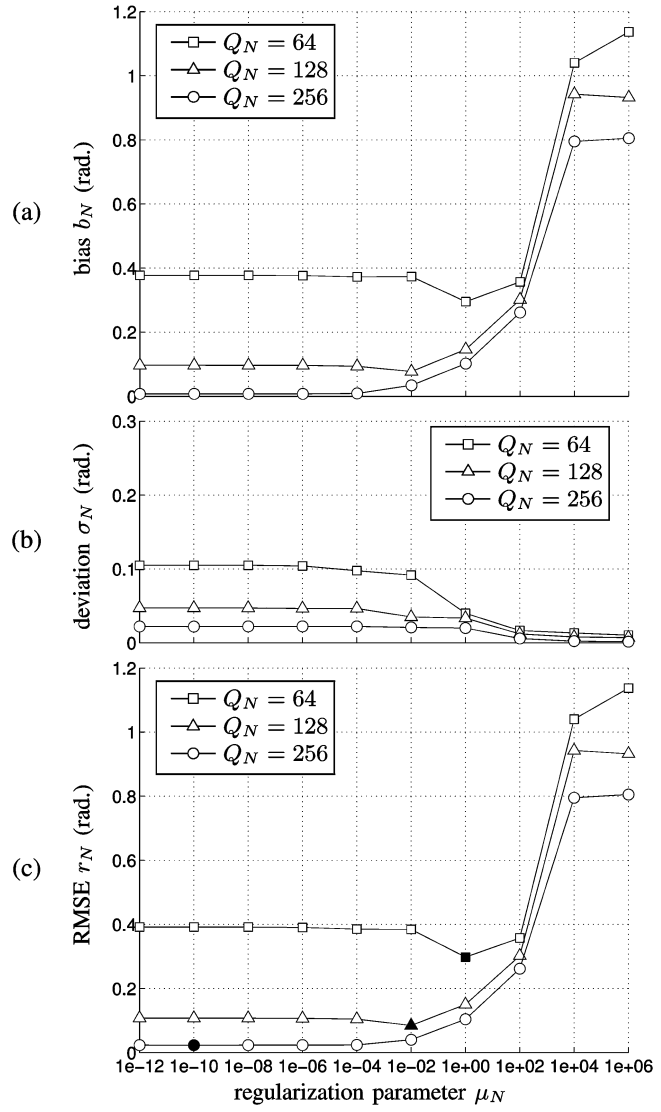


Fig. 5. Same as Fig. 4, except that observed images have been apodized using the tapering technique introduced in [20].

JLS method are then required to recover meaningful estimates. We have more specifically considered the tapering technique proposed in [20], and we have empirically verified that the latter technique is still statistically convergent.

Finally, some alternative error metrics have been introduced to replace the criterion $\mathcal{K}_N$ induced by the JLS approach [24], [25], for the sake of faster computations. An interesting perspective would be to study if such alternative error metrics are still minimum contrast functions.

APPENDIX A

Proof of Theorem 1

Let us decompose the observed images according to $\boldsymbol{i}_{kN} = \boldsymbol{y}_{kN} + \boldsymbol{n}_{kN}$, where $\boldsymbol{y}_{kN} = \left\{(\tilde{h}_k * \tilde{o})(\boldsymbol{x}), \boldsymbol{x} \in \Omega_N\right\}$. Then, we have

$$\mathcal{K}'_N = \sum_k ||\boldsymbol{i}_{kN}||^2 = \sum_k \left(||\boldsymbol{n}_{kN}||^2 + ||\boldsymbol{y}_{kN}||^2 + 2\boldsymbol{y}_{kN}^t \boldsymbol{n}_{kN}\right).$$

Hence

$$E\left[\mathcal{L}'_N\right] = Q_N^{-2} E\left[\mathcal{K}'_N\right] = K\sigma^2 + Q_N^{-2} \sum_k E\left[||\boldsymbol{y}_{kN}||^2\right]$$

according to Assumptions 1 and 2. The second term converges toward the second term of (18), according to Parseval identity. As a consequence, we have $\lim_{N\to\infty} E\left[\mathcal{L}'_N\right] = \mathcal{L}'$.

The main part of the proof is to express $\lim_{N\to\infty} E\left[\mathcal{L}''_N\right]$, where $\mathcal{L}''_N = Q_N^{-2}\mathcal{K}''_N$. In particular, we need to examine the asymptotical behavior of matrices $\mathbf{H}_{kN}$. This is done in the following technical lemma.

Lemma 1: For all k, we have $\mathbf{H}_{kN} \sim \bar{\mathbf{H}}_{kN}$, where $\bar{\mathbf{H}}_{kN}$ is the $Q_N^2 \times Q_N^2$ CBC matrix whose eigenvalues are equally distributed on γ_k, *i.e.*, $\{\gamma_k(\boldsymbol{x}/Q_N), \boldsymbol{x} \in \Omega_N\}$.

Proof: Let us introduce the DTFT of h_{kN}: $\forall\, \boldsymbol{\nu} \in \mathbb{R}^2$, $k = 1, \ldots, K$

$$\gamma_{kN}(\boldsymbol{\nu}) = \sum_{x\in\Omega_N} h_{kN}(\boldsymbol{x}) e^{-2i\pi\langle\boldsymbol{\nu},\boldsymbol{x}\rangle}$$

so that $\mathbf{H}_{kN}$ is the CBC matrix whose eigenvalues are equally distributed on γ_{kN}. According to (9) and (10), we have

$$\gamma_{kN}(\boldsymbol{\nu}) = Q_N^{-2} \sum_{y\in\Omega_N} Z_k^*\left(\boldsymbol{\nu} + \frac{\boldsymbol{y}}{Q_N}\right) Z_k(\boldsymbol{\nu}) \qquad (21)$$

where Z_k is extended over $\mathbb{R}^2$ in a periodic manner: $\forall\, \boldsymbol{\nu} \in \Gamma$, $\boldsymbol{x} \in \mathbb{Z}^2$, $Z_k(\boldsymbol{\nu} + \boldsymbol{x}) = Z_k(\boldsymbol{\nu})$. On the other hand, (6) and (7) also yield

$$\gamma_k(\boldsymbol{\nu}) = \int_\Gamma Z_k^*(\boldsymbol{\nu} + \boldsymbol{\eta}) Z_k(\boldsymbol{\eta}) d\boldsymbol{\eta}. \qquad (22)$$

It is clear from (21) and (22) that $\mathbf{H}_{kN}$ and $\bar{\mathbf{H}}_{kN}$ are uniformly bounded by $\max_{\boldsymbol{\nu}\in\Gamma} A^2(\boldsymbol{\nu})$. It is also clear from the same equations that $\gamma_{kN}(\boldsymbol{\nu})$ is a Rieman sum that uniformly converges toward $\gamma_k(\boldsymbol{\nu})$ when N grows to infinity

$$\forall\, \varepsilon > 0,\ \exists N_0 \text{ such that } \forall\, \boldsymbol{\nu},\ N \geq N_0, \quad |\gamma_{kN}(\boldsymbol{\nu}) - \gamma_k(\boldsymbol{\nu})| \leq \varepsilon.$$

As a consequence, we have

$$||\mathbf{H}_{kN} - \bar{\mathbf{H}}_{kN}||_F^2 = \sum_{x\in\Omega_N} \left|\gamma_{kN}\left(\frac{\boldsymbol{x}}{Q_N}\right) - \gamma_k\left(\frac{\boldsymbol{x}}{Q_N}\right)\right|^2 \leq Q_N^2 \varepsilon^2$$

for any ε, provided that N is large enough. Hence, $\lim_{N\to\infty} ||\mathbf{H}_{kN} - \bar{\mathbf{H}}_{kN}||_F^2 / Q_N^2 = 0$. ■

Let $\mathbf{U}_{k\ell N} = \mathbf{H}_{kN}\mathbf{R}_N^+\mathbf{H}_{\ell N}^t$ and $\mathbf{S}_{k\ell N} = E\left[\boldsymbol{i}_{kN}\boldsymbol{i}_{\ell N}^t\right] = E\left[\boldsymbol{y}_{kN}\boldsymbol{y}_{\ell N}^t\right] + \sigma^2\mathbf{I}_N$ (where $\mathbf{I}_N$ is the $Q_N^2 \times Q_N^2$ identity matrix). Given (15), we have[5]

$$E\left[\mathcal{K}''_N\right] = \sum_{k,\ell} E\left[\boldsymbol{i}_{kN}^t \mathbf{U}_{k\ell N} \boldsymbol{i}_{\ell N}\right] = \sum_{k,\ell} \mathrm{Tr}\left\{\mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N}\right\}.$$

According to Lemma 1 and to [18, Theorem 2.1], each matrix $\mathbf{U}_{k\ell N}$ is asymptotically equivalent to another CBC matrix, whose eigenvalues are equally distributed on

$$U_{k\ell}(\boldsymbol{\nu}) = \frac{\gamma_k^*(\boldsymbol{\nu})\gamma_\ell(\boldsymbol{\nu})}{D(\boldsymbol{\nu})}, \quad \boldsymbol{\nu} \in \Gamma.$$

Let us remark that Assumption 5 is needed here to ensure that $\mathbf{R}_N^+$ is uniformly bounded in strong norm.

[5]Recall that $\mathrm{Tr}\{\mathbf{A}^t\mathbf{B}\} = \mathrm{Tr}\{\mathbf{B}\mathbf{A}^t\}$ for any two matrices $\mathbf{A}$, $\mathbf{B}$ of same dimensions.

On the other hand, $\mathbf{S}_{k\ell N}$ is an intercorrelation matrix, which is TBT since $\tilde{o}$ is stationary. Moreover, $\mathbf{S}_{k\ell N}$ is generated by the crosscorrelation sequence $\tilde{s}_{kl}$ defined by

$$\tilde{s}_{kl}(\boldsymbol{x}) = \left(\tilde{h}_k^{(-)} * \tilde{h}_\ell * \tilde{r}\right)(\boldsymbol{x}) + \sigma^2 \delta_{k\ell} \delta_{x\mathbf{0}}$$

where $\tilde{h}_k^{(-)}$ is the reversed version of $\tilde{h}_k$: $\tilde{h}_k^{(-)}(\boldsymbol{x}) = \tilde{h}_k(-\boldsymbol{x})$. The sequence $\tilde{s}_{kl}$ is stable for all k, ℓ since

- $\tilde{r}$ is stable by assumption;
- $\forall\ k$, $\tilde{h}_k$ is stable since given (6) and (7), we have the following Parseval identity

$$\sum_{x \in Z^2} |\tilde{h}_k(\boldsymbol{x})| = \sum_{x \in Z^2} |\tilde{z}_k(\boldsymbol{x})|^2 = \int_\Gamma A^2(\boldsymbol{\nu}) d\boldsymbol{\nu}$$

- convolution products and sums of stable functions are stable.

Thus, according to [18, Lemma 4.6][6], $\mathbf{S}_{k\ell N}$ is asymptotically equivalent to a CBC matrix whose eigenvalues are equally distributed on $S_{k\ell}(\boldsymbol{\nu})$, $\boldsymbol{\nu} \in \Gamma$, where

$$S_{k\ell}(\boldsymbol{\nu}) = \tilde{\gamma}_k^*(\boldsymbol{\nu}) \tilde{\gamma}_\ell(\boldsymbol{\nu}) \tilde{P}(\boldsymbol{\nu}) + \sigma^2 \delta_{k\ell}.$$

Finally, according to [18, Theorem 2.1], the matrix product $\mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N}$ is also asymptotically CBC, with eigenvalues equally distributed on $S_{k\ell}^*(\boldsymbol{\nu}) U_{k\ell}(\boldsymbol{\nu})$, so we get the following converging Riemann sum

$$\begin{aligned}\lim_{N \to \infty} Q_N^{-2} \sum_{k,\ell} \mathrm{Tr}\left\{\mathbf{S}_{k\ell N}^t \mathbf{U}_{k\ell N}\right\} &= \int_\Gamma \sum_{k,\ell} S_{k\ell}^*(\boldsymbol{\nu}) U_{k\ell}(\boldsymbol{\nu}) d\boldsymbol{\nu} \\ &= \mathcal{L}''(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}).\end{aligned}$$

APPENDIX B

Proof of Theorem 2

Evaluation of $\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}})$ is straightforward

$$\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = K\sigma^2 + \int_\Gamma \frac{\mu |\rho(\boldsymbol{\nu})|^2 \tilde{P}(\boldsymbol{\nu}) - \sigma^2}{\tilde{D}(\boldsymbol{\nu})} \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 d\boldsymbol{\nu}$$

which takes an extremely simple form when $\mu = 0$

$$\mathcal{L}^0(\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\alpha}}) = (K-1)\sigma^2.$$

Furthermore, let us show that $(K-1)\sigma^2$ is actually the minimum value of $\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}})$. By Cauchy–Schwartz inequality

$$\forall\ \boldsymbol{\nu} \in \Gamma, \quad \left|\sum_k \gamma_k^*(\boldsymbol{\nu}) \tilde{\gamma}_k(\boldsymbol{\nu})\right|^2 \leq \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 \sum_k |\gamma_k(\boldsymbol{\nu})|^2$$

which allows to deduce from (18) and (19) that $\forall\ \boldsymbol{\alpha}$

$$\begin{aligned}\mathcal{L}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) \geq\ & (K-1)\sigma^2 \\ & + \mu \int_\Gamma \frac{|\rho(\boldsymbol{\nu})|^2}{D(\boldsymbol{\nu})} \left(\tilde{P}(\boldsymbol{\nu}) \sum_k |\tilde{\gamma}_k(\boldsymbol{\nu})|^2 + \sigma^2\right) d\boldsymbol{\nu}.\end{aligned}$$

When $\mu = 0$, the latter inequality drastically simplifies according to

$$\mathcal{L}^0(\boldsymbol{\alpha}, \tilde{\boldsymbol{\alpha}}) \geq (K-1)\sigma^2$$

which proves the assertion.

[6]For sake of correctness, Gray's asymptotical result only applies to Toeplitz, not necessarily Hermitian, matrices. In extensions to TBT matrices found in [15], [19], only the Hermitian case is considered. Here, we shall admit that Gray's result extends to TBT matrices, not necessarily Hermitian.

REFERENCES

[1] R. A. Gonsalves and R. Chidlaw, "Wavefront sensing by phase retrieval," in *Proc. SPIE Applications of Digital Image Processing III*, vol. 207, A. Tescher, Ed., 1979, pp. 32–39.

[2] R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.*, vol. 21, no. 5, pp. 829–832, 1982.

[3] R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations using phase diversity," *J. Opt. Soc. Amer. A*, vol. 9, no. 7, pp. 1072–1085, 1992.

[4] R. G. Paxman, B. J. Thelen, and J. H. Seldin, "Phase-diversity correction of turbulence-induced space-variant blur," *Opt. Lett.*, vol. 19, no. 16, pp. 1231–1233, Aug. 1994.

[5] R. J. A. Little and D. B. Rubin, "On jointly estimating parameters and missing data by maximizing the complete-data likelihood," *Amer. Stat.*, vol. 37, pp. 218–220, Aug. 1983.

[6] A. Blanc, J. Idier, and L. Mugnier, "Novel estimator for the aberrations of a space telescope by phase diversity," in *Proc. SPIE Int. Symp. Astronomical Telescopes and Instrumentation*, Munich, Germany, Mar. 2000, pp. 728–736.

[7] A. Blanc, L. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," *J. Opt. Soc. Amer. A*, vol. 20, no. 6, pp. 1035–1045, Jun. 2003.

[8] J. R. Fienup, J. Marron, T. J. Schulz, and J. H. Seldin, "Hubble space telescope characterized by using phase-retrieval algorithms," *Appl. Opt.*, vol. 32, pp. 1747–1767, 1993.

[9] D. J. Lee, M. C. Roggemann, and B. M. Welsh, "Cramer-Rao analysis of phase-diverse wave-front sensing," *J. Opt. Soc. Amer.*, vol. 16, no. 5, pp. 1005–1015, 1999.

[10] D. Dacunha-Castelle and M. Duflo, *Probabilités et Statistiques, 2. Problèmes à Temps Mobile*, 1st ed. Paris, France: Masson, 1983.

[11] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Amer.*, vol. 66, no. 3, pp. 207–211, 1976.

[12] B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," *IEEE Trans. Comput.*, vol. C-22, no. 8, pp. 805–812, Aug. 1973.

[13] C. R. Vogel, T. Chan, and R. Plemmons, "Fast algorithms for phase-diversity-based blind deconvolution," in *Proc. SPIE Adaptative Optical System Technologies*, vol. 3353, D. Bonaccini and R. K. Tyson, Eds., 1998, pp. 994–1005.

[14] M. Z. Nashed, *Generalized Inverses and Applications*. New York: Academic, 1976.

[15] P. A. Voois, "A theorem on the asymptotic eigenvalue distribution of Toeplitz-block-Toeplitz matrices," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 1837–1841, Jul. 1996.

[16] T. Amemiya, "Non-linear regression models," in *Handbook of Econometrics*, Z. Griliches and M. D. Intriligator, Eds. Amsterdam, The Netherlands, 1983, vol. 1, ch. 6, pp. 333–389.

[17] L. J. Gleser, "Estimation in a multivariate errors in variables regression model: Large sample results," *Ann. Stat.*, vol. 9, no. 1, pp. 24–44, 1981.

[18] R. M. Gray. (2002) Toeplitz and circulant matrices: A review. Inf. Theory Lab., Stanford Univ., Stanford, CA. [Online]. Available: http://www-ee.stanford.edu/~gray/toeplitz.html

[19] N. K. Bose and K. J. Boo, "Asymptotic eigenvalue distribution of block-Toeplitz matrices," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 858–961, Mar. 1998.

[20] M. G. Löfdahl and G. B. Scharmer, "Wavefront sensing and image restoration from focused and defocused solar images," *Astron. Astrophys.*, vol. 107, pp. 243–264, 1994.

[21] D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wavefront sensing," *Appl. Opt.*, pp. 9186–9197, Dec. 1997.

[22] A. Blanc, "Identification de Réponse Impulsionnelle et Restauration D'images: Apports de la Diversité de Phase," Ph.D. dissertation, Université de Paris-Sud, Orsay, France, 2002.

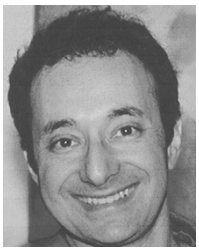
[23] J. H. Seldin and R. G. Paxman, "Phase-diverse speckle reconstruction of solar data," in *Proc. SPIE Image Reconstruction and Restoration*, vol. 2302, T. Schulz and D. Snyder, Eds., 1994, pp. 268–280.

[24] R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," *Appl. Opt.*, vol. 33, no. 27, pp. 6533–6546, 1994.

[25] M. G. Löfdahl and G. B. Scharmer, "A predictor approach to closed-loop phase-diversity wavefront sensing," in *Proc. SPIE 4013, UV, Optical, and IR Telescopes and Instrumentation VI*, Munich, Germany, 2000, pp. 737–748.

Jérôme Idier was born in France in 1966. He received the diploma degree in electrical engineering from the École Supérieure d'Électricité, Gif-sur-Yvette, France, in 1988, and the Ph.D. degree in physics from the University of Paris-Sud, Orsay, France, in 1991.

In 1991, he joined the Centre National de la Recherche Scientifique. He is currently with the Institut de Recherche en Communications et Cybernétique, Nantes, France. His major scientific interests are in probabilistic approaches to inverse problems for signal and image processing.



Laurent Mugnier graduated from the Ecole Polytechnique, France, in 1988, and received the Ph.D. degree from the Ecole Nationale Supérieure des Télécommunications (ENST), France, in 1992, for his work on the digital reconstruction of incoherent-light holograms.

In 1994, he joined ONERA, where he is currently a Senior Research Scientist in the field of inverse problems and high-resolution optical imaging. His current research interests include wavefront sensing, image restoration, and image reconstruction, in particular, for adaptive-optics corrected imaging through turbulence, for Earth observation, and for optical interferometry. His publications include a chapter in a reference book on inverse problems, 20 papers in peer-reviewed international journals, and more than 40 papers in conference proceedings.

Amandine Blanc received the Ph.D. degree in image processing from the University of Paris-Sud, Orsay, France, in 2002.

She is currently with the Laboratoire d'Astrophysique de l'Observatoire de Grenoble, Grenoble, France. Her main subjects of interest are image restoration, wavefront sensing, and adaptive optics.

Annexe K

Article Mugnier et coll., 2004 [A27]

Mugnier *et al.* Vol. 21, No. 10/October 2004/J. Opt. Soc. Am. A 1841

MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images

Laurent M. Mugnier, Thierry Fusco, and Jean-Marc Conan

Office National d'Études et de Recherches Aérospatiales, Optics Department, B.P. 72, F-92322 Châtillon cedex, France

Received August 8, 2003; revised manuscript received April 14, 2004; accepted April 15, 2004

Deconvolution is a necessary tool for the exploitation of a number of imaging instruments. We describe a deconvolution method developed in a Bayesian framework in the context of imaging through turbulence with adaptive optics. This method uses a noise model that accounts for both photonic and detector noises. It additionally contains a regularization term that is appropriate for objects that are a mix of sharp edges and smooth areas. Finally, it reckons with an imperfect knowledge of the point-spread function (PSF) by estimating the PSF jointly with the object under soft constraints rather than blindly (i.e., without constraints). These constraints are designed to embody our knowledge of the PSF. The implementation of this method is called MISTRAL. It is validated by simulations, and its effectiveness is illustrated by deconvolution results on experimental data taken on various adaptive optics systems and telescopes. Some of these deconvolutions have already been used to derive published astrophysical interpretations.

OCIS codes: 100.1830, 100.3020, 100.3190, 010.1080, 010.1330, 110.6770.

1. INTRODUCTION

The performance of high-resolution imaging with large astronomical telescopes is severely limited by atmospheric turbulence. Adaptive optics[1–3] (AO) offers real-time compensation of the turbulence. The correction is, however, only partial,[2,4–7] and the long-exposure images must be deconvolved to restore the fine details of the object.

Because of the inevitable noise in the images, great care must be taken in the deconvolution process to obtain a reliable restoration with good photometric precision. A key point is to recognize that noise makes it necessary to add some prior knowledge on the observed object into the deconvolution method; failure to do so usually results in unacceptable amplification of the noise.[8,9] Additionally, fine modeling of the noise statistics contributes to the accurate restoration of objects with a high dynamic range. Finally, the fact that the residual point-spread function (PSF) is usually not perfectly known[10,11] adds to the difficulty.

This paper presents a deconvolution method that falls within the maximum *a posteriori* (MAP) framework, or, equivalently, the penalized-likelihood framework, and that addresses these three points. It uses a prior that is well adapted to astronomical objects that are a mix of sharp structures and smooth areas, such as planets and asteroids; for pointlike objects such as binary stars, an alternative and more appropriate prior can be used. This method takes into account the presence of a mixture of photon and electronic noises. It also estimates the PSF given some prior information on the average PSF and its variability. The implementation of this method is called MISTRAL (for Myopic Iterative STep-preserving Restoration ALgorithm). Although it is presented in the context of long-exposure images recorded on AO-corrected telescopes, this method can be used in other contexts as well. In particular, it has already been successfully used for Hubble Space Telescope data.[12]

2. IMAGING MODEL AND PROBLEM STATEMENT

Within the isoplanatic angle, defined as the size of the angular patch in which the PSF due to turbulence can be considered constant, the image i of the observed object o at the focal plane of the system consisting of the atmosphere, the telescope, the AO system and the detector is given by

$$i = [h * o] \diamond n, \tag{1}$$

where $*$ denotes the convolution operator and $[\cdot]$ the sampling operator, h is the PSF of the system, n is a corruptive noise process (often predominantly photon noise), and the symbol $\diamond$ represents a pixel-by-pixel operation.[8] If the noise is additive and independent of the noiseless image $[h*o]$, then the symbol $\diamond$ simply represents addition.

In the following sections we shall consider that the object and the image are sampled on a regular grid, yielding a vectorial formulation for Eq. (1):

$$\mathbf{i} = (\mathbf{h} * \mathbf{o}) \diamond \mathbf{n}, \tag{2}$$

where $\mathbf{o}$, $\mathbf{i}$, and $\mathbf{n}$ are the vectors corresponding to the lexicographically ordered object, image, and noise, respectively.

1084-7529/2004/101841-14$15.00

Note that the raw image must be carefully preprocessed to yield an image that closely follows this imaging model. The preprocessing includes at least the correction of the background and of the flat field, the correction of the camera's bad pixels and possibly of its correlated noise, and the scaling of the image in photons. In the case of background-dominated (e.g., thermal infrared) images, one must often record tens of images to avoid detector saturation. The preprocessing then includes the (relative) recentering of these images[13] and their addition prior to deconvolution. One could contemplate processing these images jointly rather than co-adding them. Yet for AO-corrected images this would be costly in computation time while bringing very little additional information, because all PSFs are essentially the same as soon as the exposure time is long with respect to the evolution time scale of turbulence.

The deconvolution procedure needs a measurement of the PSF. The usual procedure consists in recording the corrected image of a nearby unresolved star shortly before and/or after observing the object of interest. Since the correction quality depends on the observing conditions (turbulence strength, wind speed, magnitude, and spatial extent of the source used for wave-front sensing), the unresolved star image is not a perfect measurement of the PSF associated with the image to be deconvolved.[10] A more precise estimate of the PSF can be obtained via control-loop data accumulated during the acquisition of the object of interest.[14] The PSF estimated in this fashion still has a limited precision owing to noise and corresponds to an infinite integration time, so it is also imperfect. Furthermore, it does not intrinsically include the effect of static aberrations, even though these can be calibrated either by using the image of a star and its control-loop data[14] or by a dedicated setup of the instrument.[15,16]

The problem is to obtain an estimate $\hat{\mathbf{o}}$ of the observed object $\mathbf{o}$ given the image $\mathbf{i}$, a more-or-less precise knowledge of the PSF $\mathbf{h}$, and some prior information on the noise statistics and on the object.

3. DECONVOLUTION APPROACH

A. Deconvolution with Known Point-Spread Function

We first address the classical case in which the PSF is assumed to be known. Most deconvolution techniques boil down to the minimization (or maximization) of a criterion. The first issue is the definition of a suitable criterion for the given inverse problem. The second issue is then to find the position of the criterion's global minimum, which is defined as the solution. This second issue will be addressed in Section 4.

Following the probabilistic (Bayesian) MAP approach, the deconvolution problem can be stated as follows: We look for the most likely object $\hat{\mathbf{o}}$ given the observed image $\mathbf{i}$ and our prior information on $\mathbf{o}$, which is summarized by a probability density $p(\mathbf{o})$. This reads as

$$\hat{\mathbf{o}} = \arg\max_{\mathbf{o}} p(\mathbf{o}|\mathbf{i}) = \arg\max_{\mathbf{o}} p(\mathbf{i}|\mathbf{o}) \times p(\mathbf{o}). \quad (3)$$

Equivalently, $\hat{\mathbf{o}}$ can be defined as the object that minimizes a compound criterion $J(\mathbf{o})$ defined as follows:

$$J(\mathbf{o}) = J_i(\mathbf{o}) + J_o(\mathbf{o}), \quad (4)$$

where the negative log-likelihood $J_i = -\ln p(\mathbf{i}|\mathbf{o})$ is a measure of fidelity to the data and $J_o = -\ln p(\mathbf{o})$ is a regularization or penalty term, so the MAP solution can equivalently be called a penalized-likelihood solution. Note that the Bayesian approach does not require that $\mathbf{o}$ truly be the outcome of a stochastic process; rather, $p(\mathbf{o})$ should be designed to embody the available prior information on $\mathbf{o}$, which means that J_o should have higher values for objects that are less compatible with our prior knowledge,[8] e.g., that are very oscillating. When $\mathbf{o}$ is not the outcome of a stochastic process, J_o usually includes a scaling factor or global hyperparameter, denoted by μ in the following, which adjusts the balance between fidelity to the data and fidelity to the prior information.

If no prior knowledge is used, which corresponds to setting $p(\mathbf{o})$ = constant in Eq. (3), one then maximizes $p(\mathbf{i}|\mathbf{o})$ (likelihood of the data) so that the solution is a maximum-likelihood solution. In this case the criterion of Eq. (4) is constituted only of the term J_i. The Richardson–Lucy algorithm[17] is an example of an iterative algorithm that converges toward the minimum of J_i when the noise follows Poisson statistics.

1. Noise Model

If the noise statistics are additive, stationary, white Gaussian, then the data fidelity term is J_i^{swG}, a simple least-squares difference between the actual data $\mathbf{i}$ and our model of the data for a given object, $\mathbf{h}*\mathbf{o}$.

In astronomical imaging, the noise is often predominantly photon noise, which follows Poisson statistics and has the following negative log-likelihood:

$$J_i^P(\mathbf{o}) = \sum_{l,m} (\mathbf{h}*\mathbf{o})(l, m) - \mathbf{i}(l, m)\ln[(\mathbf{h}*\mathbf{o})(l, m)], \quad (5)$$

where the sum extends over all pixels (l, m) of the image. This model can perform notably better than simple least-squares for objects with high dynamic range on a dark background, because the noise variance $(\mathbf{h}*\mathbf{o})(l, m)$ varies strongly between pixels for images of such objects. Yet in dark regions of the image the noise is usually predominantly detector noise, which follows Gaussian, and approximately stationary, statistics. A fine noise model should thus take into account both components of the noise.[18]

This is why we adopt a nonstationary white Gaussian model for the noise, which is a good approximation of a mix of photon and detector noise:

$$J_i^{\mathrm{mix}}(\mathbf{o}) = \sum_{l,m} \frac{1}{2\sigma^2(l, m)}[\mathbf{i}(l, m) - (\mathbf{o}*\mathbf{h})(l, m)]^2, \quad (6)$$

where $\sigma^2(l, m) = \sigma_{\mathrm{ph}}^2 + \sigma_{\mathrm{det}}^2$ is the sum of the photon noise and the detector noise variance. In the absence of detector noise, J_i^{mix} from Eq. (6) is actually a second-order expansion of J_i^P defined in Eq. (5). Additionally, from our experience, the use of J_i^{mix} rather than J_i^P makes the com-

putation of the solution with gradient-based techniques faster. We believe that this can be attributed to the fact that J_i^{mix} is quadratic.

The variance σ_{det}^2 and the variance map $\sigma_{\text{ph}}^2(l, m)$ can both be estimated from the image. Currently we estimate σ_{det}^2 by fitting the histogram of negative-valued pixels with the left half of a centered Gaussian, which leads to

$$\sigma_{\text{det}}^2 = \frac{\pi}{2}[\langle \mathbf{i}(l, m)\rangle_{((l,m);\mathbf{i}(l,m)\leqslant 0)}]^2. \tag{7}$$

Of course this estimate can be accurate only if the camera offset has been subtracted carefully. The photon noise variance map is also estimated before the deconvolution, as

$$\sigma_{\text{ph}}^2(l, m) = \max[\mathbf{i}(l, m), 0]. \tag{8}$$

This estimate is quite precise for the bright regions in the image, where photon noise dominates. Its poor accuracy in the dark regions of the image is unimportant because in these regions the main contribution to the noise variance is given by the detector.

Note that if the image to be processed is actually obtained by subtraction of two images (e.g., in the infrared, an image of interest and a sky background), then the noise variance map of the difference image can be obtained by adding the individual variance maps estimated by means of Eq. (8) on each of the two images.

Finally, one could also contemplate estimating the photon noise variance $\sigma_{\text{ph}}^2(l, m)$ at each iteration from the current object estimate. This may make J_i a nonconvex function of $\mathbf{o}$, so we have discarded this possibility.

2. Object Prior

This section aims at deriving an object prior for objects that are either smooth or piecewise smooth, such as planets and natural or artificial satellites.

The choice of a Gaussian prior probability distribution for the object can be justified from an information theory standpoint as being the least informative, given the first two moments of the distribution. In this case, a reasonable model of the object's power spectral density (PSD) can be found[10] and used to derive the regularization criterion J_o, which is then quadratic. The model parameters can even be estimated from the image itself, as shown recently in the context of phase diversity.[19–21]

The disadvantage of a Gaussian prior (or equivalently of a quadratic regularization term), especially for objects with sharp edges such as asteroids or artificial satellites, is that it tends to oversmooth edges. A possible remedy is to use an edge-preserving prior that is quadratic for small gradients and linear for large ones.[22] The quadratic part ensures a good smoothing of the small gradients (i.e., of noise), and the linear behavior cancels the penalization of large gradients (i.e., of edges), as explained by Bouman and Sauer.[23] Such priors are called quadratic–linear, or L_2–L_1 for short.[24]

Here we use a function that is an isotropic version of the expression suggested by Rey[25] in the context of robust estimation, used by Brette and Idier[26] for image restoration, and recently applied to imaging through turbulence[27,28]:

$$J_o(\mathbf{o}) = \mu\delta^2 \sum_{l,m} \phi(\nabla\mathbf{o}(l, m)/\delta), \tag{9}$$

where

$$\phi(x) = |x| - \ln(1 + |x|) \tag{10}$$

and where $\nabla\mathbf{o}(l, m) = [\nabla_x\mathbf{o}(l, m)^2 + \nabla_y\mathbf{o}(l, m)^2]^{1/2}$, and $\nabla_x\mathbf{o}$ and $\nabla_y\mathbf{o}$ are the object finite-difference gradients along x and y, respectively. This functional J_o is indeed L_2–L_1 because $\phi(x) \approx x^2/2$ for x close to 0 and $\phi(x)/|x| \to 1$ for $x \to \pm\infty$. Thus parameter δ is a (soft) threshold, in the sense that J_o switches, at each pixel (l, m), between the quadratic and the linear behaviors depending on whether $\nabla\mathbf{o}(l, m)$ is smaller than or greater than δ.

The global factor μ and the threshold δ have to be adjusted according to the noise level and the structure of the object. These two hyperparameters currently have to be adjusted by hand. As a rule of thumb, a reasonable set of hyperparameters for L_2–L_1 regularization is to take $\mu \approx 1$ and δ to be on the order of the image gradient's norm, i.e., generally several times smaller than the object gradient's norm $\|\nabla\mathbf{o}\| = [\Sigma_{l,m}|\nabla\mathbf{o}(l, m)|^2]^{1/2}$ (because for high signal-to-noise ratios, the image is essentially a smoothed version of the object). This is supported by Künsch's findings in the context of noise removal with L_2–L_1 regularization.[29] When adjusting these hyperparameters, one should bear in mind that their sensitivity is logarithmic; i.e., one must increase or decrease them by a factor of 10 to see a notable difference. Additionally, we have noticed that convergence is somewhat faster when the inversion is more regularized, i.e., when μ or δ or both are large. This is not surprising, as the inverse problem is then better conditioned, and it suggests that one should begin with large values of the hyperparameters and make them decrease rather than the other way around. See, e.g., Ref. 12 for additional information on a working strategy of hyperparameter adjustment.

The functional J_o is strictly convex because $\phi''(x) = 1/(1 + |x|)^2 > 0$, and J_i^{mix} of Eq. (6) is convex because it is quadratic, so the global criterion $J = J_i^{\text{mix}} + J_o$ is strictly convex. This ensures uniqueness and stability of the solution with respect to noise and also justifies the use of a gradient-based method for the minimization.

The photometric quality of the restoration is an ever-present concern for astronomers, as the interpretation of the restored image may heavily depend on it. An appealing property of the prior of Eq. (9) is that it does not bias the global photometry of the restoration. Indeed, because it is a function of the local pixel value differences, this prior is insensitive to a global offset of the object. This is notably different from other priors such as the several variants of the entropy, which incorporate prior knowledge on the pixel values and thus bias the photometry.

If the object is a stellar field, then a stronger prior can be used, namely, the fact that the unknown object is a collection of Dirac delta functions. In this case, the unknown parameters are no longer a pixel map but the po-

sitions and magnitudes of each star.[11] The implementation of this alternative object prior in our restoration method, assuming that the number of stars is known, is hereafter called δ-MISTRAL. The case of a crowded stellar field with an unknown number of stars is a difficult problem of its own often called "deconvolution of spike trains." It has been addressed in many papers, including Refs. 30–33, and is outside the scope of this paper. See, e.g., Ref. 34 for a review on this subject.

B. Myopic Deconvolution

As mentioned in Section 2, the true PSF is seldom available. The deconvolution of turbulence-degraded images with an unknown PSF is a difficult subject. It has been tackled by many authors (see, e.g. Ref. 10 or Ref. 28 for a short review) and was first cast into a probabilistic framework, as a joint maximum-likelihood problem, by Holmes.[35] Considering the PSF to be completely unknown usually leads to unacceptable results, because estimating both the object and the PSF from a single image is a highly underdetermined problem. Typically, even when the positivity[36,37] of the object and the PSF is strictly enforced, the estimated PSF (or the estimated object) can be reduced to a Dirac delta function. It is thus necessary to regularize the problem by adding more constraints, both on the object (see Subsection 3.A.2) and on the PSF. In particular, the band-limitedness of the PSF has been found to be an effective one.[35,36,38] For short-exposure images, the fact that the optical PSF is completely described by a phase function over the aperture is a very effective constraint.[36,39] Unfortunately, for long-exposure images, this constraint is not fulfilled; for such images, Conan *et al.* have proposed a scheme called myopic deconvolution[10] that constrains the transfer function softly at all frequencies, which is more informative than the sole band-limitedness of the PSF. This scheme is adopted here; it consists in jointly estimating the PSF and the object in the same MAP framework. This joint MAP estimator is

$$\begin{aligned}[\hat{\mathbf{o}}, \hat{\mathbf{h}}] &= \arg\max_{\mathbf{o},\mathbf{h}} p(\mathbf{o}, \mathbf{h}|\mathbf{i}) \\ &= \arg\max_{\mathbf{o},\mathbf{h}} p(\mathbf{i}|\mathbf{o}, \mathbf{h}) \times p(\mathbf{o}) \times p(\mathbf{h}) \\ &= \arg\min_{\mathbf{o},\mathbf{h}} [J_i(\mathbf{o}, \mathbf{h}) + J_o(\mathbf{o}) + J_h(\mathbf{h})]. \end{aligned} \quad (11)$$

The myopic criterion contains the two terms of Eq. (4), the first one now being a function of $\mathbf{o}$ and of $\mathbf{h}$, plus an additional term $J_h = -\ln p(\mathbf{h})$ that accounts for the knowledge, although partial, available on the PSF. Such a three-term criterion has also been obtained in a deterministic approach.[40] The regularization term for the PSF can quite naturally be derived from our probabilistic approach. The PSF is considered to be a Gaussian stochastic process since it is the temporal average of a large number of short-exposure PSFs. Additionally assuming that the difference between the PSF and its mean is stationary, J_h is given by[10]:

$$J_h(\mathbf{h}) = \frac{1}{2} \sum_f \frac{|\tilde{\mathbf{h}}(f) - \tilde{\mathbf{h}}_{\mathbf{m}}(f)|^2}{S_h(f)}, \quad (12)$$

where $\tilde{\mathbf{h}}_{\mathbf{m}} = E[\tilde{\mathbf{h}}]$ is the mean transfer function and $S_h = E[|\tilde{\mathbf{h}}(f) - \tilde{\mathbf{h}}_{\mathbf{m}}(f)|^2]$ is the associated spatial PSD. J_h is interpretable as the energy of a set of springs (one per spatial frequency) that draw the transfer function toward its mean with a stiffness given by the PSD of the PSF, which characterizes the variability of the transfer function at each frequency. Such a regularization obviously ensures that the estimated transfer function is close to the mean transfer function with respect to error bars given by the PSD. The regularization in particular enforces a zero value for the estimate of $\tilde{\mathbf{h}}(f)$ above the cutoff of the telescope, because $S_h(f)$ is zero above the cutoff. This ensures the aforementioned band-limitedness of the PSF.

In practice, the mean PSF and its PSD are estimated by replacing, in their definitions, the expected values ($E[\,\cdot\,]$) by an average on the different images recorded on an unresolved star. This star must be in a region where the seeing is the same as when observing the object. Additionally, the star flux should be chosen so that the wave-front sensing noise on the star is similar to the wave-front sensing noise on the object, the latter being a function of the object's spatial extent and flux.[41]

It is possible to estimate the PSF from the statistics of the wave-front sensing data,[14,42] which avoids the errors due to seeing fluctuations. The quantity estimated by this technique is actually the part of the *mean* PSF that is due to turbulence; the myopic approach is thus interesting in that it accounts for the PSF uncertainties due to constant aberration calibration errors,[11] residual speckle patterns because of the finite exposure time, and the wave-front sensing noise for faint stars.

Note that the new criterion is convex in $\mathbf{o}$ for a given $\mathbf{h}$ and convex in $\mathbf{h}$ for a given $\mathbf{o}$, but it is not necessarily convex on the whole parameter space. However, using reasonable starting points (see Section 4), we did not encounter the minimization problems associated with nonconvexity.

4. IMPLEMENTATION ISSUES

A. Minimization Method

The criterion of Eq. (11) is minimized numerically to obtain the joint MAP estimate for the object $\mathbf{o}$ and the PSF $\mathbf{h}$. The minimization is performed by a conjugate-gradient method,[43] which is usually recognized to be faster than expectation-maximization-based algorithms. The convergence of the conjugate-gradient method to a stationary point (in practice, to a local minimum) is guaranteed[44] because the criterion is continuously differentiable. We have found that the convergence is faster if the descent direction is re-initialized regularly; this can be attributed to the fact that the criterion to be minimized is not quadratic. This modified version of the conjugate gradient is known as the partial conjugate-gradient method and has similar convergence properties (see Sec. 8.5 of Ref. 45).

The simplest way to organize the unknowns for the minimization is to stack the object and the PSF together into a vector and to run the conjugate-gradient routine on this variable. Yet this can be slow, as the gradients of the

criterion with respect to the object and to the PSF may have different orders of magnitude. We have found that the minimization is speeded up by splitting it into two blocks and alternating between minimizations on the object for the current PSF estimate and minimizations on the PSF for the current object estimate. Additionally, for a given (current) object, the criterion of Eq. (11) is quadratic with respect to the PSF, so its minimum in **h**, without the positivity constraint, is analytical. This analytical expression can still be used as a starting point for the estimation of **h** if one wants to enforce positivity of the PSF.

The minimization starts by estimating the object for a fixed PSF taken as the mean PSF. The initial guess for the object is either the image itself or a Wiener-filtered version of the image. The minimization is not stopped by hand but rather when the estimated object and PSF no longer evolve (i.e., when their evolution from one iteration to the next is close to machine precision).

B. Positivity Constraint

The object intensity map is a set of positive values, which is important *a priori* information. One should therefore enforce a positivity constraint on the object. This constraint can be implemented in various ways,[46] such as criterion minimization under the positivity constraint, reparameterization of the object, or explicit modification of the *a priori* probability distribution (e.g., addition of an entropic term). The first two methods can actually be interpreted as an implicit modification of the *a priori* distribution that gives a zero probability to objects having negative pixel values. Note that with some expectation-maximization-based algorithms such as the Richardson–Lucy algorithm, the positivity constraint is automatically satisfied provided that the initial guess is positive.

The addition of an entropic term notably slows down the minimization; additionally, it degrades the photometric quality of the restored image as mentioned above. The reparameterization is easy to implement, but the only one we found that does not notably slow down the minimization is[47] $\mathbf{o}(l, m) = \mathbf{a}(l, m)^2$, which induces local minima because it is not monotonic. We have found that the best way to ensure positivity, with respect both to speed and to not introducing local minima, is to directly minimize the criterion under this constraint. We do so by means of the projected gradient, as proposed in image processing by Commenges[48] and by Nakamura *et al.*[49]

5. VALIDATION ON SIMULATED DATA

This section successively validates the contribution of the various components of our restoration method to the quality of the restored image, by means of simulations. These components are the noise model, the positivity, the edge-preserving object prior, and the myopic capability. To have a realistic PSF and PSF variability, we use as PSFs five experimental NAOS–CONICA[50] images of an unresolved bright star recorded in the Brγ narrow band (wavelength $\lambda = 2.166\ \mu$m). They are shown in Fig. 1. Their Strehl ratios range from 52% to 66%, and the Strehl ratio fluctuation is due to seeing variations. In all subsections but Subsection 5.D, the PSF is assumed to be known.

Two types of objects are used: a planetary-type object (synthetic asteroid of uniform level with a 10% brighter broad feature and three 30% brighter spots, surrounded by a few stars), and a galaxy (scaled image of M51).

The images are obtained by convolution of these objects with a PSF (the fourth of the PSFs shown in Fig. 1) and contamination by noise. The noise is the combination of photon noise (Poisson statistics with a total flux of 10^7 photons) and detector noise [stationary Gaussian statistics with a standard deviation of 10 photoelectrons/pixel (ph/pix)].

Figure 2 shows the objects and the simulated noisy images. The ideal, i.e., noiseless diffraction-limited, images are also shown for comparison.

A. Effect of the Noise Model

The restored object minimizes Eq. (4) with J_o being the quadratic regularization derived from the Gaussian prior used in Ref. 10. Figure 3 shows the best restorations obtained for each of two white-noise models when the global hyperparameter μ is varied. These models are the simple stationary model, which leads to the least-squares criterion J_i^{swG} mentioned in Subsection 3.A.1, and the nonstationary model J_i^{mix} of Eq. (6), which accounts for both photon noise and detector noise. The root-mean-square error is slightly smaller for the finer noise model (2052 ph/pix) than for the simple one (2115 ph/pix). This error is computed as the mean square difference between the restoration and the true object on all pixels where the object's value is nonzero, so it is indeed indicative of the restoration quality of the object rather than of the background. Visually, the presence of the star on the right of

Fig. 1. Five experimental VLT-NAOS-CONICA images of an unresolved bright star recorded on Sep. 29, 2002, at 20:58 UT in the Brγ narrow band and used in the simulations (wavelength $\lambda = 2.166\ \mu$m, exposure time 2 s). The fourth one is used as the true PSF. The estimated Strehl ratios are 57%, 52%, 64%, 66%, and 58%. The corresponding Fried parameters are $r_0 = 17.5$, 15.6, 20.7, 21.0 and 17.0 cm.

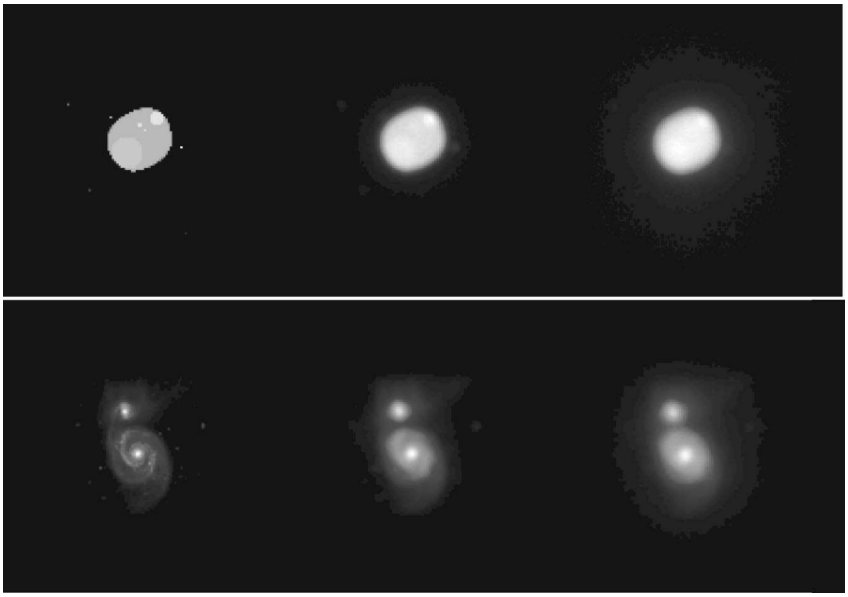


Fig. 2. From left to right: original object used for the simulation, noiseless image by a perfect diffraction-limited telescope (for comparison), and simulated image obtained by adding noise to the convolution of the object by the PSF (the fourth of the PSFs shown in Fig. 1). Top, asteroid; bottom, galaxy.

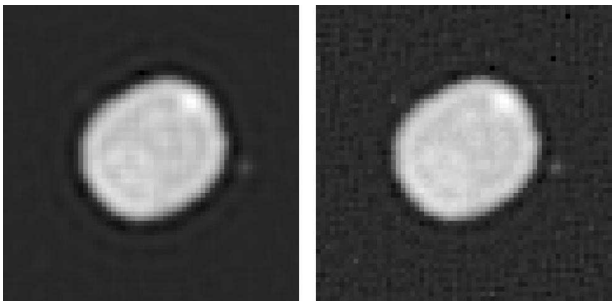


Fig. 3. Best restorations of the asteroid obtained with quadratic regularization for the two noise models, without a positivity constraint. Left, stationary model; right, nonstationary model, which accounts for both photon noise and detector noise. Restoration errors are 2115 and 2052 ph/pix, respectively. These restored images are zoomed by a factor of two for better visualization.

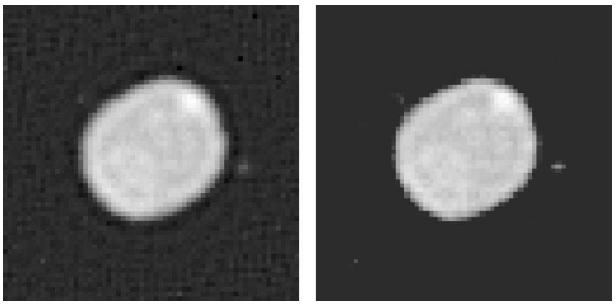


Fig. 4. Restorations of the asteroid obtained with quadratic regularization and the nonstationary noise model, without (left) and with (right) the positivity constraint. Restoration errors are 2052 and 1764 ph/pix, respectively. These restored images are zoomed by a factor of 2 for better visualization.

the asteroid is more obvious with the finer model, which is the one used from now on.

B. Effect of the Positivity Constraint

The influence of the positivity constraint is illustrated on the asteroid case in Fig. 4. It is implemented through projection, as discussed in Subsection 4.B. This constraint helps reduce noise and ringing in the dark regions of the image, i.e., where it is actually enforced. Indeed, the root-mean-square error drops from 2052 to 1764 ph/pix; this constraint will thus be used in all the following restorations.

Yet, some ringing remains inside the object because of its sharp edges and the use of a quadratic regularization; indeed, quadratic regularization precludes spectral extrapolation and thus can produce a restored object with a sharp cutoff in Fourier space. In the object space, this sharp cutoff takes the form of ringing, akin to Gibbs oscillations.[51] Such artifacts prevent any astrophysical interpretation that would rely on precise photometry.

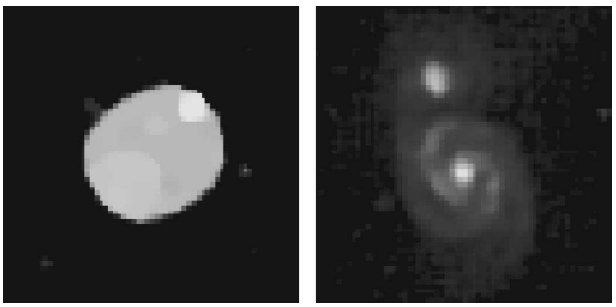


Fig. 5. Restorations obtained with the edge-preserving prior, the nonstationary noise model, and the positivity constraint for the best values of the hyperparameters. Restoration errors are 1201 ph/pix for the asteroid (left) and 985 ph/pix for the galaxy (right). These restored images are zoomed by a factor of 2 for better visualization.

C. Effect of the Edge-Preserving Object Prior

Figure 5 shows the restoration of the asteroid and of the galaxy with the edge-preserving prior of Eq. (9). They are obtained with $(\delta, \mu) = (0.03, 30)$ and $(30, 0.03)$, respectively. The restoration errors are, respectively, 1201

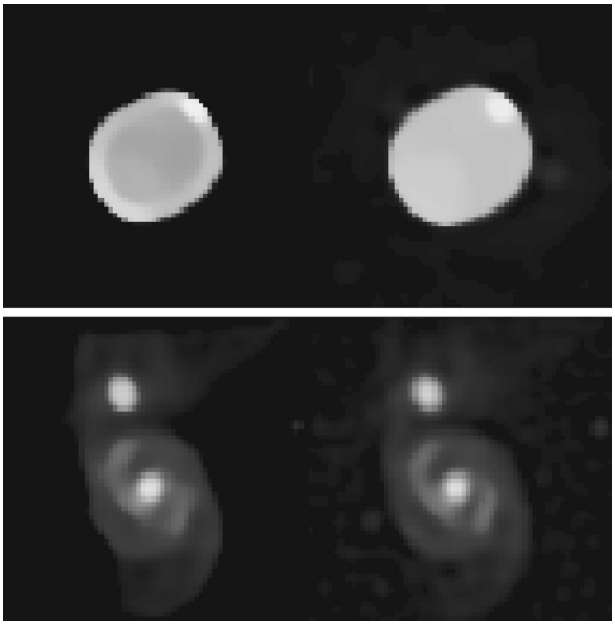


Fig. 6. Classical deconvolution using the mean PSF as the true one (left); myopic deconvolution (right). Restoration errors are 5139 and 2333 ph/pix, respectively, for the asteroid (top) and 1671 and 1365 ph/pix for the galaxy (bottom).

and 985 ph/pix. The enhancement of the photometric quality of the L_2–L_1 restoration of the asteroid, with respect to the quadratic restoration of Fig. 4, is obvious: The 10% brighter broad feature and the largest of the three 30% brighter spots are visible, and the piecewise uniformity of the asteroid is respected.

D. Effect of the Myopic Capability

This subsection assumes that the true PSF is unknown and that we instead have five noisy PSFs taken shortly before or after the object of interest. These are the five images of a nearby star shown in Fig. 1. We thus use the myopic scheme described in Subsection 3.B, which is capable of estimating the PSF and the object simultaneously. We use the edge-preserving prior of Eq. (9) and the fine noise model of Eq. (6). The mean PSF and the PSD of the PSF are estimated by replacing, in their definitions, the expected values ($E[\,\cdot\,]$) by an average on the five images. The quality of the restoration can be favorably compared with that of a "classical" (i.e., with known PSF) deconvolution by using the mean PSF: for the asteroid, the latter deconvolution gives an error of 5139 ph/pix, whereas the myopic deconvolution gives an error of 2333 ph/pix. For the galaxy, these restoration errors are 1671 and 1365 ph/pix, respectively. Figure 6 shows the restored asteroid and galaxy for the two restoration methods. As expected, the myopic deconvolution performs better than the classical deconvolution with mean PSF because the former does not assume erroneously that the PSF is known, but not as well as the classical deconvolution with the true PSF (see Fig. 5).

6. DECONVOLUTION OF EXPERIMENTAL DATA

This section presents results obtained on different types of AO systems (BOA, NAOS, HOKUPA'A, ADONIS, PUEO, Keck-AO) and different kinds of astronomical objects (planetary objects, pointlike objects, etc). The average PSF and the PSD of the PSF is computed from a set of images of a nearby star recorded after (and also before, whenever possible) the object of interest. For each of these examples, astrophysical results have been derived from the deconvolved images and have lead to scientific publications.

A. Planetary Objects

1. Ganymede Observed with BOA

Ganymede is the biggest moon of Jupiter (visual magnitude $M_v = 4.6$, diameter approximately 1.7 arcsec). It was observed by the 1.52-m telescope of Observatoire de Haute Provence by ONERA's AO bench called BOA (for Banc d'Optique Adaptative) on September 28, 1997.[52,53] The seeing conditions were particularly severe: $D/r_o \approx 23$ at the imaging wavelength 0.85 μm. The object itself is used for the wave-front sensing. The corrected image (100 s exposure time) is shown in Fig. 7(a) and does not exhibit any detail (the Strehl ratio is approximately 5%). The field of view is 3.80 arcsec. The estimated total flux is approximately 8×10^7 photons. The star θ Cap ($M_v = 4.1$), located 1.5 deg away from Ganymede, was then observed to provide a PSF calibration. A neutral density, i.e., a light attenuator that is not chromatic, was used to have approximately the same wave-front sensing conditions. The PSD of the PSF and the mean PSF were estimated from a series of 50 images recorded with an exposure time of 1 s. The difference in exposure times between the reference star and Ganymede was accounted for in the PSD computation.[10] The MISTRAL de-

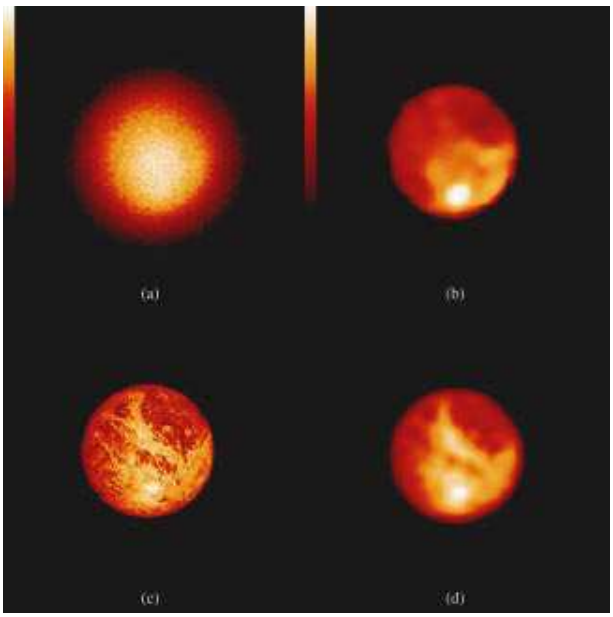


Fig. 7. (a) image of Ganymede with ONERA's AO bench at the 1.52-m telescope of Observatoire de Haute Provence on 1997/09/28, 20:18 UT. (b) MISTRAL deconvolution. (c) Synthesized image obtained by mapping broadband probe images of NASA's data base into a view of Ganymede seen from Earth at the time (a) was taken (courtesy NASA/JPL/Caltech, see http://space.jpl.nasa.gov/). (d) Same synthesized image convolved with the Airy disk of a 1.52-m telescope, for comparison with the deconvolved image.

convolution is shown on Fig. 7(b). Although a probe passing by Jupiter obviously exhibits a better resolution than a 1.52-m telescope on the ground, many features of Ganymede (various dark areas, bright spot) are visible in the deconvolved AO-corrected image and can be compared with the NASA/JPL high-resolution image [Fig. 7(c)]. The latter image was synthesized by mapping broadband probe images of NASA/JPL's database into a view of Ganymede seen from Earth at the time Fig. 7(a) was taken (see the solar system simulator at http://space.jpl.nasa.gov/). A fairer comparison can be done between the MISTRAL deconvolution and the image of Fig. 7(c) convolved with the Airy pattern of a 1.52-m telescope; the latter image is shown on Fig. 7(d).

2. Io Observed with Keck-AO

Io, the innermost Galilean satellite of Jupiter (with angular size of 1.2 arcsec) was observed in near IR in February 2001 from the 10-m Keck II telescope with use of its AO system. After deconvolution with MISTRAL, the resolution, approximately 100 km on Io's disk, is comparable with the best Galileo/NIMS (Near Infrared Mapping Spectrometer) resolution for global imaging and allows, for the first time, investigation of the very nature of individual eruptions[54] (see Figs. 8 and 9). On February 19, two volcanos, Amirani and Tvshatar, with temperatures differing from the Galileo observations, were observed. On February 20, a slight brightening near the Surt volcano was detected that turned into an extremely bright volcanic outburst two days later. Thanks to the quality of the photometric restitution obtained with MISTRAL,[55] these outburst data have been fitted with a silicate cooling model, which indicates that this is a highly vigorous eruption with a highly dynamic emplacement mechanism, akin to fire fountaining. Its integrated thermal output was close to the total estimated output of Io, making this the largest Ionian thermal outburst yet witnessed.

3. Io Observed with NAOS

Io was observed on December 5, 2001, with the recently commissioned VLT/NAOS AO system[50,56] and its infrared camera CONICA.[57] The Strehl ratio on this Bracket-γ (2.166 μm) observation is estimated to be 35%. This wavelength range mainly gives information about reflected sunlight modulated by various surface features. The image deconvolved with MISTRAL is shown in Fig. 10. Dark caldera, such as Pele and Pillan patera, are visible in the southeast area of the disk. The low-albedo area that is at the North of the center of the disk corresponds to Lei–Kung fluctus, a lava flow field. No high-temperature volcanic hot spot was detected during this observing night. Ground-based monitoring programs us-

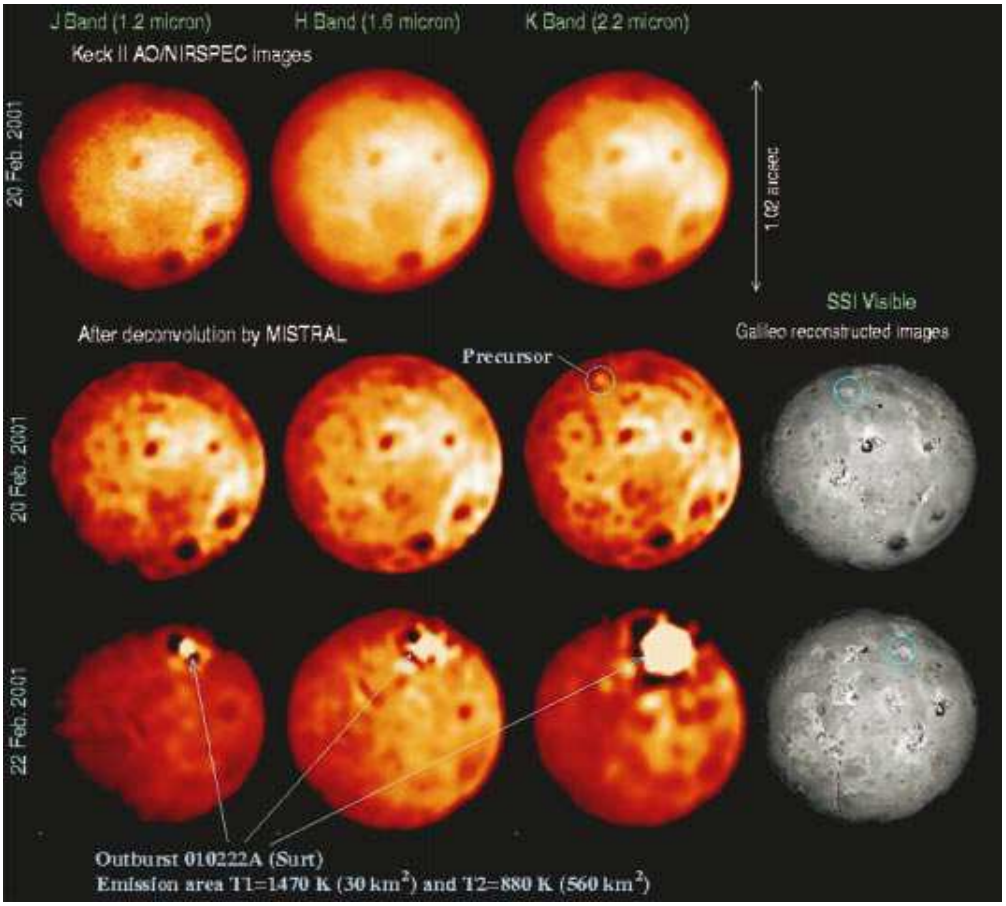


Fig. 8. Jupiter-facing hemisphere of Io observed with the Keck AO system in J, H, and K band (from left to right). The basic preprocessed images from February 20, 2001, are displayed on the first row. The second row corresponds to the same images after deconvolution. Albedo features, comparable with the 20-km resolution reconstructed Galileo/SSI image (right column) are easily detected. The last row shows the February 22, 2001, images, which are dominated by the presence of the Surt outburst.

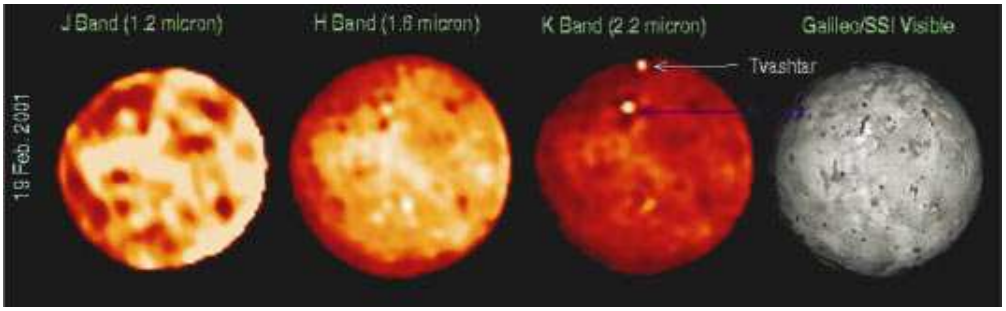


Fig. 9. Io observed by the Keck from February 19, 2001, and deconvolved with MISTRAL. Two hot spots, corresponding to Tvashtar (north) and Amirani, are clearly detected in the H and K bands. The bad seeing conditions of this observation induced very poor AO correction in the J band, which explains the poor restitution quality in the J band.

ing AO systems on 8-m–10-m-class telescopes will help astronomers characterize the time evolution of Io's volcanic activity, including the frequency, spatial distribution, and temperature of hot spots and surface changes.[58,59] Indeed, with the Galileo mission coming to an end, the future monitoring of Io's volcanism lies in the hands of terrestrial observers.

4. Neptune Observed with HOKUPA'A

The 36-actuator curvature AO system called HOKUPA'A was used to observe Neptune at the Canada–France–Hawaii Telescope both in November 1997 and in July 1998. It produced the first sharp infrared images of Neptune. These images show the fine structures of its cloud bands with high contrast, allowing the details of Neptune's atmospheric activity to be observed from the ground for the first time.[60] Figure 11 shows images of Neptune obtained on July 6, 1998 in a methane absorption band (1.72 μm) and deconvolved with MISTRAL. At this wavelength Neptune's atmosphere is very dark, and high altitude clouds appear with a high contrast. The top three images are individual 600-s exposures taken at the time indicated above each image. Note how the fine structure in the cloud bands can be followed from one frame to another as the planet rotates. The upper left and right images have been numerically rotated about Neptune's rotation axis to match the central image and added to it to form the bottom left image, thus improving the signal-to-noise ratio. The bottom right image is the same as the bottom left one except for its color table, which better shows the low-light levels. The periodic pattern of bright dots seen just above the southernmost cloud band (at the bottom of each image) is particularly remarkable. Such a regular pattern of small clouds has never before been observed on Neptune and may be indicative of gravity waves in Neptune's atmosphere.

5. Uranian System Observed with ADONIS

MISTRAL has been applied to infrared images of Uranus acquired on May 2, 1999 with the ADONIS AO system. The deconvolved images in the J and H bands exhibit structures on the planet (bright polar haze).[61] When looking at low intensity levels (see Fig. 12), one can also see the structure of the Epsilon ring and of the innermost rings, as well as very faint satellites discovered by Voyager 2 in 1986 and never reobserved since.

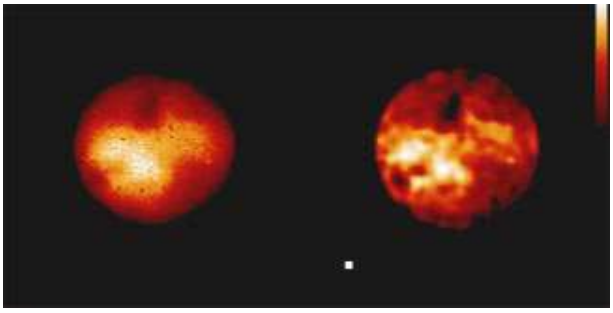


Fig. 10. Left: Io observed with NAOS–CONICA on December 5th, 2001 (7:14:59 UTC), in the Brγ band (2.166 μm); north is up and east is left. The object itself is used for the wave-front sensing. The camera pixel scale is 13.25 mas. The seeing was 0.9 arcsec, and the estimated Strehl ratio is 35%. Right, deconvolution using MISTRAL. The two images are given in the same linear color scale. The white square represents the telescope diffraction limit at the observing wavelength.

B. Disklike Objects: MBM 12 Association Observed with Pueo

In the younger association MBM 12, seven binaries and a quadruple system including a protoplanetary disk have recently been detected and deconvolved.[62] For the young protoplanetary disk LkHα 263 C seen almost perfectly edge-on, MISTRAL was applied to recover the maximum spatial information possible (see Fig. 13). The deconvolved images were then compared with synthetic images of a disk model so as to extract structural parameters such as outer radius, dust mass, and inclination.[62]

C. Pointlike Objects

1. Capricornus Association Observed with ADONIS

The young, nearby stellar associations are ideal laboratories to study the formation and evolution of circumstellar disks, brown dwarfs, and planets around solar-type stars. Owing to their proximity (closer than 100 pc), small separations can be reached to explore the faint circumstellar environment of such associations. In the Capricornus association (distance 48 pc; age less than 30 Myr), the source HD 199143 has been previously resolved as a binary system.[63] The strong IR color ($J - K = 1.37$) of

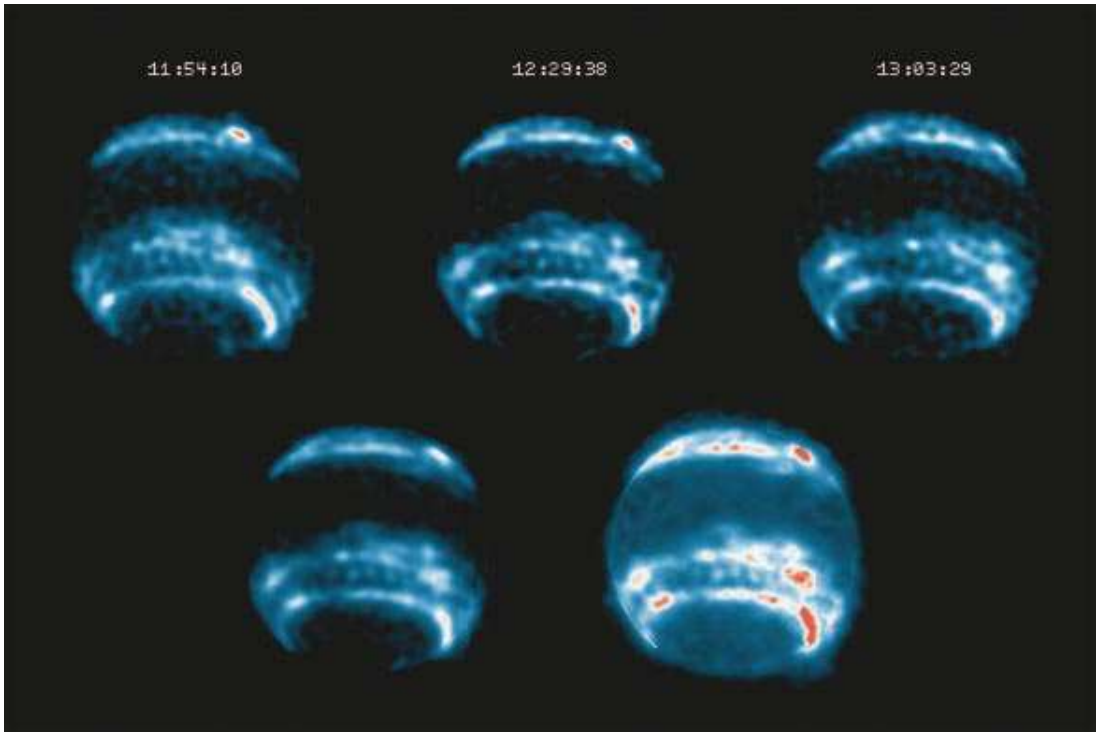


Fig. 11. False-color images of Neptune obtain at the Canada–France–Hawaii Telescope with the HOKUPA'A AO system on July 6, 1998, in a methane absorption band (1.72 μm) and deconvolved by MISTRAL. The object itself is used for the wave-front sensing. The top three images are individual 600-s exposures taken at 11:54:10, 12:29:38, and 13:03:29 UT time. The left and right images have been numerically rotated about Neptune's rotation axis to match the central image and co-added to it to form the bottom images, thus improving the signal-to-noise ratio. The color scale in the bottom right image shows the fainter details.

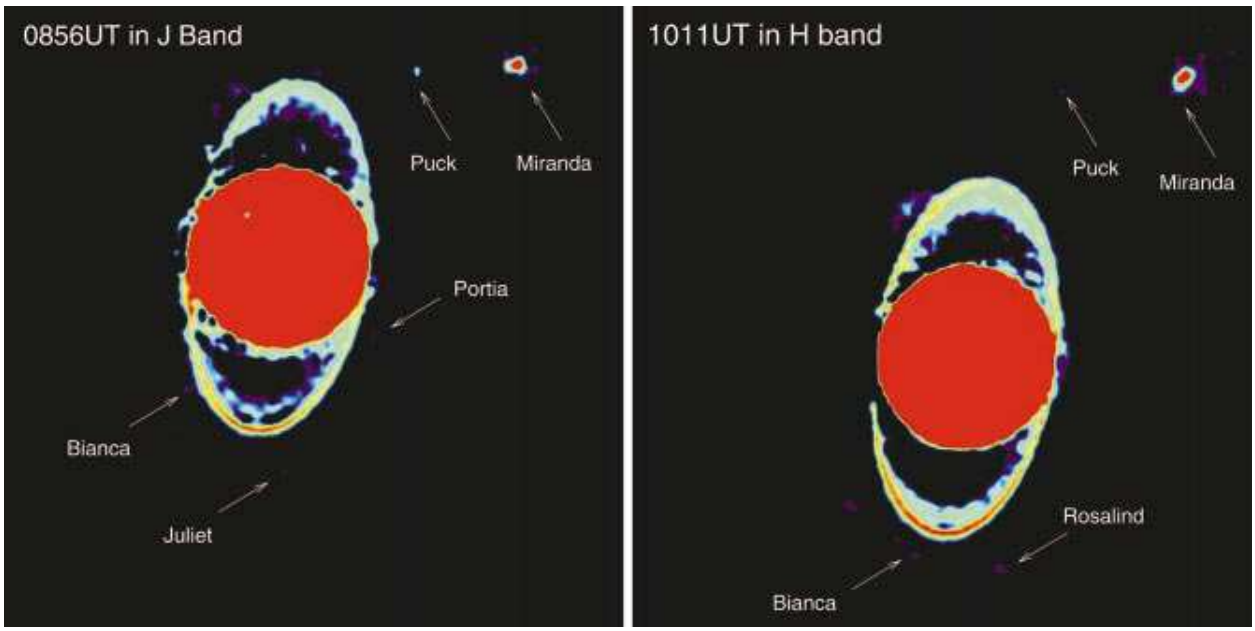


Fig. 12. Logarithmic display of the images restored by MISTRAL; north is up and east is left. The chosen color scale saturates the high levels so as to make visible the faint details. Innermost rings and faintest satellites, first observed with Voyager in 1986, are also detected.

the companion had been estimated from an image deconvolved with IDAC (Iterative Deconvolution Algorithm in C)[38] and attributed to the presence of a "circumsecondary" disk. New observations of Chauvin *et al.*, after deconvolution by δ-MISTRAL (see an example of deconvolution with MISTRAL in the J band in Fig. 14), did not confirm the photometry in the J band of HD 199143 B. Instead, these new results showed that no disk is needed to explain the re-estimated IR color ($J - K = 0.81$) of HD 199143 B, now interpreted as a late M2 dwarf.[64]

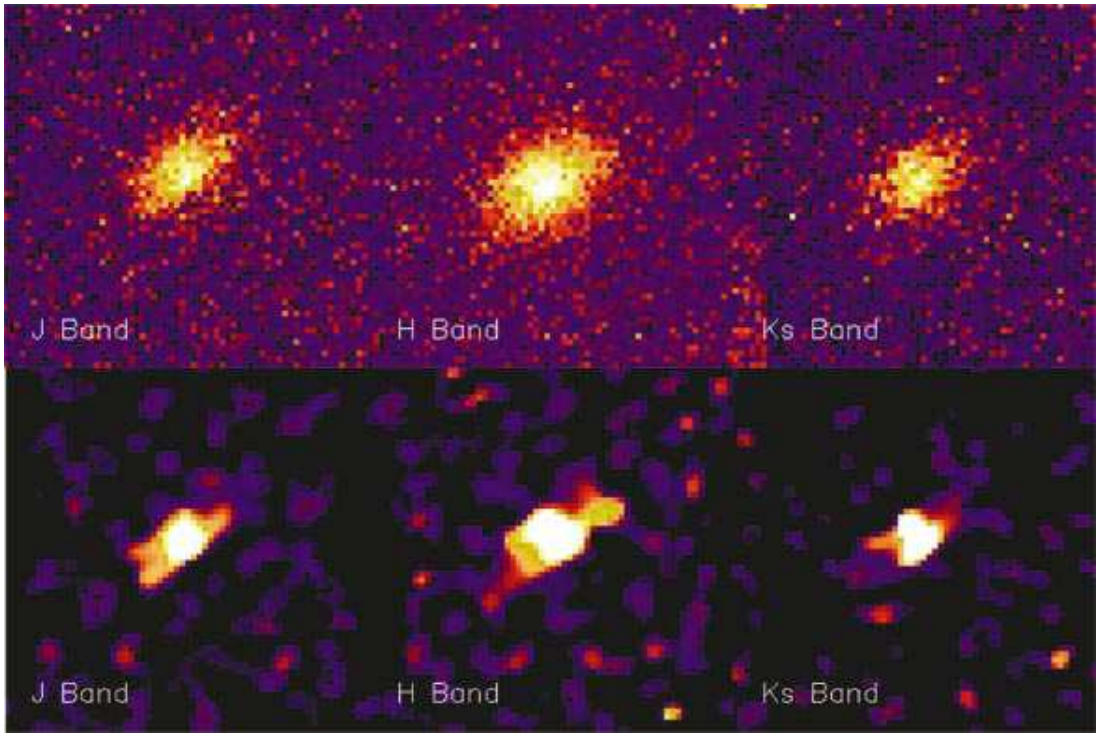


Fig. 13. Top row, raw images of the circumstellar disk surrounding LkHα 263 C in J, H, and Ks (from left to right). Bottom row, corresponding deconvolved images. On all images the field of view is 2.2 arcsec, with north up and east to the left.

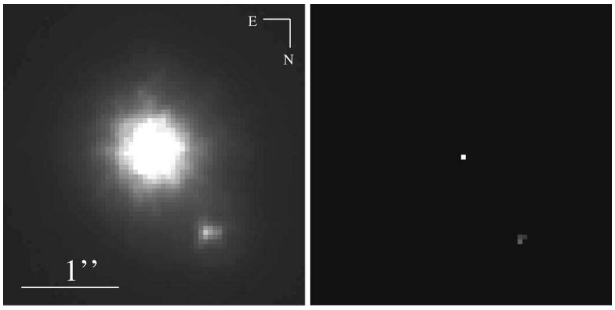


Fig. 14. Left, ADONIS image in the J band of HD 199143 A and B. The object itself is used for the wave-front sensing. Right, resulting image obtained with MISTRAL.

2. GJ263 Observed with NAOS

Figure 15 shows how diffraction-limited imaging with NAOS–CONICA at a wavelength of 1.257 μm shows the individual components of the close binary star GJ 263. The angular distance between the two stars is only 0.040 arcsec. Spatially resolved observations combined with precise photometric deconvolution of binary stars like this one will allow the determination of orbital parameters and ultimately of the masses of the individual binary star components.

D. Other Objects and Other Adaptive Optics Systems

A nonexhaustive list of astronomical images taken on various systems and deconvolved with MISTRAL is presented below with their corresponding publications.

- Titan observed with PUEO[65] and with Keck-AO[66]: detection of albedo surface features.

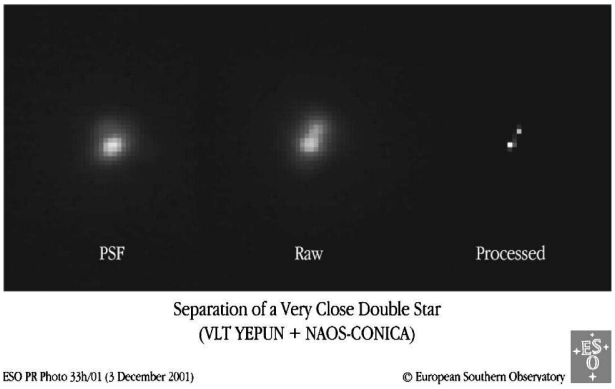


Fig. 15. NAOS–CONICA image of the double star GJ 263; the angular distance between the two components is 0.040 arcsec. The raw image, as directly recorded by CONICA, is shown in the middle, with a MISTRAL deconvolved version to the right. The recorded PSF is shown to the left. The object itself was used for the wave-front sensing. The C50S camera (0.01325 arcsec/pixel) was used with an FeII filter at the near-infrared wavelength 1.257 μm. The exposure time was 10 s. (See http://www.eso.org/outreach/press-rel/pr-2001/pr-25-01.html for more details).

- Tucana–Horologium association observed with ADONIS[67]: precise photometry on several binaries.
- The main belt asteroid 216-Kleopatra observed with ADONIS[68]: bifurcated shapes, density, and origin.
- The main belt asteroid Vesta observed with Keck-AO: shape, mineralogy, etc.

7. CONCLUSION AND PERSPECTIVES

A deconvolution method has been derived in a Bayesian framework. Its three main components are a fine noise

model, a PSF estimation capability, and a carefully designed object regularization term. The nonstationary noise model accounts for both photonic and detector noise and yields a large dynamic range in the restored images. Additionally, this method reckons with the usually imperfect knowledge of the PSF by estimating the PSF simultaneously with the object under soft constraints that embody our uncertainty about the PSF. Finally, this method comprises a regularization term that is appropriate for a wide class of objects, namely, objects that are a mix of sharp edges and smooth areas. This regularization does not bias the photometry and can restore sharp edges without ringing effects. The implementation of this method, called MISTRAL, allows a positivity constraint to be enforced both on the object and on the PSF.

The contributions of the different components of this method to the overall quality of the restoration have been validated by simulations. The effectiveness of MISTRAL has been illustrated by several results on experimental and scientific data taken on various AO systems and telescopes. Additionally, this method has already been successfully employed in a number of astronomical publications to derive astrophysical results.

Future work should include the automatic tuning of the hyperparameters. The deconvolution of images larger than the isoplanatic patch also deserves further study. At least when the PSF is known, the result of the use of MISTRAL "as is" on such data can be predicted. Indeed, it has been shown that for AO-corrected images, the PSF within an angular patch away from the guide star is the on-axis PSF convolved with an anisoplanatism PSF,[69] which is a delta function on axis and widens as the considered patch direction gets farther from the guide star. Hence MISTRAL should be able to deconvolve the image from the on-axis PSF and restore an object smoothed only by this anisoplanatism PSF.

Finally, integrating the PSF out of the problem rather than estimating it jointly with the object should also be studied. Indeed, the reverse operation, i.e., integrating the object out of the problem when estimating the PSF of an instrument, has already been proven to be a successful approach in the context of phase diversity, since it offers robustness to noise and should asymptotically remove the local minima associated with the joint estimation.[20,21]

ACKNOWLEDGMENTS

We are grateful to our colleagues of the High-Resolution Imaging team at ONERA for making the Banc d'Optique Adaptative happen as well as for many fruitful discussions. We warmly thank all the astronomers who participated in the data reduction and used the deconvolved images for their astrophysical interpretations, in particular Franck Marchis, Gaël Chauvin and the NAOS–CONICA team, François and Claude Roddier, Christophe Dumas, and Athena Coustenis. We acknowledge the referees for their most careful reviews and constructive comments.

Corresponding author Laurent Mugnier can be reached at Laurent.Mugnier@onera.fr.

REFERENCES

1. J. W. Hardy, J. E. Lefevbre, and C. L. Koliopoulos, "Real time atmospheric compensation," J. Opt. Soc. Am. **67**, 360–369 (1977).
2. G. Rousset, J.-C. Fontanella, P. Kern, P. Gigan, F. Rigaut, P. Léna, C. Boyer, P. Jagourel, J.-P. Gaffard, and F. Merkle, "First diffraction-limited astronomical images with adaptive optics," Astron. Astrophys. **230**, 29–32 (1990).
3. F. Roddier, ed., *Adaptive Optics in Astronomy* (Cambridge U. Press, Cambridge, UK, 1999).
4. F. Rigaut, G. Rousset, P. Kern, J.-C. Fontanella, J.-P. Gaffard, F. Merkle, and P. Léna, "Adaptive optics on a 3.6-m telescope: results and performance," Astron. Astrophys. **250**, 280–290 (1991).
5. M. C. Roggemann, "Limited degree-of-freedom adaptive optics and image reconstruction," Appl. Opt. **30**, 4227–4233 (1991).
6. J. M. Conan, P. Y. Madec, and G. Rousset, "Image formation in adaptive optics partial correction," in *Active and Adaptive Optics*, F. Merkle, ed., ESO Conference and Workshop Proceedings (Garching bei München, Germany, 1994), pp. 181–186.
7. J.-M. Conan, "Étude de la correction partielle en optique adaptative," Ph.D. thesis (Université Paris XI Orsay, Orsay, France, 1994).
8. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust. Speech Signal Process. **37**, 2024–2036 (1989).
9. J. Idier, ed., *Approche bayésienne pour les problèmes inverses* (Hermès, Paris, 2001).
10. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
11. T. Fusco, J.-P. Véran, J.-M. Conan, and L. Mugnier, "Myopic deconvolution method for adaptive optics images of stellar fields," Astron. Astrophys. Suppl. Ser. **134**, 1–10 (1999).
12. A. D. Storrs, C. Dunne, J.-M. Conan, L. Mugnier, B. P. Weiss, and B. Zellner, "A closer look at main belt asteroids 1: WF/PC images," Icarus (to be published).
13. D. Gratadour, L. M. Mugnier, and D. Rouan, "Image centering with a maximum likelihood estimator: application to infrared astronomical imaging," in *ADASS XIII*, F. Ochsenbein, M. Allen, and D. Egret, eds., Astron. Soc. Pacific Conf. Series 30 (Publications of the Astronomical Society of the Pacific, San Francisco, Calif., 2004); manuscript available from the authors, Damien.Gratadour@onera.fr.
14. J.-P. Véran, F. Rigaut, H. Maitre, and D. Rouan, "Estimation of the adaptive optics long-exposure point-spread function using control loop data," J. Opt. Soc. Am. A **14**, 3057–3069 (1997).
15. A. Blanc, L. M. Mugnier, and J. Idier, "Marginal estimation of aberrations and image restoration by use of phase diversity," J. Opt. Soc. Am. A **20**, 1035–1045 (2003).
16. M. Hartung, A. Blanc, T. Fusco, F. Lacombe, L. M. Mugnier, G. Rousset, and R. Lenzen, "Calibration of NAOS and CONICA static aberrations. Experimental results," Astron. Astrophys. **399**, 385–394 (2003).
17. W. H. Richardson, "Bayesian-based iterative method of image restoration," J. Opt. Soc. Am. **62**, 55–59 (1972).
18. J. Nunez and J. Llacer, "A general Bayesian image reconstruction algorithm with entropy prior: preliminary application to HST data," Publ. Astron. Soc. Pac. **105**, 1192–1208 (1993).
19. A. Blanc, J. Idier, and L. M. Mugnier, "Novel estimator for the aberrations of a space telescope by phase diversity," in *UV, Optical, and IR Space Telescopes and Instruments*, J. B. Breckinridge and P. Jakobsen, eds., Proc. SPIE **4013**, 728–736 (2000).
20. A. Blanc, "Identification de réponse impulsionnelle et restauration d'images: apports de la diversité de phase," Ph.D. thesis (Université Paris XI Orsay, Orsay, France, 2002).
21. A. Blanc, T. Fusco, M. Hartung, L. M. Mugnier, and G.

Rousset, "Calibration of NAOS and CONICA static aberrations. Application of the phase diversity technique," Astron. Astrophys. **399**, 373–383 (2003).
22. P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," IEEE Trans. Med. Imaging **9**, 84–93 (1990).
23. C. Bouman and K. Sauer, "A generalized gaussian image model for edge-preserving MAP estimation," IEEE Trans. Image Process. **2**, 296–310 (1993).
24. J. Idier and L. Blanc-Féraud, "Déconvolution en imagerie," in *Approche bayésienne pour les problèmes inverses*, J. Idier, ed. (Hermès, Paris, 2001), Chap. 6, pp. 139–165.
25. W. J. J. Rey, *Introduction to Robust and Quasi-Robust Statistical Methods* (Springer-Verlag, Berlin, 1983).
26. S. Brette and J. Idier, "Optimized single site update algorithms for image deblurring," in *Proceedings of the International Conference on Image Processing* (IEEE Computer Society, Press, Los Alamitos, Calif., 1996), pp. 65–68.
27. L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem, "Myopic deconvolution from wave-front sensing," J. Opt. Soc. Am. A **18**, 862–872 (2001).
28. L. Mugnier and G. Le Besnerais, "Problèmes inverses en imagerie optique à travers la turbulence," in *Approche bayésienne pour les problèmes inverses*, J. Idier, ed. (Hermès, Paris, 2001), Chap. 10, pp. 241–270.
29. H. R. Künsch, "Robust priors for smoothing and image restoration," Ann. Inst. Statist. Math. **46**, 1–19 (1994).
30. J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," Astron. Astrophys. Suppl. Ser. **15**, 417–426 (1974).
31. P. Stetson, "DAOPHOT: a computer program for crowded-field stellar photometry," Publ. Astron. Soc. Pac. **99**, 191–222 (1987).
32. P. Magain, F. Courbin, and S. Sohy, "Deconvolution with correct sampling," Astrophys. J. **494**, 472–477 (1998).
33. K. F. Kaaresen, "Evaluation and applications of the iterated window maximization method for sparse deconvolution," IEEE Trans. Signal Process. **46**, 609–624 (1998).
34. F. Champagnat, Y. Goussard, S. Gautier, and J. Idier, "Déconvolution impulsionnelle," in *Approche bayésienne pour les problémes inverses*, J. Idier, ed. (Hermès, Paris, 2001), Chap. 5, pp. 115–138.
35. T. J. Holmes, "Blind deconvolution of speckle images quantum-limited incoherent imagery: maximum-likehood approach," J. Opt. Soc. Am. A **9**, 1052–1061 (1992).
36. E. Thiébaut and J.-M. Conan, "Strict *a priori* constraints for maximum-likelihood blind deconvolution," J. Opt. Soc. Am. A **12**, 485–492 (1995).
37. D. A. Fish, A. M. Brinicombe, E. R. Pike, and J. G. Walker, "Blind deconvolution by means of the Richardson–Lucy algorithm," J. Opt. Soc. Am. A **12**, 58–65 (1995).
38. S. M. Jefferies and J. C. Christou, "Restoration of astronomical images by iterative blind deconvolution," Astrophys. J. **415**, 862–874 (1993).
39. T. J. Schulz, "Multiframe blind deconvolution of astronomical images," J. Opt. Soc. Am. A **10**, 1064–1073 (1993).
40. Y.-L. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," IEEE Trans. Image Process. **5**, 416–428 (1996).
41. G. Rousset, "Wave-front sensors," *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, Cambridge, UK, 1999), Chap. 5, pp. 91–130. See also Ref. 3.
42. S. Harder, "Reconstruction de la réponse impulsionnelle du système d'optique adaptative ADONIS à partir des mesures de son analyseur de surface d'onde et Étude photométrique de la variabilité des étoiles YY Orionis," Ph.D. thesis (Université Joseph Fourier, Grenoble 1, France, 1999).
43. Groupe Problèmes Inverses, "GPAV: une grande œuvre collective," internal report (Laboratoire des Signaux et Systèmes, CNRS/Supélec/Université Paris-Sud, Orsay, France, 1997).
44. D. P. Bertsekas, *Nonlinear Programming* (Athena Scientific, Belmont, Mass., 1995).
45. D. G. Luenberger, *Introduction to Linear and Nonlinear Programming* (Addison-Wesley, Reading, Mass., 1973).
46. D. M. Titterington, "General structure of regularization procedures in image reconstruction," Astron. Astrophys. **144**, 381–387 (1985).
47. Y. Biraud, "A new approach for increasing the resolving power by data processing," Astron. Astrophys. **1**, 124–127 (1969).
48. D. Commenges, "The deconvolution problem: fast algorithms including the preconditioned conjugate-gradient to compute a MAP estimator," IEEE Trans. Autom. Control **29**, 229–243 (1984).
49. O. Nakamura, S. Kawata, and S. Minami, "Optical microscope tomography. II. Nonnegative constraint by a gradient-projection method," J. Opt. Soc. Am. A **5**, 554–561 (1988).
50. G. Rousset, F. Lacombe, P. Puget, N. Hubin, E. Gendron, T. Fusco, R. Arsenault, J. Charton, P. Gigan, P. Kern, A.-M. Lagrange, P.-Y. Madec, D. Mouillet, D. Rabaud, P. Rabou, E. Stadler, and G. Zins, "NAOS, the first AO system of the VLT: on sky performance," in *Adaptive Optical System Technology II*, P. L. Wizinowich and D. Bonaccini, eds., Proc. SPIE **4839**, 140–149 (2002).
51. D. L. Snyder, M. I. Miller, L. J. Thomas, Jr., and D. G. Politte, "Noise and edge artifacts in maximum-likelihood reconstruction for emission tomography," IEEE Trans. Med. Imaging **6**, 228–237 (1987).
52. J.-M. Conan, T. Fusco, L. Mugnier, F. Marchis, C. Roddier, and F. Roddier, "Deconvolution of adaptive optics images: from theory to practice," in *Adaptive Optical Systems Technology*, P. Wizinowich, ed., Proc. SPIE **4007**, 913–924 (2000).
53. T. Fusco, "Correction partielle et anisoplanétisme en optique adaptative: traitement *a posteriori* et optique adaptative multiconjuguée," Ph.D. thesis (Université de Nice-Sophia Antipolis, Nice, France, 2000).
54. F. Marchis, I. de Pater, A. G. Davies, H. G. Roe, T. Fusco, D. Le Mignant, P. Descamps, B. A. Macintosh, and R. Prangé, "High-resolution Keck adaptive optics imaging of violent volcanic activity on Io," Icarus **160**, 124–131 (2002).
55. F. Marchis, "Imagerie à haute résolution angulaire des objets planétaires en optique adaptative. Application au volcanisme de Io," Ph.D. thesis (Université Paul Sabatier (Toulouse III), Toulouse, France, 2000).
56. A.-M. Lagrange, G. Chauvin, T. Fusco, E. Gendron, D. Rouan, M. Hartung, F. Lacombe, D. Mouillet, G. Rousset, P. Drossart, R. Lenzen, C. Moutou, W. Brandner, N. Hubin, Y. Clenet, A. Stolte, R. Schoedel, G. Zins, and J. Spyromilio, "First diffraction limited images at VLT with NAOS and CONICA," in *Instrumental Design and Performance for Optical/Infrared Ground-Based Telescopes*, M. Iye and A. F. M. Moorwood, eds., Proc. SPIE **4841**, 860–868 (2002).
57. R. Lenzen, M. Hartung, W. Brandner, G. Finger, N. N. Hubin, F. Lacombe, A.-M. Lagrange, M. D. Lehnert, A. F. M. Moorwood, and D. Mouillet, "NAOS-CONICA first on-sky results in a variety of observing modes," in *Instrumental Desigtn and Performance for Optical/Infrared Ground-Based Telescopes*, M. Iye and A. F. M. Moorwood, eds., Proc. SPIE **4841**, 944–952 (2002).
58. F. Marchis, R. Prangé, and T. Fusco, "A survey of Io's volcanism by adaptive optics observations in the 3.8 micron thermal band (1996–1999)," J. Geophys. Res. **106**, 33141–33160 (2001).
59. R. R. Howell, J. R. Spencer, J. D. Goguen, F. Marchis, R. Prangé, T. Fusco, B. D. L. , G. J. Veeder, J. A. Rathbun, G. S. Orton, A. J. Groeholski, and J. A. Stansberry, "Ground-based observations of volcanism on Io in 1999 and early 2000," J. Geophys. Res. **106**, 33–129 (2002).
60. F. Roddier, C. Roddier, J. E. Graves, M. J. Northcott, and T. Owen, "Neptune cloud structure and activity: ground based monitoring with adaptive optics," Icarus **136**, 168–172 (1998).
61. J.-M. Conan, T. Fusco, L. M. Mugnier, and F. Marchis, "MISTRAL: Myopic Deconvolution Method Applied to ADONIS and simulated VLT-NAOS Images," ESO Messenger **99**, 38–45 (2000).
62. G. Chauvin, F. Ménard, T. Fusco, A.-M. Lagrange, J.-L. Beuzit, D. Mouillet, and J.-C. Augereau, "Adaptive optics

imaging of MBM 12 association. Seven binaries and edge-on disk in a quadruple system," Astron. Astrophys. **394**, 949–956 (2002).
63. R. Jayawardhana and A. Brandeker, "Discovery of close companions to the nearby young stars HD 199143 and HD 358623," Astrophys. J. Lett. **561**, L111–L113 (2001).
64. G. Chauvin, T. Fusco, A.-M. Lagrange, D. Mouillet, J.-L. Beuzit, M. Thomson, J.-C. Augereau, F. Marchis, C. Dumas, and P. Lowrance, "No disk needed around HD 199143 B," Astron. Astrophys. **394**, 219–223 (2002).
65. A. Coustenis, E. Gendron, O. Lai, J.-P. Véran, J. Woillez, M. Combes, L. Vapillon, T. Fusco, L. Mugnier, and P. Rannou, "Images of Titan at 1.3 and 1.6 micron with adaptive optics at the CFHT," Icarus **154**, 501–515 (2001).
66. S. Gibbard, B. A. Macintosh, C. E. Max, I. de Pater, F. Marchis, H. Roe, D. S. Acton, O. Lai, P. Wizinowitch, P. Stomski, E. F. Young, and C. P. McKay, "Near-infrared observations of Titan from the W. M. Keck Telescope adaptive optics system," manuscript available from the authors, e.g., fmarchis@astron.berkeley.edu.
67. G. Chauvin, M. Thomson, C. Dumas, J.-L. Beuzit, P. Lowrance, T. Fusco, A.-M. Lagrange, B. Zuckerman, and D. Mouillet, "Adaptive optics imaging survey of the Tucana–Horologium association," Astron. Astrophys. **404**, 157–162 (2003).
68. D. Hestroffer, F. Marchis, T. Fusco, and J. Berthier, "Adaptive optics observations of asteroid (216) Kleopatra," Research Note, Astron. Astrophys. **394**, 339–343 (2002).
69. T. Fusco, J.-M. Conan, L. Mugnier, V. Michau, and G. Rousset, "Characterization of adaptive optics point spread function for anisoplanatic imaging. Application to stellar field deconvolution," Astron. Astrophys. Suppl. Ser. **142**, 149–156 (2000).

Annexe L

Article Mugnier et coll., 2001 [A34]

862 J. Opt. Soc. Am. A/Vol. 18, No. 4/April 2001 Mugnier *et al.*

Myopic deconvolution from wave-front sensing

Laurent M. Mugnier, Clélia Robert, Jean-Marc Conan, Vincent Michau, and Sélim Salem

Optics Department, Office National d'Études et de Recherches Aérospatiales, B.P. 72, F-92322 Chatillon cedex, France

Received November 30, 1999; revised manuscript received August 1, 2000; accepted October 19, 2000

Deconvolution from wave-front sensing is a powerful and low-cost high-resolution imaging technique designed to compensate for the image degradation due to atmospheric turbulence. It is based on a simultaneous recording of short-exposure images and wave-front sensor (WFS) data. Conventional data processing consists of a sequential estimation of the wave fronts given the WFS data and then of the object given the reconstructed wave fronts and the images. However, the object estimation does not take into account the wave-front reconstruction errors. A joint estimation of the object and the respective wave fronts has therefore been proposed to overcome this limitation. The aim of our study is to derive and validate a robust joint estimation approach, called myopic deconvolution from wave-front sensing. Our estimator uses all data simultaneously in a coherent Bayesian framework. It takes into account the noise in the images and in the WFS measurements and the available *a priori* information on the object to be restored as well as on the wave fronts. Regarding the *a priori* information on the object, an edge-preserving prior is implemented and validated. This method is validated on simulations and on experimental astronomical data.

OCIS codes: 010.1330, 010.7350, 100.1830, 100.3020, 100.3190, 110.6770.

1. INTRODUCTION

The performance of high-resolution imaging with large optical instruments is severely limited by atmospheric turbulence. Deconvolution from wave-front sensing (DWFS) is a powerful high-resolution imaging technique designed to compensate for the image degradation that is due to atmospheric turbulence. It can be an interesting, light-weight, and low-cost alternative to adaptive optics. It is based on a simultaneous recording of monochromatic short-exposure images and associated wave-front sensor (WFS) data. This technique was originally proposed by Fontanella in 1985.[1] Conventional data processing consists of a sequential estimation of the wave fronts given the WFS data and then of the object given the reconstructed wave fronts and the images. This estimator was tested by Primot *et al.* on laboratory data.[2,3] The first astronomical results were obtained in the early 1990's.[4,5] Primot[3] also gave an analytical expression of the signal-to-noise ratio (SNR) of the technique. DWFS should be more efficient than speckle interferometry for two reasons: DWFS is not limited by speckle noise at high flux,[6] and the SNR is better for extended objects, provided that a bright source is available for the WFS. This second point was confirmed in recent publications.[7,8] The estimator proposed by Primot was, however, shown to be biased[9] since it does not take into account the wave-front reconstruction errors in object estimation. An unbiased estimator can be obtained with additional calibration data recorded on an unresolved star,[10] but its high-flux performance is limited by speckle noise.[11] An alternative approach is to perform a joint estimation of the object and the wave fronts, accounting for the noise both in the images and in the WFS data.[11–14]

The aim of this paper is to propose a robust joint estimator derived in a Bayesian framework that takes advantage of all the available statistical information: the noise statistics in the images and in the WFS as well as the available *a priori* knowledge about the object structure and about the turbulent-phase statistics. This novel method, recently presented in Ref. 15, is called myopic deconvolution from wave-front sensing (MDWFS).

The idea of using WFS data and images jointly can be found in the pioneering work of Schulz.[12] Our contribution is twofold: First, we model our WFS data as slopes deduced from wide-band images, not as narrow-band images. This is mandatory for the processing of low-flux experimental data; indeed, in DWFS all the photons not used for speckle imaging (which is necessarily narrow band) can and should be used for wave-front sensing. Second, we incorporate regularization into the data processing, for both the phase and the object. More specifically, our approach is a joint (object and phase) maximum *a posteriori* (MAP) estimation, which is more robust to noise than an unregularized maximum-likelihood estimation. The prior on the turbulent phase is derived from Kolmogorov statistics. The prior used to regularize the estimation of the object can be taken as Gaussian, in which case the Bayesian interpretation of regularization provides a means to avoid any manual tuning of a regularization parameter. As an alternative, an edge-preserving object prior is also implemented and validated; it is particularly suitable for restoring asteroids or man-made objects such as satellites.

We use the term myopic deconvolution instead of blind deconvolution to underline the fact that during the deconvolution, the phase is neither completely unknown (WFS measurements are used) or assumed to be perfectly known through the WFS measurements.

The outline is as follows. The imaging models for the images and for the WFS are presented in Section 2. In Section 3 we recall the conventional framework of DWFS, which consists of a sequential estimation of (a) the wave

0740-3232/2001/040862-11$15.00

fronts, given the WFS data, and (b) the object, given the wave-front estimates and the images; we also introduce the various terms of the criteria that can be minimized in this sequential estimation for optical wave-front estimation and for edge-preserving object restoration. Then in Section 4 we introduce our myopic-deconvolution approach, which replaces the two previous minimizations (wave-front estimation and object estimation) with a single joint minimization. This scheme is validated by simulations in Section 5 for the case of satellite imaging, and some experimental results are given in Section 6.

2. IMAGING MODEL AND PROBLEM STATEMENT

A linear shift-invariant model is assumed for all M recorded short-exposure images i_t of the observed object o:

$$i_t = o \star h_t + n_t, \quad 1 \leqslant t \leqslant M, \tag{1}$$

where $\star$ denotes the convolution operator, h_t is the instantaneous point-spread function (PSF), at time t, of the system consisting of the atmosphere, of the telescope, and of the detector, and n_t is an additive noise (often predominantly photon noise). The shift-invariance assumption is valid if the field of view is within the isoplanatic patch. If the field of view is larger, then the phase variations in the field should be modeled, taking into account possible wave-front measurements in different directions as well as *a priori* information on the wave-front spatial variations.[16,17]

The integration time of each image is considered to be short enough (typically 10 ms or less) to "freeze" the turbulence. Assuming that scintillation is negligible (near-field approximation), the PSF at time t is completely characterized by the turbulent phase φ_t in the pupil of the instrument:

$$h_t = |\mathrm{FT}^{-1}[P \exp(j\varphi_t)]|^2, \tag{2}$$

where P is the pupil function (1 inside, 0 outside) and FT denotes Fourier transformation. Owing to this relationship, which will be used throughout the paper, the knowledge of the wave front φ_t is equivalent to that of the PSF h_t. The phase is decomposed on the basis of Zernike polynomials[18] $\{Z_k\}$:

$$\varphi_t(r) = \sum_k \phi_t^k Z_k(r), \tag{3}$$

so the unknown phase at time t is the vector $\boldsymbol{\phi}_t = (\phi_t^1, \ldots, \phi_t^k, \ldots)^T$, where T denotes transposition. In practice the summation on k will be limited to some high number K (typically one hundred to a few hundreds) chosen to keep the computational burden reasonable. The turbulent phase in the pupil is considered zero-mean Gaussian with Kolmogorov statistics; its covariance matrix C_ϕ is thus known and given by Noll's formula.[18]

In the following sections we shall consider that the object and the image are sampled on a regular grid, hence a vectorial formulation for Eq. (1):

$$\mathbf{i}_t = \mathbf{o} \star \mathbf{h}_t + \mathbf{n}_t = \mathcal{H}_t \mathbf{o} + \mathbf{n}_t, \tag{4}$$

where $\mathbf{o}$, $\mathbf{i}_t$ and $\mathbf{n}_t$ are the vectors corresponding to the lexicographically ordered object, image, and zero-mean noise, respectively. $\mathcal{H}_t$ is the Toeplitz matrix corresponding to the convolution by the PSF $\mathbf{h}_t$.

The WFS is assumed to be working in a linear domain, which is very reasonable for the Shack–Hartmann (SH) WFS considered in this paper. The WFS data recorded at time t can thus be written as

$$\mathbf{s}_t = \mathcal{D}\boldsymbol{\phi}_t + \mathbf{n}'_t, \tag{5}$$

where $\mathbf{s}_t$ is the vector concatenating the WFS measurements (wave-front slopes in the case of a SH sensor), $\mathbf{n}'_t$ is a zero-mean Gaussian noise (of covariance $C_{n'}$), $\boldsymbol{\phi}_t$ is the turbulent phase vector, and $\mathcal{D}$ is the so-called interaction matrix, which depends on the chosen WFS and on its geometrical configuration as well as on the basis chosen for $\boldsymbol{\phi}$. For example, for a SH WFS and a phase decomposed on the Zernike polynomials, the columns of $\mathcal{D}$ consist of the collection of the responses of the WFS to each Zernike polynomial, which are close to the spatial derivatives of these polynomials.[3]

The problem at hand is to estimate the observed object $\mathbf{o}$ given a set of images $\mathbf{i}_t$ and a synchronously recorded set of wave-front measurements $\mathbf{s}_t$ $(1 \leqslant t \leqslant M)$.

3. SEQUENTIAL ESTIMATION OF WAVE FRONTS AND OBJECT

The conventional processing of DWFS data consists of a sequential estimation of the wave fronts, given the WFS data, and then of the object, given the reconstructed wave fronts (and thus the PSF's) and the images. Both the wave-front reconstruction and the image restoration are ill-posed inverse problems and must be regularized (see Refs. 19 and 20 for reviews on regularization), in the sense that some *a priori* information must be introduced in their resolution in order for the solution to be unique and robust to noise.

A. Wave-Front Reconstruction

Let us first tackle the wave-front reconstruction problem. A common solution for the estimation of a wave front $\boldsymbol{\phi}$, given slope measurements $\mathbf{s}$ made by a SH WFS, is the least-squares one:

$$\hat{\boldsymbol{\phi}} = (\mathcal{D}^T\mathcal{D})^{-1}\mathcal{D}^T\mathbf{s}. \tag{6}$$

This solution can be interpreted as a maximum-likelihood solution if the noise on the WFS is stationary white Gaussian. The matrix $\mathcal{D}^T\mathcal{D}$ is often ill-conditioned or even not invertible, because the number of measurements is finite (twice the number of subapertures N_{sub}), whereas the dimension of the true phase vector $\boldsymbol{\phi}$ is theoretically infinite (in practice a high number K, possibly greater than $2N_{\mathrm{sub}}$). The usual remedy is to reduce the dimension K of the vector space of the unknown $\boldsymbol{\phi}$. This kind of regularization is not very satisfactory, because the choice of the dimension is difficult and somewhat *ad hoc*: The best choice depends in particular on the SNR of the WFS. A more rigorous approach is to turn to a probabilistic solution that makes use of prior information on the turbulent phase. Such a solution is, for instance, the linear minimum-variance estimator proposed by Wallner[21]

(similar to the Wiener filter familiar in image restoration) and first applied to DWFS by Welsh and VonNiederhausern.[22] Another such solution is given by a MAP approach, i.e., by searching for the most likely phase given the measurements and our prior information. When, as in this paper, the noise and the turbulent phase are considered Gaussian, these two approaches are equivalent.[23] The maximization of the posterior probability of the phase is equivalent to the minimization of the neg-log-posterior probability of $\boldsymbol{\phi}$, called J_{MAP}^{ϕ} in the following. With use of Bayes' rule, it is straightforward to show that J_{MAP}^{ϕ} takes the following form:

$$J_{\mathrm{MAP}}^{\phi} = J_s + J_{\phi}, \tag{7}$$

where

$$J_s = \tfrac{1}{2}(\mathbf{s} - \mathcal{D}\boldsymbol{\phi})^T C_{n'}^{-1}(\mathbf{s} - \mathcal{D}\boldsymbol{\phi}), \tag{8}$$

$$J_{\phi} = \tfrac{1}{2}\boldsymbol{\phi}^T C_{\phi}^{-1}\boldsymbol{\phi}. \tag{9}$$

J_s is a term of fidelity to the WFS data and is the opposite of their log likelihood, and J_{ϕ} is a term of fidelity to the prior. The solution is analytical and can be written in matrix form[24] as follows:

$$\hat{\boldsymbol{\phi}}_{\mathrm{MAP}} = (\mathcal{D}^T C_{n'}^{-1}\mathcal{D} + C_{\phi}^{-1})^{-1}\mathcal{D}^T C_{n'}^{-1}\mathbf{s}. \tag{10}$$

Note that there is an equivalent expression for this solution that involves the inversion of a smaller matrix when the phase $\boldsymbol{\phi}$ is expanded on more modes than there are WFS measurements:

$$\hat{\boldsymbol{\phi}}_{\mathrm{MAP}} = C_{\phi}\mathcal{D}^T(\mathcal{D}C_{\phi}\mathcal{D}^T + C_{n'})^{-1}\mathbf{s}. \tag{11}$$

This phase estimate is more accurate than the least-squares one of Eq. (6) and in addition does not require an *ad hoc* tuning of the dimension K of the unknown phase. To the best of our knowledge, the phase estimate has so far always been used as the true phase [for computing the PSF's by use of Eq. (2)] for the image restoration.[8–10] In contrast, the phase estimate of Eq. (10) will be used as a *starting point* for our myopic method, and the corresponding criterion of Eq. (7) will be *part* of the criterion derived in the myopic method. We delay the presentation of this myopic method to the following section, and for the time being we consider, as is classically done, that the wave fronts estimated by the MAP approach above are the true ones.

B. Multiframe Image Restoration

Let us now turn to the image restoration problem, in the classical setting where the PSF's are known. Most deconvolution techniques boil down to the minimization (or maximization) of a criterion. The first issue is the definition of a suitable criterion for the given inverse problem. The second issue is then to find the position of the criterion's global minimum, which is defined as the solution. In some rare cases (when the criterion to be minimized is quadratic) the solution is given by an analytical expression (e.g., Wiener filtering), but most of the time one must resort to an iterative numerical method to solve the problem. For our applications we use a conjugate-gradient method.

Similarly to wave-front reconstruction, a common approach for the image restoration is to use a (multiframe) least squares. The solution, i.e., the minimum of this criterion, is analytical and is the multiframe inverse filter proposed in DWFS by Primot *et al.*[3] This approach has a maximum-likelihood interpretation when the noise can be assumed to be stationary white Gaussian. It leads to unacceptable noise amplification for high noise levels and must therefore be regularized.

A natural way to regularize the inversion is functional regularization. It consists in defining the solution as the minimum of a compound criterion with two terms, say, $J_i + \lambda J_o$. Term J_i of the criterion enforces fidelity to the data (it can be, e.g., a least squares) while term J_o expresses fidelity to some prior information about the solution. Two choices must be made: One is the regularization functional, i.e., the expression of J_o, and the other is the regularization parameter λ, a scalar that adjusts the trade-off between the two terms.

The choice of the regularization parameter can be made automatically during the restoration process in a constrained least-squares formulation of the regularization.[25] Yet the regularization functional remains to be chosen by the user in a somewhat arbitrary fashion.

An alternative is to use the Bayesian interpretation of regularization, which can provide "natural" regularization functionals. This is the approach taken in this paper: The object is endowed with an *a priori* distribution $p(\mathbf{o})$, and Bayes' rule combines the likelihood of the M images $p(\{\mathbf{i}_t\}_{t=1}^{M}|\mathbf{o},\{\boldsymbol{\phi}_t\}_{t=1}^{M})$ with this *a priori* distribution into the *a posteriori* probability distribution $p(\mathbf{o}|\{\mathbf{i}_t\}_{t=1}^{M}, \{\boldsymbol{\phi}_t\}_{t=1}^{M})$:

$$p(\mathbf{o}|\{\mathbf{i}_t\}_{t=1}^{M}, \{\boldsymbol{\phi}_t\}_{t=1}^{M}) \propto p(\{\mathbf{i}_t\}_{t=1}^{M}|\mathbf{o}, \{\boldsymbol{\phi}_t\}_{t=1}^{M}) \times p(\mathbf{o}). \tag{12}$$

We shall make two assumptions in the following: First, that the noise is independent between images and second, that the time between two successive data recordings is greater than the typical evolution time of turbulence. With these two reasonable assumptions, the likelihood of the set of images can be rewritten as the product of the likelihoods of the individual images, each of the conditioned only by the object and by the true phase at the same time. The restored object is defined as the most probable one given the data, i.e., the one that minimizes:

$$J_{\mathrm{MAP}}^{o}(\mathbf{o}) = \sum_{t=1}^{M} J_i(\mathbf{o}; \boldsymbol{\phi}_t, \mathbf{i}_t) + J_o(\mathbf{o}), \tag{13}$$

where

$$J_i(\mathbf{o}; \boldsymbol{\phi}_t, \mathbf{i}_t) = -\ln p(\mathbf{i}_t|\mathbf{o}, \boldsymbol{\phi}_t), \tag{14}$$

$$J_o(\mathbf{o}) = -\ln p(\mathbf{o}). \tag{15}$$

J_i is the neg-log-likelihood of image number t and J_o is the neg-log-prior probability of the object. The minimization of this criterion is performed on $\mathbf{o}$ only, and the $\boldsymbol{\phi}_t$'s are a reminder of the dependency of the criterion on the phases, assumed here to be known. The forms taken by

criteria J_i and J_o depend on the statistical assumptions made on the noise and on the object, respectively, and are discussed next.

1. Noise Statistics

If the noise is zero mean Gaussian with covariance matrix C_n, then J_i is quadratic for any image $\mathbf{i}$:

$$J_i(\mathbf{o};\ \boldsymbol{\phi}, \mathbf{i}) = \tfrac{1}{2}(\mathbf{h} \star \mathbf{o} - \mathbf{i})^T C_n^{-1}(\mathbf{h} \star \mathbf{o} - \mathbf{i}), \quad (16)$$

which depends on $\boldsymbol{\phi}$ through $\mathbf{h}$ [see Eq. (2)]. In particular, if the noise is, in addition, stationary and white, of variance σ_n^2, then Eq. (16) reduces to the familiar least squares:

$$J_i(\mathbf{o};\ \boldsymbol{\phi}, \mathbf{i}) = \frac{1}{2\sigma_n^2}\|\mathbf{h} \star \mathbf{o} - \mathbf{i})\|^2 = \frac{1}{2\sigma_m^2}\sum_{l,m} |[\mathbf{h} \star \mathbf{o}](l, m) - \mathbf{i}(l, m)|^2. \quad (17)$$

In astronomical imaging, the noise is often predominantly photon noise, which follows Poisson statistics. One possibility is then to derive the true MAP criterion for photon noise statistics, which is the neg-log-likelihood of the Poisson law. In this paper J_i is taken for simplicity as the least-squares term of Eq. (17) with a uniform noise variance equal to the mean number of photons per pixel. This can be considered as a first approximation of the photon noise in the case of a rather bright and extended object.

2. Object Prior

The choice of a Gaussian prior probability distribution for the object can be justified from an information theory standpoint as being the least informative, given the first two moments of the distribution. In this case a reasonable model of the object's power spectral density (PSD) can be found[26] and used to derive the criterion J_o. This is more satisfactory than an *ad hoc* regularization functional such as the identity or the traditional Laplacian and gives better object estimates.[26] In addition, because J_o is derived from a probability distribution, there is no regularization parameter (scaling factor between the J_i's and J_o) to be adjusted. Finally, the solution is analytical and is a multiframe Wiener filter.[27]

The disadvantage of a Gaussian prior, especially for objects with sharp edges such as asteroids or artificial satellites, is that it tends to oversmooth edges. A possible remedy is to use an edge-preserving prior such as an L_2–L_1 criterion, quadratic for small gradients and linear for large ones.[28] The quadratic part ensures a good smoothing of the small gradients (i.e., of noise), and the linear behavior cancels the penalization of large gradients (i.e., of edges).[29] Here we use a function that is an isotropic version of the expression suggested by Rey[30] in the context of robust estimation, used by Brette and Idier[31] for image restoration, and recently applied to imaging through turbulence[32]:

$$J_o(\mathbf{o}) = \mu\delta^2 \sum_{l,m}\left[\left(\frac{\nabla\mathbf{o}(l, m)}{\delta}\right) - \ln\left(1 + \frac{\nabla\mathbf{o}(l, m)}{\delta}\right)\right], \quad (18)$$

where $\nabla\mathbf{o}(l, m) = [\nabla_x\mathbf{o}(l, m)^2 + \nabla_y\mathbf{o}(l, m)^2]^{1/2}$, and $\nabla_x\mathbf{o}$ and $\nabla_y\mathbf{o}$ are the object finite-difference gradients along x and y, respectively.

The global factor μ and the threshold δ have to be adjusted according to the noise level and the structure of the object. This is currently done by hand but an automatic procedure is under study. This function is convex, as is the global criterion, which ensures uniqueness and stability of the solution with respect to noise and also justifies the use of a gradient-based method for the minimization. Note that if the object is a stellar field, then a stronger prior can be used, namely, the fact that the unknown object is a collection of Dirac delta functions.[33]

4. JOINT ESTIMATION OF WAVE FRONTS AND OBJECT

A. Motivation

The conventional processing DWFS data consists of a sequential estimation of the wave fronts given the WFS data and then of the object given the images and the reconstructed wave fronts, which are regarded as the true ones. So, obviously, the information about the wave fronts is gathered only in the WFS data, not in the images.

Yet the wave-front estimates are inevitably noisy, so it would be useful to exploit both the WFS data and the images in the wave-front estimation. And there *is* some exploitable information about the wave front (or the PSF) in a short-exposure turbulence-degraded image. To prove this point, it is enough to recall that multiframe blind deconvolution for such images (blind meaning without a WFS) is feasible, at least at high SNR's; indeed the PSF reparameterization of Eq. (2) by the phase, which was proposed by Schulz,[34] is a strong constraint that allows this technique to work in practice, even on experimental data.[34,35]

A drawback of blind criteria is that they usually exhibit local minima, and the PSF reparameterization by the phase is probably not a strong enough constraint to give the blind-deconvolution problem a unique solution. So the WFS data should definitely not be thrown away but rather used in conjunction with the images.

Our aim is thus to estimate the object and the PSF's while taking into account all the measurements (WFS data and images) simultaneously in a coherent framework.

B. Myopic-Deconvolution Approach

The myopic-deconvolution approach consists in a joint estimation of the object $\mathbf{o}$ and the turbulent phases $\boldsymbol{\phi}_t$ that are the most likely, given the images $\mathbf{i}_t$, the WFS data $\mathbf{s}_t$ and the *a priori* information on $\mathbf{o}$ and the $\boldsymbol{\phi}_t$'s. If we use Bayes' rule in a way similar to its use in Subsection 3.B and the independence assumptions therein, the joint posterior probability distribution of the object and the phases is

866 J. Opt. Soc. Am. A/Vol. 18, No. 4/April 2001 Mugnier *et al.*

Table 1. Block Diagram of the Algorithm Used for Myopic Deconvolution from Wave-Front Sensing

Step	Operation Performed	Implementation
1. Initialization	$\boldsymbol{\phi}_t^0 = \hat{\boldsymbol{\phi}}_t^{\mathrm{MAP}} = (\mathcal{D}^T\mathcal{C}_{n'}^{-1}\mathcal{D} + \mathcal{C}_\phi^{-1})^{-1}\mathcal{D}^T\mathcal{C}_{n'}^{-1}\mathbf{s}_t \ \ \forall t$	Matrix Multiplication
	$\mathbf{o}^0 = \arg\min_{\mathbf{o}}\left(\frac{1}{2\sigma_n^2}\sum_t \|\mathbf{h}(\boldsymbol{\phi}_t^0) \star \mathbf{o} - \mathbf{i}_t\|^2 + J_o(\mathbf{o})\right)$	One conjugate-gradient minimization
2. qth phase iteration	$\boldsymbol{\phi}_t^q = \arg\min_{\boldsymbol{\phi}}\left(\frac{1}{2\sigma_n^2}\|\mathbf{h}(\boldsymbol{\phi}) \star \mathbf{o}^{q-1} - \mathbf{i}_t\|^2 + \frac{1}{2}(\mathbf{s}_t - \mathcal{D}\boldsymbol{\phi})^T\mathcal{C}_{n'}^{-1}(\mathbf{s}_t - \mathcal{D}\boldsymbol{\phi}) + \frac{1}{2}\boldsymbol{\phi}^T\mathcal{C}_\phi^{-1}\boldsymbol{\phi}\right)$	M Partial conjugate-gradient minimizations
3. qth object iteration	$\mathbf{o}^q = \arg\min_{\mathbf{o}}\left(\frac{1}{2\sigma_n^2}\sum_t \|\mathbf{h}(\boldsymbol{\phi}_t^q) \star \mathbf{o} - \mathbf{i}_t\|^2 + J_o(\mathbf{o})\right)$	One partial conjugate-gradient minimization
4. Stopping rule	if $\|\mathbf{o}^q - \mathbf{o}^{q-1}\| > 10^{-7}\|\mathbf{o}^q\|$ or if $\exists t;\ \|\boldsymbol{\phi}_t^q - \boldsymbol{\phi}_t^{q-1}\| > 10^{-7}\|\boldsymbol{\phi}_t^q\|$ then go to (2), else stop.	

$$
\begin{aligned}
p(\mathbf{o}, \{\boldsymbol{\phi}_t\}_{t=1}^M | \{\mathbf{i}_t\}_{t=1}^M, \{\mathbf{s}_t\}_{t=1}^M) &\propto p(\{\mathbf{i}_t\}_{t=1}^M, \{\mathbf{s}_t\}_{t=1}^M | \mathbf{o}, \{\boldsymbol{\phi}_t\}_{t=1}^M) \times p(\mathbf{o}) \times p(\{\boldsymbol{\phi}_t\}_{t=1}^M) \\
&\propto \prod_{t=1}^M p(\mathbf{i}_t | \mathbf{o}, \boldsymbol{\phi}_t) \times p(\mathbf{o}) \times \prod_{t=1}^M p(\mathbf{s}_t | \boldsymbol{\phi}_t) \\
&\quad \times \prod_{t=1}^M p(\boldsymbol{\phi}_t). \qquad (19)
\end{aligned}
$$

So the joint MAP estimates $(\hat{\mathbf{o}}, \{\hat{\boldsymbol{\phi}}_t\})$ are the ones that minimize:

$$
\begin{aligned}
J_{\mathrm{MAP}}(\mathbf{o}, \{\boldsymbol{\phi}_t\}) &= \sum_{t=1}^M J_i(\mathbf{o}, \boldsymbol{\phi}_t; \mathbf{i}_t) + J_o(\mathbf{o}) \\
&\quad + \sum_{t=1}^M J_s(\boldsymbol{\phi}_t; \mathbf{s}_t) + \sum_{t=1}^M J_\phi(\boldsymbol{\phi}_t), \qquad (20)
\end{aligned}
$$

where

- The J_i's are terms of fidelity to image data; $J_i(\mathbf{o}, \boldsymbol{\phi}_t; \mathbf{i}_t)$ is the neg-log-likelihood of the tth image; for a Gaussian noise, $J_i(\mathbf{o}, \boldsymbol{\phi}_t; \mathbf{i}_t) = (1/2)(\mathbf{h}_t \star \mathbf{o} - \mathbf{i}_t)^T C_n^{-1}(\mathbf{h}_t \star \mathbf{o} - \mathbf{i}_t)$, which is the same as Eq. (16) except that it is now a function of the object *and* the phases;
- $J_o(\mathbf{o})$ is the object prior, which can be taken either as quadratic (with a simple parametric model for the object PSD[26]) or as an edge-preserving criterion [see Eq. (18)];
- The J_s's are terms of fidelity to WFS data; under our assumptions they are quadratic and given by Eq. (8);
- The J_ϕ's are the phase priors; for Kolmogorov statistics they, too, are quadratic and given by Eq. (9).

C. Implementation of the Myopic Deconvolution

The criterion of Eq. (20) is minimized numerically to obtain the joint MAP estimate for the object $\mathbf{o}$ and the phases $\boldsymbol{\phi}_t$. The minimization is performed on the unknown object $\mathbf{o}$ and on the unknown phases $\boldsymbol{\phi}_t$ by a fast conjugate-gradient method.[36] Because the criterion is continuously differentiable, the convergence of the conjugate-gradient method to a stationary point (in practice, to a local minimum) is guaranteed.[37] We have found that the convergence is faster if the descent direction is reinitialized regularly (typically every ten iterations for the sizes of our images and phase vectors); this result can be attributed to the fact that the criterion to be minimized is not quadratic. This modified version of the conjugate gradient is known as the *partial conjugate-gradient method* and has similar convergence properties (see Sec. 8.5 of Ref. 38).

A practical but severe minimization problem may occur if one simply stacks together the object and the phases into a vector of unknowns, because the gradients of the criterion with respect to the object and to the phase may have very different orders of magnitude. One solution can be to scale the unknown object. The scaling must be adjusted so that the gradients with respect to the object and to the phase will be of comparable magnitudes.[39] We have found that it is safer and faster to split the minimization into two blocks and to alternate between minimizations on the object for the current phase estimates and minimizations on the phases for the current object estimate, so that this scaling problem is avoided. Additionally, as mentioned above, we use a partial conjugate-gradient method; i.e., we perform a fixed number of conjugate-gradient iterations within each block (object or phases) and then switch to the other block.

Finally, a careful inspection of the criterion J_{MAP} of Eq. (20) allows a significant reduction of the computational burden: First, only the $M + 1$ first terms of J_{MAP} depend on the object, so the last $2M$ terms of the criterion need not be computed when minimizing on the object. Second, when J_{MAP} is minimized on the phases $\{\boldsymbol{\phi}_t\}_{t=1}^M$ for the current object estimate, the minimization can be decoupled into M minimizations, each being performed on one single phase screen $\boldsymbol{\phi}_t$ and involving only three terms $(J_i + J_s + J_\phi)$.

As all numerical minimizations do, this one needs an initialization point or initial guess; in order both to speed up the descent and to avoid, in practice, the local minima associated with joint criteria, we use as an initial guess the MAP estimates obtained by the sequential processing described in Section 3. More precisely, the initialization phases are those obtained by a MAP processing of the sole WFS data, i.e., by the minimization of the two right-hand

terms of Eq. (20); these phases are given analytically by Eq. (10). The initialization object is the one obtained by a MAP processing of the sole images given the initialization phases, i.e., by the minimization of the two left-hand terms of Eq. (20); this object has an analytical expression when both the object and the noise are assumed to be Gaussian (it is the multiframe Wiener filter applied to the image sequence).

The minimization is not stopped by hand but rather when the estimated object and phases no longer evolve (i.e., when their evolution from one iteration to the next is close to machine precision). The metric used to asses the quality of the object estimate is the standard mean square error (MSE) between the true object and the estimated one, expressed in photons per pixel. We have verified that the estimated objects are not shifted by more than a fraction of pixel with respect to the true object, so the MSE's mentioned below are a good measure of the restored object's quality.

The conjugate-gradient routine needs to repeatedly compute the criterion and its gradients with respect to the object $\mathbf{o}$ and to the phases $\boldsymbol{\phi}_t$. These gradients can be computed analytically; they are given in Appendix A for completeness. The whole algorithm is summarized in the block diagram of Table 1. In the next section we validate the proposed myopic approach on simulated data.

5. VALIDATION BY SIMULATIONS

A set of 100 noisy images and associated wave-front measurements are simulated. The 100 turbulent wave fronts are obtained with a modal method[40]: The phase is expanded on the Zernike polynomial basis, and it is given Kolmogorov statistics.[18] The strength of the turbulence corresponds to a ratio D/r_0 of the telescope diameter to the Fried parameter equal to 10.

Each of these turbulent wave fronts leads to a 128 $\times$ 128 Shannon-sampled short-exposure image with use

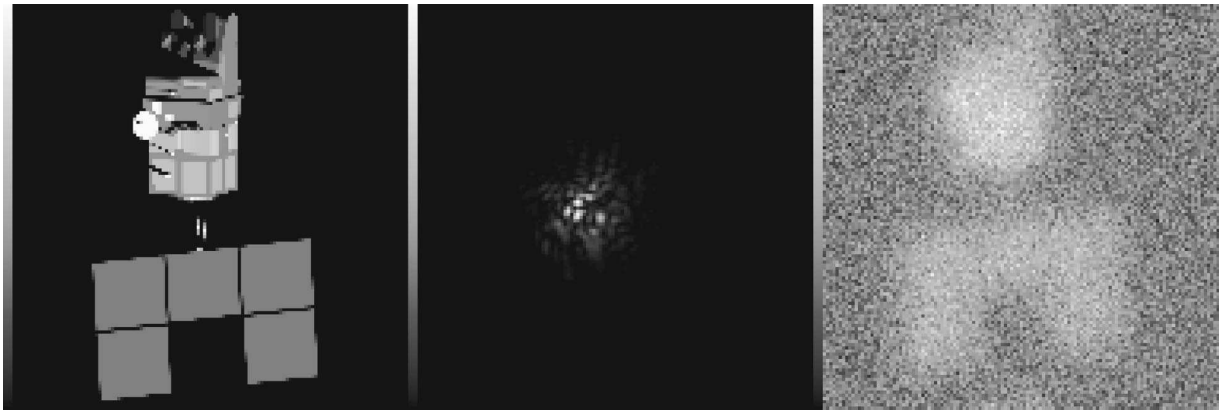


Fig. 1. Original object (SPOT satellite, left) one of the 100 turbulent PSF's ($D/r_0 = 10$, center) and corresponding noisy image (flux of 10^4 photons, right).

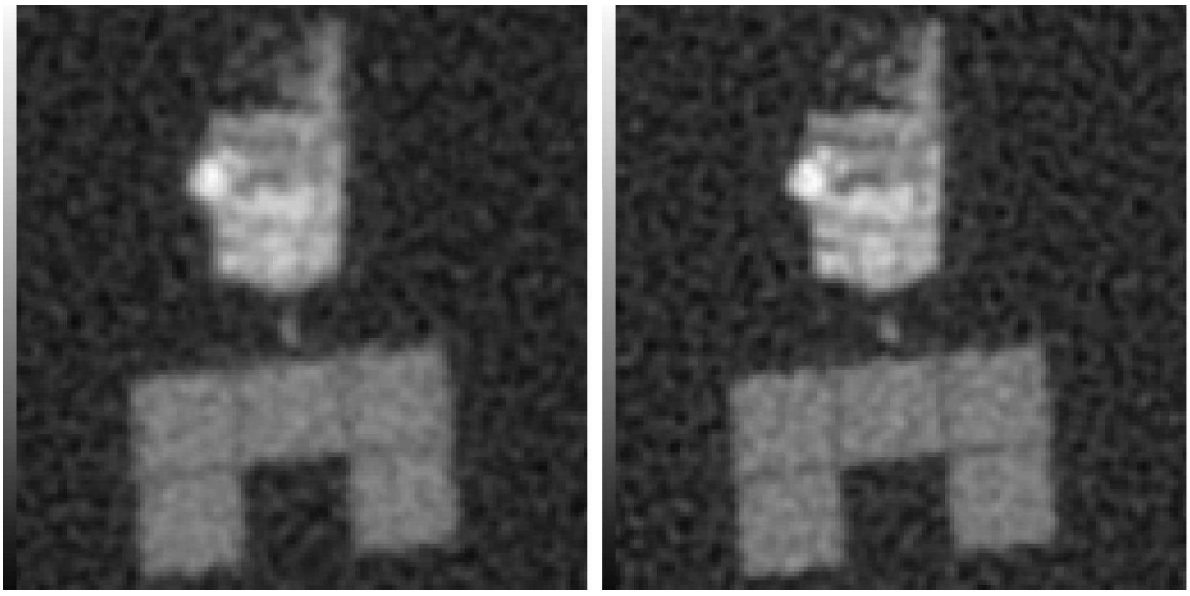


Fig. 2. Restored object with conventional method (MAP wave-front estimation followed by a multiframe Wiener filter, left), to be compared with the restored object by myopic deconvolution (right). In both cases the same Gaussian prior is used for the object. The MSE is 0.48 photon (left) and 0.45 photon (right).

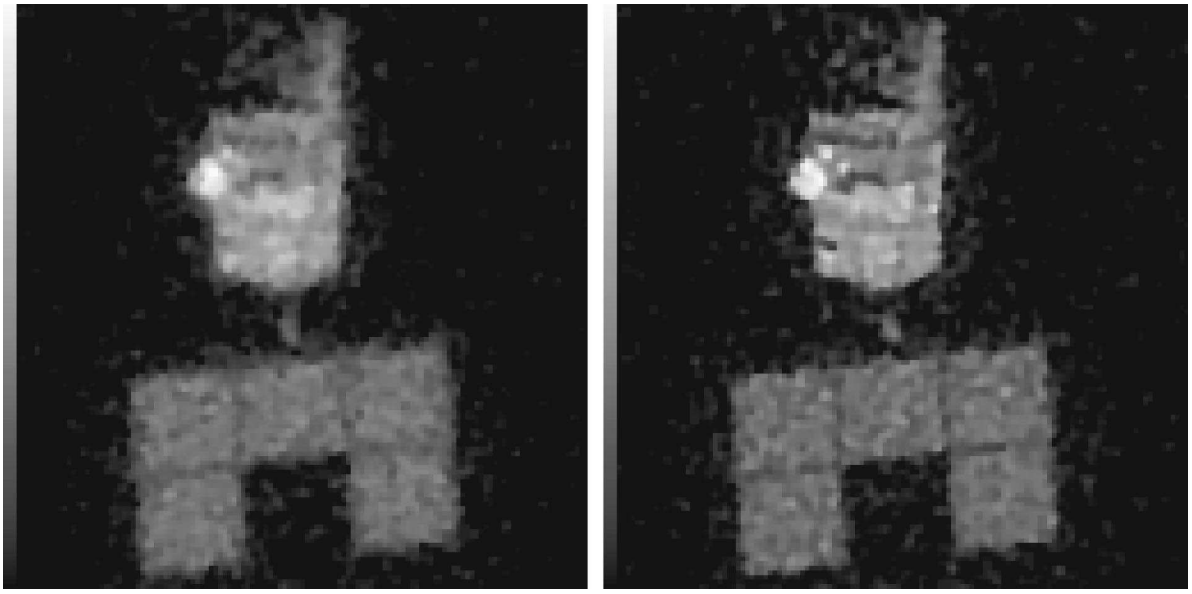


Fig. 3. Restored object with conventional method (left): MAP wave-front estimation followed by a multiframe edge-preserving restoration. Restored object by myopic deconvolution (right). In both cases the same edge-preserving prior is used for the object, with an additional positivity constraint. The MSE is 0.45 photon (left) and 0.39 photon (right).

of Eqs. (1) and (2). The image noise is a uniform Gaussian noise with a variance equal to the mean flux, or $10^4/(128 \times 128) = 0.61$ photons/pixel. It approximates the photon noise for a uniform distribution of a 10^4 photon total flux over the image sensor. Figure 1 shows the original object (a model of the SPOT satellite), one of the 100 turbulence PSF's, and the corresponding noisy image.

The WFS is of SH type with 20×20 subapertures, of which 308 are useful, and without a central obscuration. A white Gaussian noise is added to the computed slopes [see Eq. (5)] so that the SNR on the measured slopes, defined as the ratio of the turbulence slope variance over the noise variance, is 1.

The Zernike expansion of the phase is limited to 190 Zernike polynomials (radial degree 18), only to keep the computational burden reasonably small; indeed, with the SNR chosen for the WFS, it can be shown that the reconstructed phase modes and the reconstructed noise have the same level for the 55th polynomial (radial degree 9).

Figure 2 compares the conventional sequential estimation and our myopic estimation for the same Gaussian object model with the same PSD. On the left, the conventional restoration, consisting of a MAP wave-front estimation followed by a multiframe Wiener filter, gives a MSE of 0,48 photon. On the right, the myopic joint estimation gives a MSE on the object of 0.45 photon. Several details are visibly better restored on the myopic restoration, in particular the two little bright rectangles near the dish and the separations between the solar panels. Figure 3 shows the same comparison between the conventional and the myopic restorations (left and right, respectively), but with the edge-preserving prior of Eq. (18) plus a positivity constraint instead of the Gaussian prior. The MSE is 0.45 photon and 0.39 photon, respectively. A particularly noticeable difference between these two images is the thin shadow on the lower left part of the satellite body, which is visible only in the edge-preserving myopic

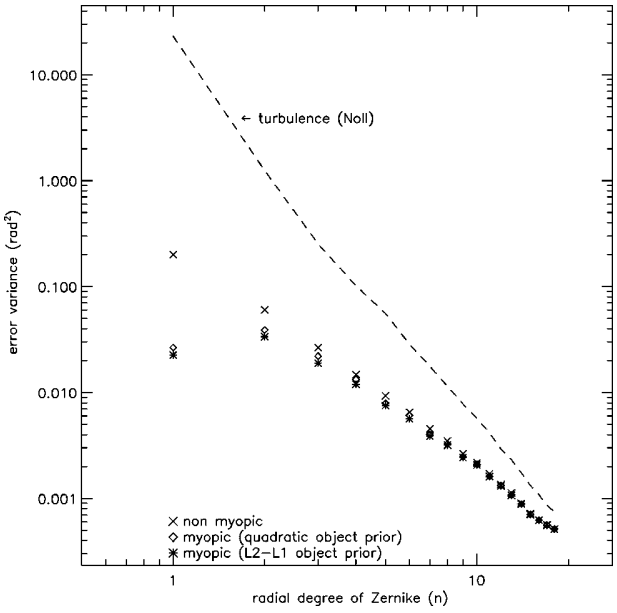


Fig. 4. Comparison of the reconstructed phase errors, as a function the Zernike mode radial degree, for the conventional (MAP) and the myopic (join MAP) restoration schemes. The turbulent-phase variance is shown for comparison (dashed line).

restoration. In both cases the images restored with our myopic method (right side of Figs. 2 and 3) are less blurred than the ones restored with the sequential scheme (left side of the same figures). And the two restorations obtained with the edge-preserving prior are also crisper and closer to the true object than are their quadratic counterparts. The restorations shown in Fig. 3 were obtained with respectively 210 s (left) and 1450 s (right) of CPU time on a Sun Sparc. Processing 100 frames with MDWFS therefore leads to a quite reasonable computation time.

It is remarkable that, simultaneously with the better object estimation, the myopic deconvolution provides a better estimation of the phase. We believe that the latter is due to the fact that all the available data, images included, are used for the phase estimation. A more quantitative explanation of this improvement remains to be done and is worth further study. This estimation improvement is shown in Fig. 4; the tip-tilt, the defocus, and the astigmatism happen to be particularly better estimated than by use solely of WFS data. One can note that the use of a prior that is well suited to the object (an L_2–L_1 criterion for an object that is known to have sharp edges) enhances not only the object restoration but also the phase estimation.

6. EXPERIMENTAL RESULTS

We have applied our MDWFS method to experimental images of the binary star Capella taken at the 4.2-m William Herschel Telescope, located at La Palma in the Canary Islands. Ten short-exposure images of size 128×128 were taken on November 8, 1990, at 4:00 universal time and processed. The exposure time is 5 ms on both the imaging and the wave-front sensing channel. Note that for fainter objects it has been shown that in speckle imaging it can be useful to increase the exposure time so as to obtain a better SNR in the images,[41] even though the speckles get blurred. Preliminary internal studies tend to show that this is not the case for DWFS, because the wave-front measurements degrade quickly when the exposure time is greater than the evolution time of turbulence.

The experimental conditions were the following: A flux of 67,500 photons per frame and an estimated D/r_0 of 13. The imaging wavelength was 0.66 μm, which gives a diffraction-limit resolution λ/D of 32.10^{-3} arc sec. The WFS is of SH type with 29×29 subapertures, of which 560 are useful because of the telescope's 29% central obscuration. The estimated WFS SNR is 5, which gives the WFS noise level required in J_s. Note that the noise level can be artificially increased to account for noncommon-path aberrations in the optical setup. When this noise level goes to infinity, the algorithm becomes a multiframe regularized blind deconvolution. Figure 5 shows one of the ten processed short exposures (left) and the corresponding computed long exposure obtained by adding the ten short exposures (right); the latter image illustrates the loss in resolution associated with the long exposure.

The left panel of Fig. 6 shows the result of conventional estimation: A multiframe Wiener filter was applied to the images, with wave front estimates obtained by a MAP reconstruction from the slopes. The PSD is taken as a constant, as is natural for a spike-like object, and the wave front is reconstructed on the first 190 Zernike polynomials. The binary nature of the star is not clearly visible. The center panel of Fig. 6 shows an improved conventional estimation, in which we have used the same flat PSD but added a positivity constraint to the restoration. The binary nature of the star begins to be visible, but the binary star is still embedded in strong fluctuations.

The right panel of Fig. 6 shows the result of our MDWFS method: the object and the wave fronts are jointly estimated by minimizing the criterion of Eq. (20); the prior used is the same flat PSD with the positivity constraint, and the starting point of the minimization is the estimate on the center. There is a clear gain in image quality with the MDWFS method. In particular, the two components of Capella are clearly resolved, and the estimated angular separation is 57.10^{-3} arc sec, in good agreement with the one predicted by orbit data at the date of the experiment (55.10^{-3} arc sec).

The PSD used in the regularization is the same for the conventional and for the myopic estimations: It is deduced from the average flux in the images and is not tuned manually; one may consider underestimating the PSD (i.e., performing overregularization) to make the con-

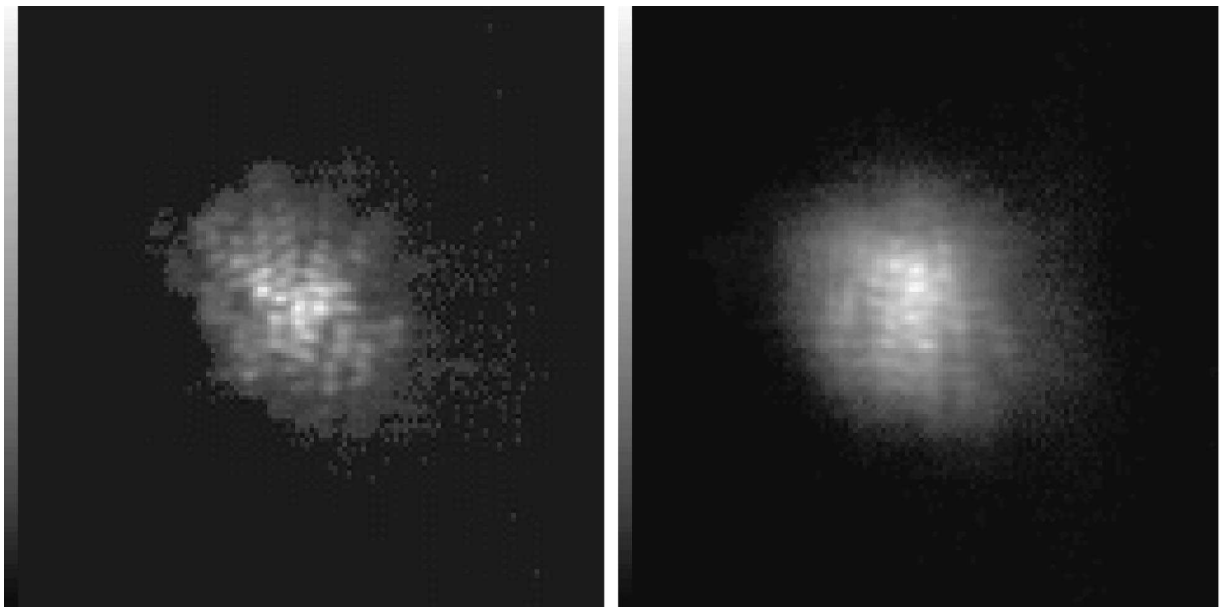


Fig. 5. Experimental short-exposure image of Capella taken on November 8, 1990 (left); corresponding long exposure (average of ten short exposures, right).

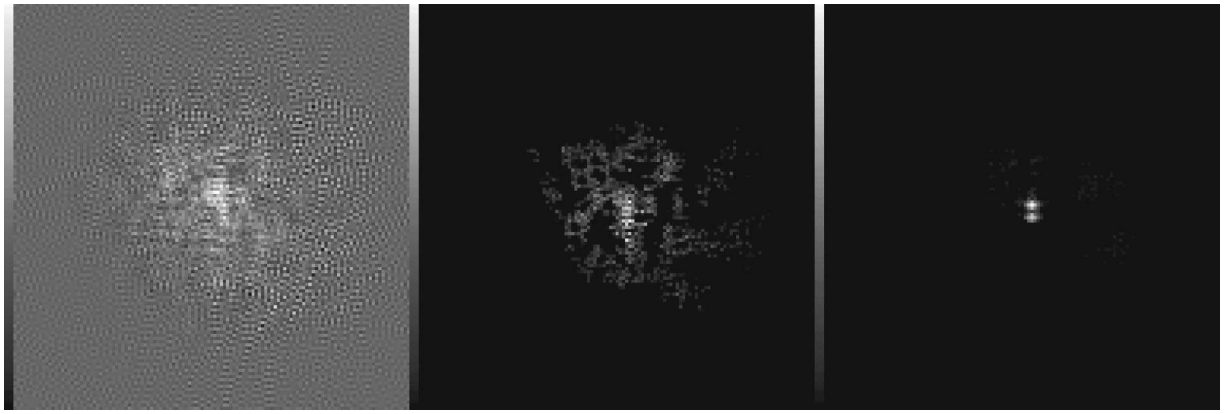


Fig. 6. Restored object with conventional methods: MAP wave-front estimation followed by a multiframe Wiener filter (left); MAP wave-front estimation followed by a quadratic restoration with a positivity constraint (center), to be compared with the object restored by the myopic deconvolution (right), with use of the same quadratic regularization and the same positivity constraint.

ventional estimation smoother, but no manual tuning of the PSD can make the conventionally restored images both sharp and free from spurious ripples.

The restoration of Capella without the use of WFS data by use of the bispectrum has also been reported,[42] but such higher-order methods usually require many more images to produce a result of similar quality (typically 100 images even for very simple objects such as a binary star).

The conventional restoration with positivity constraint (center panel of Fig. 6) required 40 s of CPU time, and the myopic restoration (right part of the figure) took 110 s, which remains fairly fast.

7. CONCLUSION

We have presented a novel approach for deconvolution from wave-front sensing, called myopic deconvolution form wave-front sensing (MDWFS). This approach consists in a joint estimation of the object of interest and the unknown turbulent phases; it considers all the data (WFS slopes and images) simultaneously in a coherent Bayesian framework. This method takes into account the noise in the images and in the wave-front sensor measurements, as well as the available prior information on the object to be restored and on the turbulent phases.

This approach has been validated on simulations and has led to a better object estimation than that obtained with the sequential processing of the WFS data and the images. An edge-preserving object prior has been implemented and should be very effective for asteroids or man-made objects such as satellites. An initial experimental astronomical application of MDWFS on Capella has also been presented.

APPENDIX A: DERIVATION OF THE GRADIENTS WITH RESPECT TO OBJECT AND PHASE PARAMETERS

The image deconvolutions shown in this paper, whether conventional or myopic, are performed by means of a conjugate-gradient routine. This routine needs to repeatedly compute the considered criterion and its gradients with respect to the object $\mathbf{o}$ (and with respect to the phases $\boldsymbol{\phi}_t$ in the case of myopic deconvolution). The gradients of all the terms of the criteria considered here are given below.

1. Gradients of the Image Data Term

The gradient of the least-squares criterion $J_i(\mathbf{o}, \boldsymbol{\phi}_t; \mathbf{i}_t)$ with respect to the unknown object is easily computed in matrix form and can be expressed in Fourier space for implementation by fast Fourier transform:

$$\frac{\mathrm{d}J_i}{\mathrm{d}\mathbf{o}} = \frac{1}{\sigma_n^2}\mathcal{H}_t^T(\mathcal{H}_t\mathbf{o} - \mathbf{i}_t) = \frac{1}{\sigma_n^2}\,\mathrm{FT}^{-1}(\tilde{\mathbf{h}}_t^*(\tilde{\mathbf{h}}_t\tilde{\mathbf{o}} - \tilde{\mathbf{i}}_t)), \tag{A1}$$

where the tilde and $\mathrm{FT}^{-1}(\cdot)$ denote Fourier transformation and its inverse, respectively.

The gradient of the least-squares criterion $J_i(\mathbf{o}, \boldsymbol{\phi}_t; \mathbf{i}_t)$ with respect to the unknown phase $\boldsymbol{\phi}_t$ is done in three steps. The derivation begins by calculating the gradient of J_i with respect to the PSF $\mathbf{h}_t$:

$$\frac{\mathrm{d}J_i}{\mathrm{d}\mathbf{h}_t} = \frac{1}{\sigma_n^2}\mathrm{FT}^{-1}(\tilde{\mathbf{o}}^*(\tilde{\mathbf{h}}_t\tilde{\mathbf{o}} - \tilde{\mathbf{i}}_t)). \tag{A2}$$

It then consists in calculating the gradient of the PSF with respect to the phase $\varphi_t(l, m)$ at pixel (l, m) in the pupil and applying the chain rule. The result, already derived in Ref. 35 in a slightly different form, is

$$\frac{\mathrm{d}J_i}{\mathrm{d}\varphi_t(l,m)} = \frac{2}{A}\mathcal{I}\left\{P(l,m)\exp[-j\varphi_t(l,m)] \times \left[\mathrm{FT}\left(\frac{\mathrm{d}J_i}{\mathrm{d}\mathbf{h}_t}\right) \star P\exp(j\varphi_t)\right](l,m)\right\}, \tag{A3}$$

where A is the area of the pupil [numerically, the number of pixels where $P(l,m) = 1$], $\mathcal{I}$ denotes the imaginary part of a complex number, and $\star$ denotes a convolution operator.

Finally, because we expand the unknown phases on the Zernike basis, we need the gradient of J_i with respect to the Zernike coefficients ϕ_t^k of phase ϕ_t; it is easily deduced from the previous expression:

$$\frac{\mathrm{d}J_i}{\mathrm{d}\phi_t^k} = \sum_{l,m} \frac{\mathrm{d}J_i}{\mathrm{d}\varphi_t(l,m)} Z_k(l,m)$$

$$= \frac{2}{A} \sum_{l,m} Z_k(l,m) \mathcal{I}\left\{ P(l,m) \exp[-j\varphi_t(l,m)] \times \left[\mathrm{FT}\left(\frac{\mathrm{d}J_i}{\mathrm{d}\mathbf{h}_t} \right) \star P \exp(j\varphi_t) \right](l,m) \right\}. \quad \text{(A4)}$$

For a Poisson model of the image noise, the gradient of the corresponding criterion J_i could be derived similarly.[43]

2. Gradient of the Object Prior Term

The gradient of the edge-preserving prior of Eq. (18) can be derived and implemented by using the finite-difference operators $(\nabla_x \mathbf{o})(l,m) = o(l,m) - o(l-1,m)$ and $(\nabla_y \mathbf{o})(l,m) = o(l,m) - o(l,m-1)$. One can show that

$$\frac{\mathrm{d}J_o}{\mathrm{d}\mathbf{o}} = \mu \left[\nabla_x^T \left(\frac{\nabla_x \mathbf{o}}{1 + \sqrt{(\nabla_x \mathbf{o})^2 + (\nabla_y \mathbf{o})^2}/\delta} \right) + \nabla_y^T \left(\frac{\nabla_y \mathbf{o}}{1 + \sqrt{(\nabla_x \mathbf{o})^2 + (\nabla_y \mathbf{o})^2}/\delta} \right) \right], \quad \text{(A5)}$$

where the fraction line denotes pointwise division, $(\nabla_x^T \mathbf{o})(l,m) = o(l,m) - o(l+1,m)$, and $(\nabla_y^T \mathbf{o})(l,m) = o(l,m) - o(l,m+1)$.

3. Gradients of the Slope Data and of the Phase Prior Terms

The gradient of criterion J_s is straightforward to compute, because this criterion is quadratic. Furthermore, in practice the wave-front noise covariance matrix is taken as $C_{n'} = \sigma_{n'}^2 I$ (where I denotes the identity matrix) so that $J_s = (1/2\sigma_{n'}^2)\|\mathbf{s}_t - \mathcal{D}\boldsymbol{\phi}_t\|^2$, and its gradient reads as

$$\frac{\mathrm{d}J_s}{\mathrm{d}\boldsymbol{\phi}_t} = \frac{1}{\sigma_{n'}^2} \mathcal{D}^T(\mathcal{D}\boldsymbol{\phi}_t - \mathbf{s}_t). \quad \text{(A6)}$$

Similarly, the gradient of $J_\phi = \frac{1}{2}\boldsymbol{\phi}_t^T C_\phi^{-1} \boldsymbol{\phi}_t$ reads as

$$\frac{\mathrm{d}J_\phi}{\mathrm{d}\boldsymbol{\phi}_t} = C_\phi^{-1} \boldsymbol{\phi}_t. \quad \text{(A7)}$$

ACKNOWLEDGMENTS

This work was supported by contracts from SPOTI, Ministère de la Défense, France. The authors thank Thierry Fusco, Gérard Rousset, Jérôme Idier, and Amandine Blanc for fruitful discussions. Many thanks also to Ruy Deron, Joseph Montri, Francis Mendez, Jean Lefèvre, Didier Rabaud, and Christophe Coudrain, who made the experimental setup happen.

Corresponding author Laurent Mugnier can be reached by e-mail at Laurent.Mugnier@onera.fr.

REFERENCES

1. J.-C. Fontanella, "Analyse de surface d'onde, déconvolution et optique active," J. Mod. Opt. **16**, 257–268 (1985).
2. J. Primot, G. Rousset, and J.-C. Fontanella, "Image deconvolution from wavefront sensing: atmospheric turbulence simulation cell results," in *Very Large Telescopes and Their Instrumentation*, Vol. II, M.-H. Ulrich, ed., ESO Conf. and Workshop Proc. No. 30 (European Southern Observatory, Garching, Germany, 1988), pp. 683–692.
3. J. Primot, G. Rousset, and J.-C. Fontanella, "Deconvolution from wavefront sensing: a new technique for compensating turbulence-degraded images," J. Opt. Soc. Am. A **7**, 1598–1608 (1990).
4. J. D. Gonglewski, D. G. Voelz, J. S. Fender, D. C. Dayton, B. K. Spielbusch, and R. E. Pierson, "First astronomical application of postdetection turbulence compensation: images of α Aurigae, ν Ursae Majoris, and α Geminorum using self-referenced speckle holography," Appl. Opt. **29**, 4527–4529 (1990).
5. T. Marais, V. Michau, G. Fertin, J. Primot, and J. C. Fontanella, "Deconvolution from wavefront sensing on a 4 m telescope," in *High-Resolution Imaging by Interferometry II,* J. M. Beckers and F. Merkle, eds., ESO Conf. and Workshop Proc. No. 39 (European Southern Observatory, Garching, Germany, 1992), pp. 589–597.
6. F. Roddier, "Passive versus active methods in optical interferometry," in *High-Resolution Imaging by Interferometry Part II*, F. Merkle, ed., ESO Conf. and Workshop Proc. No. 29 (European Southern Observatory, Garching, Germany, 1988), pp. 565–574.
7. M. C. Roggemann, C. A. Hyde, and B. M. Welsh, "Fourier phase spectrum estimation using deconvolution from wavefront sensing and bispectrum reconstruction," in *Adaptive Optics,* Vol. 12 of 1996 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1996), pp. 133–135.
8. D. Dayton, J. Gonglewski, and S. Rogers, "Experimental measurements of estimator bias and the signal-to-noise ratio for deconvolution from wave-front sensing," Appl. Opt. **36**, 3895–3902 (1997).
9. M. C. Roggemann and B. M. Welsh, "Signal to noise ratio for astronomical imaging by deconvolution from wave-front sensing," Appl. Opt. **33**, 5400–5414 (1994).
10. M. C. Roggemann, B. M. Welsh, and J. Devey, "Biased estimators and object-spectrum estimation in the method of deconvolution from wave-front sensing," Appl. Opt. **33**, 5754–5763 (1994).
11. J.-M. Conan, V. Michau, and G. Rousset, "Signal-to-noise ratio and bias of various deconvolution from wavefront sensing estimators," in *Image Propagation through the Atmosphere*, J. C. Dainty and L. R. Bissonnette, eds., Proc. SPIE **2828**, 332–339 (1996).
12. T. J. Schulz, "Estimation-theoretic approach to the deconvolution of atmospherically degraded images with wavefront sensor measurements," in *Digital Image Recovery and Synthesis II*, P. S. Idell, ed., Proc. SPIE **2029**, 311–320 (1993).
13. L. M. Mugnier, J.-M. Conan, V. Michau, and G. Rousset, "Imagerie travers la turbulence par déconvolution myope multi-trame," in *Seizième Colloque sur le Traitement du Signal et des Images,* J.-M. Chassery and C. Jutten, eds. (Gretsi, Grenoble, France, 1997), pp. 567–570.
14. D. C. Dayton, S. C. Sandven, and J. D. Gonglewski, "Expectation maximization approach to deconvolution from wavefront sensing," in *Image Reconstruction and Restoration II*, T. J. Schulz, ed., Proc. SPIE **3170**, 16–24 (1997).
15. L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem, "Regularized multiframe myopic deconvolution from wavefront sensing," in *Propagation through the Atmosphere III*, M. C. Roggemann and L. R. Bissonnette, eds., Proc. SPIE **3763**, 134–144 (1999).
16. T. Fusco, J.-M. Conan, V. Michau, L. M. Mugnier, and G. Rousset, "Phase estimation for large field of view: application to multiconjugate adaptive optics," in *Propagation-*

through the Atmosphere III, M. C. Roggemann and L. R. Bissonnette, eds., Proc. SPIE **3763**, 125–133 (1999).
17. T. Fusco, J.-M. Conan, V. Michau, L. Mugnier, and G. Rousset, "Efficient phase estimation for large-field-of-view adaptive optics," Opt. Lett. **24**, 1472–1474 (1999).
18. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
19. C. M. Titterington, "General structure of regularization procedures in image reconstruction," Astron. Astrophys. **144**, 381–387 (1985).
20. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024–2036 (1989).
21. E. P. Wallner, "Optimal wave-front correction using slope measurements," J. Opt. Soc. Am. **73**, 1771–1776 (1983).
22. B. M. Welsh and R. N. VonNiederhausern, "Performance analysis of the self-referenced speckle-holography image-reconstruction technique," Appl. Opt. **32**, 5071–5078 (1993).
23. H. L. Van Trees, *Detection, Estimation, and Modulation Theory* (Wiley, New York, 1968).
24. P. A. Bakut, V. E. Kirakosyants, V. A. Loginov, C. J. Solomon, and J. C. Dainty, "Optimal wavefront reconstruction form a Shack–Hartmann sensor by use of a Bayesian algorithm," Opt. Commun. **109**, 10–15 (1994).
25. S. D. Ford, B. M. Welsh, and M. C. Roggemann, "Constrained least-squares estimation in deconvolution from wave-front sensing," Opt. Commun. **151**, 93–100 (1998).
26. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images by use of object and point-spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
27. L. P. Yaroslavsky and H. J. Caulfield, "Deconvolution of multiple images of the same object," Appl. Opt. **33**, 2157–2162 (1994).
28. P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM Algorithm," IEEE Trans. Med. Imaging **9**, 84–93 (1990).
29. C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," IEEE Trans. Image Process. **2**, 296–310 (1993).
30. W. J. Rey, *Introduction to Robust and Quasi-Robust Statistical Methods* (Springer-Verlag, Berlin, 1983).
31. S. Brette and J. Idier, "Optimized single site update algorithms for image deblurring," in *Proceedings of the International Conference on Image Processing* (IEEE Computer Society Press, Los Alamitos, Calif., 1996), pp. 65–68.
32. J.-M. Conan, T. Fusco, L. M. Mugnier, E. Kersalé, and V. Michau, "Deconvolution of adaptive optics images with imprecise knowledge of the point spread function: results on astronomical objects," in *Astronomy with Adaptive Optics: Present Results and Future Programs,* D. Bonaccini, ed., ESO Conf. and Workshop Proc. No. 56 (European Southern Observatory, Garching, Germany, 1999), pp. 121–132.
33. T. Fusco, J.-P. Véran, J.-M. Conan, and L. Mugnier, "Myopic deconvolution method for adaptive optics images of stellar fields," Astron. Astrophys., Suppl. Ser. **134**, 1–10 (1999).
34. T. J. Schulz, "Multiframe blind deconvolution of astronomical images," J. Opt. Soc. Am. A **10**, 1064–1073 (1993).
35. E. Thiébaut and J.-M. Conan, "Strict *a priori* constraints for maximum-likelihood blind deconvolution," J. Opt. Soc. Am. A **12**, 485–492 (1995).
36. Groupe Problèmes Inverses, "GPAV: une grande oeuvre collective," Internal Report, Laboratoire des Signaux et Systèmes (Centre National de la Recherche Scientifique/Supélec/Université Paris-Sud, Paris, 1997).
37. D. P. Bertsekas, *Nonlinear Programming* (Athena Scientific, Belmont, Mass., 1995).
38. D. G. Luenberger, *Introduction to Linear and Nonlinear Programming* (Addison-Wesley, Reading, Mass., 1973).
39. Y.-L. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," IEEE Trans. Image Processing **5**, 416–428 (1996).
40. N. Roddier, "Atmospheric wavefront simulation using Zernike polynomials," Opt. Eng. **29**, 1174–1180 (1990).
41. D. W. Tyler and C. L. Matson, "Speckle imaging detector optimization and comparison," Opt. Eng. **32**, 864–869 (1993).
42. E. Thiébaut, "Speckle imaging with the bispectrum and without reference star," in *International Astronomical Union Symposium on Very High Angular Resolution Imaging,* R. J. G. Robertson and W. J. Tango, eds. (Kluwer Academic, Dordrecht, The Netherlands, 1994), Vol. 158, p. 209.
43. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072–1085 (1992).

Annexe M

Article Mugnier et coll., 1996 [A41]

Mugnier *et al.* Vol. 13, No. 12/December 1996/J. Opt. Soc. Am. A 2367

Aperture configuration optimality criterion for phased arrays of optical telescopes

Laurent M. Mugnier, Gérard Rousset, and Frédéric Cassaing

Office National d'Etudes et de Recherches Aérospatiales, Division Imagerie Optique à Haute Résolution, BP 72, F-92322 Châtillon cedex, France

Received January 16, 1996; accepted May 14, 1996; revised manuscript received July 10, 1996

We address the optimization of the relative arrangement (aperture configuration) of a phased array of optical telescopes, coherently combined to form images of extended objects in a common focal plane. A novel optimality criterion, which is directly linked to the restoration error of the original object from the recorded image, is derived. This criterion is then refined into a second criterion to accommodate the possible knowledge of the noise spectrum. The optimal configuration is a function of the maximum spatial frequency of interest (or desired resolution) and takes into account the diameters of the elementary telescopes. Simulations illustrate the usefulness of this criterion for designing a synthetic-aperture optical instrument with three, four, and five telescopes. © 1996 Optical Society of America.

Key words: synthetic aperture, phased arrays, aperture configuration, interferometry, optical imaging, image restoration.

1. INTRODUCTION

The relative arrangement of the elementary telescopes (the so-called aperture configuration, or pupil configuration) is a key aspect of the design of a synthetic-aperture instrument. There is an abundant literature on this subject in radio astronomy (see, in particular, the pioneering work of Moffet[1] and of Golay[2] and the papers by Cornwell[3] and by Lannes *et al.*[4]). More recently, many papers have discussed this subject with respect to optical instruments.[5–15]

The currently operating synthetic-aperture optics (SAO) instruments are two-aperture interferometers, which provide only visibility measurements[16]—although new instruments are under development for imaging purposes[17]—so that optimization of the aperture configuration is a relatively new topic in optics. Papers dealing with the aperture configuration optimization of a SAO instrument often use various criteria based on the shape of the point-spread function (PSF), such as the full width at half-maximum, the encircled energy, and the sidelobe level.[5–7,10,12] In these papers the best PSF is implicitly taken as that of the full-aperture telescope. Nevertheless, it has already been pointed out that the choice of an optimal aperture configuration should be based on Fourier domain considerations.[7]

In contrast, radio astronomers, because their data consist of sparse frequency plane samples of the object spectrum, have considered Fourier domain aperture optimization and have developed a number of data processing algorithms to obtain an estimate of the object. Since even very simple digital processing of the data (i.e., of the recorded image) can yield a better object estimate than the raw image itself, we believe that such data processing (i.e., an image restoration) should be done for an imaging SAO instrument. This image restoration can even be regarded as part of the observation system, the first part being the instrument itself. In the following, we assume that such processing is performed.

Some papers dealing with the aperture configuration optimization of an SAO system do take a quality criterion based on the uniform filling of the spatial-frequency plane[8,13] (the so-called $u-v$ plane) or on the maximization of the contiguous central core diameter of the optical transfer function,[14] (OTF) rather than on the shape of the PSF, but this uniformity is not very precisely defined. Also, the frequency coverage given by the elementary telescopes—which can be an advantage of optical wavelengths over radio wavelengths—is rarely[9,15] taken into account.

The importance of a compact configuration (i.e., one with no zeros in the spatial-frequency coverage) for imaging an extended object such as the Sun has already been stressed.[15] Indeed, when the object's support lies within the field of view, a constraint support can be used in the object estimation to recover frequencies that have not been recorded,[4] and, for a given desired resolution, the smaller the support, the more effective the support constraint.[18] For such objects, one can consider diluted configurations, still taking advantage of the frequency coverage of elementary telescopes. But this is not the case when the object (e.g., the Earth viewed from a satellite) extends over the whole field of view. A compact configuration is therefore a necessary condition for the imaging of extended objects without ambiguity, but this condition is not sufficient to determine the aperture configuration uniquely.

The purpose of this paper is to derive a criterion for aperture configuration optimization in the case of an instrument that images extended objects. The problem is to design the aperture array under external constraints such as the desired resolution (i.e., the maximum spatial frequency of interest), the total collecting surface (i.e., the signal-to-noise ratio requirement), and the system com-

0740-3232/96/1202367-08$10.00

2368 J. Opt. Soc. Am. A/Vol. 13, No. 12/December 1996 Mugnier *et al.*

plexity (e.g., the number of elementary telescopes or the total size of the array).

In Section 2 a criterion is derived that minimizes the restoration error, i.e., the difference between the original object and the one estimated from the recorded image. This criterion, first presented in Ref. 19, defines rigorously what kind of frequency-plane uniformity is desirable to obtain an optimal configuration, and it explicitly takes into account the diameters of the elementary telescopes. In Section 3 this criterion is refined to accommodate the possible knowledge of the noise statistics. Then, in Section 4, computer- simulation results obtained with the defined criteria are presented.

2. APERTURE CONFIGURATION OPTIMALITY CRITERION

We consider a synthetic-aperture optical instrument that records images, that is, an instrument equivalent to a single telescope. This is in particular achieved with a phased array of elementary telescopes recombined homothetically[20] to form an image in a common focal plane. The recording process is modeled as

$$i = Ho + n, \tag{1}$$

where i is the recorded image, o is the original object, n is an additive noise, and H is the imaging operator in a Hilbert space $\mathbb{H}$ (e.g., the set of square integrable functions of two variables). The field-dependent aberrations are neglected in the following, so that the system is linear and shift invariant, and H is consequently a convolution operator of kernel h (the instrument's PSF):

$$i = h \star o + n. \tag{2}$$

If we let G be the restoration operator ($G = H^{-1}$, if H^{-1} exists, being the inverse filter), the estimated object reads as

$$o_e = Gi = GHo + Gn. \tag{3}$$

We define the restoration error by

$$\epsilon = \|o_e - o\| = \|(GH - I)o + Gn\|, \tag{4}$$

where I is the identity operator and $\|\cdot\|$ is the norm induced by the scalar product in $\mathbb{H}$.

We base our aperture configuration optimization on the minimization of the restoration error ϵ. Indeed, this error assesses the capability of the instrument (plus the restoration operator) to recover the object properly. We begin by deriving a bound on this error that is directly related to H and to G. With use of the triangular inequality,

$$\epsilon \leqslant \epsilon_o + \epsilon_n, \tag{5}$$

where

$$\epsilon_o = \|(GH - I)o\| \quad \text{and} \quad \epsilon_n = \|Gn\|. \tag{6}$$

The error term ϵ_o is a systematic type of error, which depends on the object o. It equals zero in particular if H is invertible and $G = H^{-1}$ and also if o belongs to the null space of $GH - I$ (that is, as we can see from the following, if GH is the identity up to the last frequencies of o). Nevertheless, in general, i.e., for an object o of infinite spectrum, H and G cannot be chosen so as to cancel ϵ_o. For a well-chosen G, ϵ_o is essentially due to the frequencies of the object above the cutoff of H; that is, ϵ_o is essentially determined by the choice of the instrument's resolution. In this paper we shall assume that the resolution (or, equivalently, the maximum frequency of interest) is already chosen by considerations regarding the types of object to be observed, and we shall optimize the configuration by minimizing the other term, ϵ_n, of the error.

This choice of resolution, which amounts to the choice of the frequency coverage of G, is similar to but different from the choice of the compromise between fidelity to the data (G close to H^{-1} and consequently ϵ_o small) and fidelity to the *a priori* information (smoothness of the solution, i.e., ϵ_n small), which is classical in ill-posed inverse problems.[21] Indeed, one should keep in mind that the present aim is not to best recover an object observed with a given instrument (which would involve the minimization of ϵ), but to *design* the instrument for a given resolution, so that our goal will be to minimize ϵ_n, the noise amplification that will occur during the restoration process, rather than ϵ. The relative noise amplification is defined by

$$\epsilon_n' = \frac{\|Gn\|}{\|o\|}. \tag{7}$$

Using the inequality $\|Ax\| \leqslant \|A\| \cdot \|x\|$, valid for any operator A and any vector x by definition of the norm of an operator, we see that

$$\epsilon_n \leqslant \|G\|\|n\| \tag{8}$$

$$\epsilon_n' \leqslant \|G\|\|H\| \frac{\|n\|}{\|Ho\|}. \tag{9}$$

It should be noted that the factor $\|n\|/\|Ho\|$ in Eq. (9) is the inverse of a signal-to-noise ratio (SNR) for the recorded image, since the L_2 norm is the square root of the integral of the spectral density of the signal. Likewise, ϵ_n' is the inverse of an SNR for the restored object. Thus the factor

$$c = \|G\|\|H\| \tag{10}$$

in relation (9) is a parameter that characterizes the degradation of the SNR (i.e., the noise amplification) during the imaging (H) plus restoration (G) process. It is the so-called condition number of numerical analysis when $G = H^{-1}$. It is this parameter c that will be used as a quality criterion for aperture configurations. Let us see now how to express c as a function of the OTF of the system.

First, the norm of an operator H is related to the eigenvalues of H^*H, where H^* is the adjoint of H. For a wide class of PSF's h (e.g., if h is square integrable), H is compact.[21] So H^*H is compact self-adjoint and, according to the Hilbert–Schmidt theorem (see, e.g., Ref. 22), has an eigenvalue decomposition. Additionally, H^*H is positive and $\|H^*H\| = \Lambda_s$, where Λ_s is the least upper bound (or supremum, which in fact is a maximum) of the eigenvalues of H^*H. Moreover, $\|H\| = \sqrt{\|H^*H\|}$ (see, e.g., Ref. 23) so that

$$\|H\| = \sqrt{\Lambda_s}. \tag{11}$$

If H is invertible then so is H^*H, and $\|(H^*H)^{-1}\| = {\Lambda_i}^{-1}$ where Λ_i is the greatest lower bound (or infimum) of the eigenvalues of H^*H [which is always 0 if the range of H is of infinite dimension, so that $(H^*H)^{-1}$ is unbounded]. Using $(H^{-1})^* = (H^*)^{-1}$, we can readily show that

$$\|H^{-1}\| = (\sqrt{\Lambda_i})^{-1}. \tag{12}$$

Second, the eigenvalues of H can be related to the OTF of the system. Indeed, in the discrete case the operator H is a matrix; since H is assumed, in this paper, to be a convolution operator, the matrix H has a block Toeplitz structure, which can be approximated by a block circulant matrix.[21,24] Within this approximation, which corresponds to periodizing the PSF h, H is diagonalized in the basis of the discrete Fourier exponentials $\exp[-2i\pi/N(m\mu + n\nu)]$, $0 \leqslant m,\ n \leqslant N-1$, and its eigenvalues are equal to the discrete Fourier transform values of the sampled PSF,[25] i.e., to the numerical OTF denoted hereafter by $\tilde{h}$.

Third, since the eigenvectors of H (the discrete Fourier exponentials) form an orthonormal basis, H^*H is diagonalized in the same basis as H, and its eigenvalues are the square moduli of those of H. In other words, the singular-value decomposition of H is in fact an eigenvalue decomposition, which in turn is a Fourier decomposition.

From this and Eq. (11), it is readily seen that

$$\|H\| = \max|\tilde{h}|, \quad \text{and (if } H^{-1} \text{ exists)}$$
$$\|H^{-1}\| = (\min|\tilde{h}|)^{-1}, \tag{13}$$

where $|\tilde{h}|$ is the modulation transfer function (MTF). Moreover, since the maximum value of the MTF is by convention normalized to unity (which corresponds to keeping the collecting surface constant), we obtain $\|H\| = 1$ so that the noise amplification ϵ_n and the relative noise amplification $\epsilon_n{}'$ are proportional to $c = \|G\|\|H\| = \|G\|$.

If H is invertible and if we take $G = H^{-1}$, we get

$$c = 1/\min|\tilde{h}|. \tag{14}$$

One should note that $\mathbb{H}$, the domain of definition of H, is the set of considered objects, so that the invertibility of H means that $|\tilde{h}|$ does not drop to zero on the frequency support of the considered objects. For objects with greater frequency support, H will not be invertible, and one should limit the resolution of the estimated object o_e (that is, the frequency support of the restoration filter G) to some maximum frequency $\omega_{\max}$ given *a priori* by the user (typically, the inverse of the desired resolution).

Given a maximum frequency of interest $\omega_{\max}$, we shall take for G the simple following linear filter (this filter, in operator theory terms, is the truncated singular value decomposition method[21,26]):

$$\tilde{g}(\boldsymbol{\omega}) = 1/\tilde{h}(\boldsymbol{\omega}) \quad \text{for} \quad \boldsymbol{\omega} \in D = \{\boldsymbol{\omega}: |\boldsymbol{\omega}| \leqslant \omega_{\max}\}$$
$$= 0 \quad \text{otherwise}, \tag{15}$$

where $D = \{\boldsymbol{\omega}: |\boldsymbol{\omega}| \leqslant \omega_{\max}\}$ is the frequency domain of interest, here a disk of radius $\omega_{\max}$ centered at the origin. Thus parameter c is

$$c = \|G\| \times 1 = 1/\min_{\boldsymbol{\omega} \in D}|\tilde{h}(\boldsymbol{\omega})|. \tag{16}$$

Relations (8), (9), and (16) can be interpreted as follows: The noise amplification during the restoration process is (at most) proportional to c, which is the inverse of the minimum value of the MTF in the frequency domain of interest. In particular, the relative noise amplification (inverse of the SNR of the restored object) is bounded by the ratio of c over the SNR of the recorded image.

It should be noted that the actual restoration filter used when the instrument is operating will most likely be more sophisticated than this basic one, e.g., a Wiener filter. Nevertheless, it will (if it is linear) be a variation along the idea embodied by G, i.e., it will be an MTF equalizer.

The aperture configuration quality can be assessed by the value of c, and the optimization consists in finding a configuration that minimizes c, i.e., that maximizes the minimum value of the MTF over the frequency domain of interest. In this sense, the optimal configuration is the one that is the flattest, or that has the most uniform frequency coverage. Also, compact configurations arise naturally—in the present setting, where no support constraint is available—since they are the ones with finite c. And the "practical resolution limit" defined by Harvey and Rockwell[9] coincides with the maximum value of $\omega_{\max}$ for which c is finite.

It is important to note that if a support constraint is available, the relationship between the norm of an operator and the discrete Fourier transform of the corresponding kernel [as in Eq. (13)] is no longer valid. The eigenvalues of H^*H are typically "pushed upwards" by such a constraint and no longer linked to the MTF, on which zeros may then be tolerated (see Lannes[4,18] on this subject).

Finally, this optimality criterion can be refined to accommodate the possible knowledge of the statistics of the noise, as explained in Section 3.

3. REFINED CRITERION FOR KNOWN NOISE STATISTICS

If the second-order statistics of the (zero-mean) noise n are known, it is possible to derive a better estimation of the noise amplification ϵ_n than the bound given in Eq. (8). However, it should be noted that this estimate will be in expected value, whereas the bounds given in relations (8) and (9) hold for any outcome of the noise.

The restoration operator G is still assumed to be a linear filter, so that (similarly to the developments given for H) its singular values are in fact eigenvalues, which are in turn approximately equal to the discrete Fourier transform values of its sampled PSF. The singular-value decomposition of G is then a discrete Fourier decomposition, and the square of the noise amplification is given by

$${\epsilon_n}^2 = \sum_{\boldsymbol{\omega}} |\tilde{g}(\boldsymbol{\omega})|^2 |\tilde{n}(\boldsymbol{\omega})|^2, \tag{17}$$

where $\tilde{n}(\boldsymbol{\omega})$ is the Fourier transform of the noise n. Let ${\sigma_{\tilde{n}}}^2(\boldsymbol{\omega}) = E\{|\tilde{n}(\boldsymbol{\omega})|^2\}$ be the so-called average power or average intensity of $\tilde{n}(\boldsymbol{\omega})$ (Ref. 27, Sec. 9-1); taking the expected value of Eq. (17) yields

$$E(\epsilon_n{}^2) = \sum_{\boldsymbol{\omega}} |\tilde{g}(\boldsymbol{\omega})|^2 \sigma_{\tilde{n}}{}^2(\boldsymbol{\omega})$$

$$= \sum_{\boldsymbol{\omega}\in D} \frac{1}{|\tilde{h}(\boldsymbol{\omega})|^2} \sigma_{\tilde{n}}{}^2(\boldsymbol{\omega}), \tag{18}$$

where $\tilde{g}$ is the filter defined in Eq. (15). If the average power $\sigma_{\tilde{n}}{}^2(\boldsymbol{\omega})$ of the noise is known, minimizing this expression will yield an aperture configuration that is optimal in the sense that the variance of the noise amplification in the restored image will be minimal.

In particular, if n is white, then $\tilde{n}$ is stationary,[27] i.e., the average power $\sigma_{\tilde{n}}{}^2(\boldsymbol{\omega})$ is constant, so that

$$E(\epsilon_n{}^2) = \left(\sum_{\boldsymbol{\omega}\in D} \frac{1}{|\tilde{h}(\boldsymbol{\omega})|^2} \right) \times \sigma_{\tilde{n}}{}^2 \propto \left\langle \frac{1}{|\tilde{h}(\boldsymbol{\omega})|^2} \right\rangle_{\boldsymbol{\omega}\in D}, \tag{19}$$

where $\langle \cdot \rangle_{\boldsymbol{\omega}\in D}$ denotes the average on all frequencies of the support of $\tilde{g}$. The optimal configuration is then obtained by minimizing the following refined criterion c':

$$c' = \sqrt{\left\langle \frac{1}{|\tilde{h}(\boldsymbol{\omega})|^2} \right\rangle_{\boldsymbol{\omega}\in D}}. \tag{20}$$

It must be noted that if n is truly stationary, then $\tilde{n}$ is white,[27] which in particular means that $\sigma_{\tilde{n}}{}^2(\boldsymbol{\omega})$, which is equal to the autocorrelation of $\tilde{n}$ for a zero shift $E\{\tilde{n}(\boldsymbol{\omega})\tilde{n}^*(\boldsymbol{\omega} + \mathbf{0})\}$, is infinite. This mathematical difficulty and the link between the average power $\sigma_{\tilde{n}}{}^2(\boldsymbol{\omega})$ and the power spectrum of n are explained in Appendix A.

If the noise statistics are not known, then Eq. (18) can still be used to yield the following bound:

$$E(\epsilon_n{}^2) \leqslant \frac{1}{\min_{\boldsymbol{\omega}\in D} |\tilde{h}(\boldsymbol{\omega})|^2} \times \sum_{\boldsymbol{\omega}\in D} \sigma_{\tilde{n}}{}^2(\boldsymbol{\omega}), \tag{21}$$

and the optimization reduces to the minimization of the previously derived criterion c [see Eq. (16)].

The configurations obtained with this refined criterion c' are typically slightly more compact (in the usual sense of the word) than those obtained with c, as shown in the simulations of Section 4. This can be understood intuitively as follows: On the one hand, the value of c for the c-optimal configuration is, in practice, determined by the value of the MTF at the highest frequency of interest (since the MTF typically goes down when the spatial frequency increases). On the other hand, criterion c' will "tolerate" (i.e., yield a configuration with) smaller values of the MTF at the highest frequencies (which means closer telescopes), because all frequencies of interest contribute to the value of c'.

The interpretation of the values of c and c' of a given configuration is the following: c characterizes the amplification of the noise at the spatial frequency inside the domain of interest that is the most attenuated by the optical system, whatever the noise statistics may actually be, whereas c' characterizes the average amplification of the noise at all spatial frequencies within the domain of interest for a white noise.

4. SIMULATIONS

The above criteria for aperture configuration optimality were implemented and tested for a given maximum frequency of interest for three, four, and five telescopes that are assumed to be identical and whose diameter is determined by the total collecting surface, which is kept constant in all simulations. In order to express the diameter of the telescopes in the same unit as the maximum frequency, the latter is best expressed as an equivalent length, namely, the wavelength divided by the angular resolution. This length is called the maximum-frequency equivalent diameter. With this convention, the cutoff frequency D/λ of a monolithic telescope of diameter D (of, say, 40 arbitrary units) would be taken as 40. Thus, diameters and frequencies will be expressed in the same arbitrary units—pixels—in the following.

In the presented simulations, the telescopes are constrained to lie on a circle, whose radius is allowed to vary. This is especially justified in order to simplify the design of an optical space instrument[8,11,13] and also limits the search of the algorithm (except for three telescopes, since three nonaligned points are always on a circle).

Since the elementary telescopes are assumed to be identical, the global OTF of the array, OTF, can be computed as the sum of replicas of the elementary telescope OTF, OTF_e, placed at the correlation peaks of the array:

$$OTF = OTF_e \star RTF, \tag{22}$$

where RTF is the radio transfer function, a set of delta functions placed at the correlation peaks of the array, the one at the origin being of height 1, and the others of height given by the redundancy (e.g., $1/N_t$ for a nonredundant array, where N_t is the number of telescopes). This computation is more accurate than performing a numerical correlation of the aperture, especially for frequencies close to the cutoff. The elementary telescope OTF is computed by the analytical formula given in Appendix B to take into account the central obscuration.

The algorithm used is essentially an exhaustive search of the possible positions of the telescopes. The search is pruned by keeping the first of the N_t telescopes fixed and by limiting the angle between the first two telescopes to $2\pi/N_t$; indeed, if we consider a configuration in which this angle is larger than $2\pi/N_t$, there will exist another pair of telescopes separated by an angle smaller than this

Table 1. Optimal Configurations with Criterion c[a]

Number of Telescopes	c	c'	Support Diameter (pixels)	Angular Position (deg)
3	26.1	8.9	72	−180, −60, 60
4	13.8	7.7	72	−180, −120, −32, 92
5	13.6	7.0	74	−180, −108, −37, 35, 108

[a] The second column (c) gives a bound for noise amplification at any frequency; the third (c') gives the average noise amplification at all frequencies. The fourth and fifth columns give the diameter of the circle supporting the telescopes and their angular positions, respectively. The maximum frequency of interest is 70 pixels; the collecting surface is kept constant and corresponds to a telescope diameter of 40 pixels for three telescopes; all telescopes have a 33% central obscuration.

Table 2. Optimal Configurations with Criterion c'[a]

Number of telescopes	c	c'	Support Diameter (pixels)	Angular Position (deg)
3	31.1	7.5	62	−180, −59, 59
4	21.1	7.1	64	−180, −92, −26, 64
5	14.7	6.7	66	−180, −108, −35, 35, 108

[a] The second column (c) gives a bound for noise amplification at any frequency; the third (c') gives the average noise amplification at all frequencies. The fourth and fifth columns give the diameter of the circle supporting the telescopes and their angular positions, respectively. The maximum frequency of interest is 70 pixels; the collecting surface is kept constant and corresponds to a telescope diameter of 40 pixels for three telescopes; all telescopes have a 33% central obscuration.

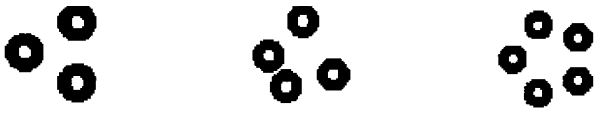


Fig. 1. Optimal configurations with criterion c for three, four, and five telescopes (see Table 1).

Fig. 2. Optimal configurations with criterion c' for three, four, and five telescopes (see Table 2).

value, so that the configuration will be equivalent, apart from a rotation, to one in which the first two telescopes are less than $2\pi/N_t$ away.

In all our simulations the scale was set by taking 20 pixels as the radius of the telescopes of the three-telescope array. The radii of the telescopes of the other arrays are determined by keeping the collecting surface constant. We considered a typical value of 0.33 for the central obscuration and a frequency of interest of 70. The angular increment is 2 deg, and the radial increment is at most 2 pixels. The optimal configurations for three, four, and five telescopes are shown in Tables 1 (criterion c) and 2 (criterion c'), and are depicted in Figs. 1 and 2, respectively.

For this value of $\omega_{\max}$ and criterion c (Table 1), the five-telescope optimal configuration is slightly better than the four-telescope configuration, which in turn is far better than the three-telescope configuration.

For criterion c' also (Table 2), the five-telescope configuration is better than the four-telescope one, which in turn is better than the three-telescope one, but the differences in the values of c' are much smaller, owing to the average in the computation of c'.

For three telescopes the optimum configuration is an equilateral triangle, but for four telescopes it is not a square. For five telescopes the optimum configuration is a regular pentagon (within the precision of our simulations). In other words, the chosen criterion leads to configurations that appear to minimize naturally the redundancy of the array without imposing it *a priori* in the search. One should nevertheless bear in mind that for an optical array the elementary telescopes themselves introduce some redundancy, owing to their nonzero diameter.

In order to check whether the optimal configuration we obtained (for a given number of telescopes and a given criterion) was pathological, we compared it with the ten configurations that had the closest values of the criterion (e.g., from 7.06 for the optimum to 7.09 for the tenth-best four-telescope configuration according to criterion c'). We indeed noticed that all these configurations were very close to the one that ranked first, apart from a possible symmetry and/or rotation.

Even for the simple case of three telescopes, criterion c or c' can be useful in determining the optimal spacing between telescopes and the expected quality of the frequency coverage for a given resolution and telescope diameter. For three telescopes the optimum configuration is always an equilateral triangle. Their spacing increases with the desired resolution (or maximum fre-

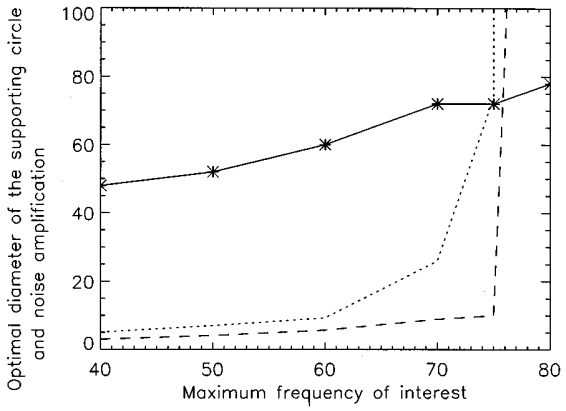


Fig. 3. Evolution of the optimal spacing of a three-telescope array with the maximum frequency (solid curve) and corresponding noise amplification parameters c (dotted curve) and c' (dashed curve). The optimization is done with criterion c, the telescope diameter is 40 pixels, and the frequency is expressed in pixels.

Table 3. Optimal Diameter of the Circle Supporting the Telescopes and Noise Amplification Parameters as a Function of the Maximum Frequency, for a Three-Telescope Array[a]

Maximum Frequency (pixels)	Noise Amplification Parameters		Optimal Diameter of the Support (pixels)
	c	c'	
40	5.1	3.0	48
50	7.0	4.1	52
60	9.3	5.7	60
70	26.0	8.9	72
75	74.0	10.0	72
80	4×10^5	4×10^2	78
90	∞	∞	NA

[a] Telecopes are of diameter 40. Optimization is done with criterion c. NA, not applicable.

quency), as shown in Fig. 3. As expected intuitively, since the telescopes are constrained to lie on a circle, the optimal diameter of the circle supporting the telescopes is never far from the diameter of a monolithic telescope having the maximum frequency of interest as its cutoff frequency. Also, not surprisingly, the quality of the frequency coverage degrades when the maximum frequency increases, and this quality can be quantified by parameters c and c' (see Fig. 3 and summary Table 3). Finally, if the maximum frequency of interest is too high (in practice above 80 in the simulations presented here), no compact configuration can reach it, so that c and c' become infinite.

5. CONCLUSION

A criterion has been derived to find an optimal aperture configuration for a synthetic-aperture optical instrument that provides images of extended objects. This criterion is based on the minimization of the restoration error: the difference between the original object and the one that will be estimated from the recorded image. It explicitly takes into account the resolution to be achieved and the diameter of the elementary telescopes. This criterion was shown to be equal to the inverse of the minimum value of the MTF in the frequency domain of interest. It gives a bound on the noise amplification in the imaging-plus-restoration process at all frequencies of interest, for any outcome of the noise. Furthermore, it has been refined to take into account the possible knowledge of the noise second-order statistics. For a white noise the refined criterion is the average of the inverse of the squared MTF in the frequency domain of interest. It is proportional to the expected value of the noise amplification.

The results of the optimization of arrays of three, four, and five telescopes, under the constraints of a given collecting surface and resolution, can be summarized as follows: First, the three- and five- telescope optimal configurations are regular polygons, whereas the four-telescope one is not. We interpret this as being because our criterion naturally minimizes the redundancy of the array without imposing it *a priori*. Second, the optimal diameter of the circle supporting the telescopes was found always to be close to that of a monolithic telescope having the maximum frequency of interest as its cutoff.

Future work should assume an optimal linear (i.e., Wiener) filter instead of the truncated inverse filter used here, even if the truncated inverse filter is close to the Wiener filter in the case of a high SNR. Additionally, in our approach, the SNR is controlled through the choice of the collecting surface. An extension of this work could then include the incorporation of constraints, such as positivity and/or support (when the object's support lies within the field of view), to allow zeros in the transfer function while still taking into account the frequency coverage of the elementary telescopes.

APPENDIX A: FOURIER TRANSFORM OF A STATIONARY NOISE

In this appendix we examine the relationship between the power spectrum of the noise n and the average power of its Fourier transform $\tilde{n}$. See Papoulis (Ref. 27, Sec. 11-3) for an overview of the second-order properties of the Fourier transform of random processes.

Let n be a noise (a zero-mean stochastic process, of the one-dimensional variable t for the sake of clarity). The Fourier transform $\tilde{n}(\omega)$ of $n(t)$ is also a zero-mean stochastic process of the variable ω. The autocorrelation of $\tilde{n}(\omega)$ is $E\{\tilde{n}(\omega_1)\tilde{n}^*(\omega_2)\}$, and, in particular, the average power or average intensity of $\tilde{n}(\omega)$ is by definition (Ref. 27, Sec. 9-1) $\sigma_{\tilde{n}}{}^2(\omega) = E\{|\tilde{n}(\omega)|^2\}$.

If $n(t)$ is white, then $\tilde{n}(\omega)$ is stationary, i.e., $\sigma_{\tilde{n}}{}^2(\omega)$ does not depend on ω. Conversely, if n is stationary, then $\tilde{n}$ is white; i.e., its autocorrelation is a Dirac function, and in particular the average power $\sigma_{\tilde{n}}{}^2$ is infinite. The origin of this infinite value is that a stationary noise is implicitly observed from $t = -\infty$ to $t = +\infty$.

Indeed, let n_T be the observed noise, equal to the true noise n (assumed to be stationary) in the observation time window $[-T/2, T/2]$ and zero outside this window. Then n_T is not stationary and its average power is finite, and we can express it as a function of the power spectrum of the noise n. Let $\Pi_T(t)$ be the characteristic function of the observation time window and $R(\tau)$ be the autocorrelation of n; then the autocorrelation of n_T and $R(\tau)$ are related through

$$E[n_T(t)n_T(t+\tau)] = \Pi_T(t)\Pi_T(t+\tau)R(\tau). \quad \text{(A1)}$$

If the support of R is much smaller than the observation time window, then

$$E[n_T(t)n_T(t+\tau)] \approx \Pi_T(t)R(\tau). \quad \text{(A2)}$$

The autocorrelation of the Fourier transform of the observed noise is then

$$\begin{aligned} E[\tilde{n}_T(\omega_1)\tilde{n}_T{}^*(\omega_2)] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} E\{n_T(t)n_T(t+\tau)\}\exp[-2i\pi(\omega_1-\omega_2)t]\exp[2i\pi\omega_2\tau]\mathrm{d}t\,\mathrm{d}\tau \\ &\approx \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Pi_T(t)\exp[-2i\pi(\omega_1-\omega_2)t]R(\tau)\exp(2i\pi\omega_2\tau)\mathrm{d}t\,\mathrm{d}\tau \\ &\approx T\ \mathrm{sinc}\{(\omega_1-\omega_2)T\}S(\omega_2), \end{aligned} \quad \text{(A3)}$$

where $S(\omega)$ is the power spectrum of the true noise n, i.e., the Fourier transform of $R(\tau)$. This expression, as expected, shows that the autocorrelation of $\tilde{n}_T$ tends towards a Dirac function when the observing time tends to infinity. It shows in particular that the average power of

$\tilde{n}_T$ is finite and directly proportional to the power spectrum of the noise (a result derived in a different manner in Ref. 27):

$$\sigma_{\tilde{n}_T}{}^2(\omega) = TS(\omega). \tag{A4}$$

Of course, if $n(t)$ is additionally white, then $\sigma_{\tilde{n}_T}{}^2(\omega)$ is constant.

APPENDIX B: OPTICAL TRANSFER FUNCTION WITH CENTRAL OBSCURATION

To the best of our knowledge, an analytic expression of the OTF of a telescope with central obscuration was first derived by Perrier.[28] Unfortunately, the expression appeared with what is most likely a typographic error (in the expression for function H_1). Thus we take this opportunity to derive an equivalent and somewhat simpler expression.

The basis of the computation is to derive the correlation $C_U(x)$ between two disks, one of diameter 1 and one of diameter U, where U is the (linear) central obscuration, defined as the ratio of the obscuration diameter to the telescope pupil diameter. Geometric considerations and elementary trigonometry yield the following result [apart from a $4/\pi$ normalization factor to ensure that $C_1(0) = 1$]:

$$C_U(x) = \begin{cases} U^2 & : x \leqslant \dfrac{1-U}{2} \\ 0 & : x \geqslant \dfrac{1+U}{2} \\ \dfrac{1}{\pi}\arccos\left(x + \dfrac{1-U^2}{4x}\right) & \\ \quad + \dfrac{U^2}{\pi}\arccos\left[\dfrac{1}{U}\left(x - \dfrac{1-U^2}{4x}\right)\right] & : \dfrac{1-U}{2} \leqslant x \leqslant \dfrac{1+U}{2} \\ \quad - \dfrac{2x}{\pi}\sqrt{1 - \left(x + \dfrac{1-U^2}{4x}\right)^2} & \end{cases} . \tag{B1}$$

In addition, the pupil function of a telescope with central obscuration can be written as $P = P_+ - P_-$, where P_+ is the pupil function of the telescope without central obscuration and P_- is the pupil function of the central obscuration (1 inside the obscuration, and 0 outside). The OTF of the telescope with central obscuration reads as

$$OTF = P \otimes P = P_+ \otimes P_+ + P_- \otimes P_- - 2P_+ \otimes P_-, \tag{B2}$$

where $\otimes$ denotes correlation. With proper normalization the OTF is then readily expressed as a function of C_U:

$$OTF(x) = \frac{1}{1-U^2}[C_1(x) + U^2 C_1(x/U) - 2C_U(x)], \tag{B3}$$

where $x = \lambda f/D$ is the reduced frequency.

ACKNOWLEDGMENTS

This research was supported by Direction des Recherches Études et Techniques. The authors thank their colleagues B. Sorrente, J.-M. Conan, and M. Séchaud for fruitful discussions.

REFERENCES

1. A. T. Moffet, "Minimum-redundancy linear arrays," IEEE Trans. Antennas Propag. **AP-16**, 172–175 (1968).
2. M. J. E. Golay, "Point arrays having compact, nonredundant autocorrelations," J. Opt. Soc. Am. **61**, 272–273 (1971).
3. T. J. Cornwell, "A novel principle for optimization of the instantaneous Fourier plane coverage of correlation arrays," IEEE Trans. Antennas Propag. **AP-36**, 1165–1167 (1988).
4. A. Lannes, É. Anterrieu, L. Koechlin, and G. Fitoussi, "On the concept of field-to-resolution ratio in aperture synthesis," in *Space Optics 1994: Earth Observation and Astronomy*, G. Cerutti-Maori and P. Roussel, eds., Proc. SPIE **2209**, 402–412 (1994).
5. R. V. Shack, J. D. Rancourt, and H. Morrow, "Effects of dilution on a six-element synthetic aperture," Appl. Opt. **10**, 257–259 (1971).
6. A. B. Meinel, M. P. Meinel, and N. J. Woolf, "Multiple aperture telescope diffraction images," in *Applied Optics and Optical Engineering*, R. R. Shannon and J. C. Wyant, eds. (Academic, New York, 1983), Vol. 9, Chap. 5, pp. 149–201.
7. J. E. Harvey, A. B. Wissinger, and A. N. Bunner, "A parametric study of various synthetic aperture telescope configurations for coherent imaging applications," in *Infrared, Adaptive, and Synthetic Aperture Optical Systems*, R. B. Johnson, W. L. Wolfe, and J. S. Fender, eds., Proc. SPIE **643**, 194–207 (1986).
8. A. Labeyrie, G. Lemaitre, and L. Kœchlin, "The optical very large array," in *Advanced Technology Optical Telescopes III*, L. D. Barr, ed., Proc. SPIE **628**, 323–332 (1986).
9. J. E. Harvey and R. A. Rockwell, "Performance characteristics of phased arrays and thinned aperture optical telescopes," Opt. Eng. (Bellingham) **27**, 762–768 (1988).
10. S. M. Watson and J. P. Mills, "Two-point resolution criterion for multiaperture optical telescopes," J. Opt. Soc. Am. A **5**, 893–903 (1988).
11. P. Y. Bely, "HARDI: a high angular resolution deployable interferometer for space," in *New Technologies for Astronomy*, J.-P. Swings, ed., Proc. SPIE **1130**, 92–100 (1989).
12. ESO/VLT Interferometry Panel, "The VLT Interferometer Implementation Plan," VLT Report 59b (European Southern Observatory, Garching, Germany, 1989).
13. M. Faucherre, F. Merkle, and F. Vakili, "Beam combination in aperture synthesis from space: field of view limitations and (u, v) plane coverage optimization," in *New Technolo-*

gies for Astronomy, J.-P. Swings, ed., Proc. SPIE **1130**, 138–145 (1989).
14. J. P. Fitch and T. W. Lawrence, "Placement of multiple apertures for imaging telescopes," in *Amplitude and Intensity Spatial Interferometry*, J. B. Breckinridge, ed., Proc. SPIE **1237**, 61–69 (1990).
15. L. Damé and T.-D. Guyenne, "Study of an optimized configuration for interferometric imaging of complex and extended solar structures," in *Targets for Space-Based Interferometry* (European Space Agency, Noordwijk, The Netherlands, 1992), Vol. SP-354, pp. 201–208.
16. J. M. Beckers and F. Merkle, eds., *High-Resolution Imaging by Interferometry II. Part II: Multiple Aperture Interferometry*, No. 39 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 1992).
17. J. B. Breckinridge, ed., *Amplitude and Intensity Spatial Interferometry II*, Proc. SPIE **2200** (1994).
18. A. Lannes, S. Roques, and M.-J. Casanove, "Stabilized reconstruction in image and signal processing; part I: partial deconvolution and spectral extrapolation with limited field," J. Mod. Opt. **34**, 161–226 (1987).
19. L. M. Mugnier and G. Rousset, "Pupil configuration optimality criterion in synthetic aperture optics," in *Spaceborne Interferometry II*, R. D. Reasenberg, ed., Proc. SPIE **2477**, 124–131 (1995).
20. W. A. Traub, "Combining beams from separated telescopes," Appl. Opt. **25**, 528–532 (1986).
21. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust. Speech Signal Process. **37**, 2024–2036 (1989).
22. V. Trénoguine, *Analyse Fonctionnelle* (Mir, Moscow, 1985).
23. L. Schwartz, *Analyse—Topologie Générale et Analyse Fonctionnelle*, No. 11 of Collection Enseignement des Sciences (Hermann, Paris, 1970).
24. A. K. Katsaggelos, "Iterative image restoration algorithms," Opt. Eng. (Bellingham) **28**, 735–748 (1989).
25. B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," IEEE Trans. Comp. **C-22**, 805–812 (1973).
26. C. K. Rushforth, "Signal restoration, functional analysis, and Fredholm integral equations of the first kind," in *Image Recovery: Theory and Application*, H. Stark, ed. (Academic, 1987), Chap. 1.
27. A. Papoulis, *Signal Analysis* (McGraw-Hill, New-York, 1977).
28. C. Perrier, "Amplitude estimation from speckle interferometry," in *Proceedings of the NATO Advanced Study Institute on Diffraction-Limited Imaging with Very Large Telescopes*, D. M. Allouin and J.-M. Mariotti, eds., Vol. 274 of Series C: Mathematical and Physical Sciences (Kluwer, Dordrecht, The Netherlands, 1989), pp. 99–111.

Problèmes inverses en Haute Résolution Angulaire

Les travaux exposés portent sur les techniques d'imagerie optique à haute résolution et plus particulièrement sur les méthodes, dites d'inversion, de traitement des données associées à ces techniques. Ils se situent donc à la croisée des chemins entre l'imagerie optique et le traitement du signal et des images.

Ces travaux sont appliqués à l'astronomie depuis le sol ou l'espace, l'observation de la Terre, et l'imagerie de la rétine.

Une partie introductive est dédiée au rappel de caractéristiques importantes de l'inversion de données et d'éléments essentiels sur la formation d'image (diffraction, turbulence, techniques d'imagerie) et sur la mesure des aberrations (analyse de front d'onde).

La première partie des travaux exposés porte sur l'étalonnage d'instrument, c'est-à-dire l'estimation d'aberrations instrumentales ou turbulentes. Ils concernent essentiellement la technique de diversité de phase : travaux méthodologiques, travaux algorithmiques, et extensions à l'imagerie à haute dynamique en vue de la détection et la caractérisation d'exoplanètes. Ces travaux comprennent également des développements qui n'utilisent qu'une seule image au voisinage du plan focal, dans des cas particuliers présentant un intérêt pratique avéré.

La seconde partie des travaux porte sur le développement de méthodes de traitement (recalage, restauration et reconstruction, détection) pour l'imagerie à haute résolution. Ces développements ont été menés pour des modalités d'imagerie très diverses : imagerie corrigée ou non par optique adaptative (OA), mono-télescope ou interférométrique, pour l'observation de l'espace ; imagerie coronographique d'exoplanètes par OA depuis le sol ou par interférométrie depuis l'espace ; et imagerie 2D ou 3D de la rétine humaine.

Enfin, une dernière partie présente des perspectives de recherches.

Mots-clés : PROBLÈME INVERSE ; DÉCONVOLUTION ; RESTAURATION IMAGE ; RECONSTRUCTION IMAGE ; DÉTECTION ; IMAGERIE HYPERSPECTRALE ; ANALYSE FRONT D'ONDE; COPHASAGE ; DIVERSITÉ PHASE ; OPTIQUE ADAPTATIVE ; OPTIQUE ACTIVE; INTERFÉROMÉTRIE ; SYNTHÈSE D'OUVERTURE ; ASTRONOMIE ; OPHTALMOLOGIE ; IMAGERIE RÉTINIENNE ; MICROSCOPIE.

Office National d'Études et de Recherches Aérospatiales
BP 72 - 29 avenue de la Division Leclerc
92322 CHATILLON CEDEX
Tél. : +33 1 46 73 40 40 - Fax : +33 1 46 73 41 41
http://www.onera.fr