

**UNIVERSITÉ PARIS XI
UFR SCIENTIFIQUE D'ORSAY**

THÈSE

Présentée

Pour obtenir

**Le GRADE de DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS XI ORSAY**

PAR

Amandine BLANC

**IDENTIFICATION DE RÉPONSE IMPULSIONNELLE
ET RESTAURATION D'IMAGES :
APPORTS DE LA DIVERSITÉ DE PHASE**

Soutenue le 5 juillet 2002 devant la commission d'examen

M. Guy DEMOMENT	Examineur
M. Jérôme IDIER	Directeur de thèse
M. Henri MAITRE	Président
M. Laurent MUGNIER	Co-directeur de thèse
Mme Sylvie ROQUES	Rapporteur
M. Éric THIÉBAUT	Rapporteur

Table des matières

1	Imagerie et Analyse de surface d'onde	11
1.1	La formation d'image	11
1.1.1	Réponse impulsionnelle d'un instrument	11
1.1.2	Relation entre objet et image	14
1.1.3	Les dégradations de l'image	15
1.2	La RI: origine des dégradations	16
1.2.1	Introduction	16
1.2.2	La turbulence atmosphérique	17
1.2.3	Les défauts propres de l'instrument	19
1.2.4	Base des Zernike et décomposition de la phase	19
1.3	L'analyse de surface d'onde	21
1.3.1	Shack-Hartmann	21
1.3.2	Courbure	22
1.3.3	La diversité de phase	23
2	Estimation conjointe des aberrations et de l'objet	29
2.1	Introduction	29
2.2	L'estimateur MAPJ	30
2.2.1	Construction du critère MAPJ	30
2.2.2	Écriture du critère dans le domaine de Fourier	32
2.2.3	Choix et réglage des hyperparamètres	33
2.2.4	Minimisation du critère	35
2.3	Propriétés asymptotiques	36
2.3.1	Définitions: biais, variance et écart quadratique moyen	36
2.3.2	Comportement asymptotique	37
2.4	Influence des hyperparamètres	44
2.4.1	Variation des hyperparamètres autour de leur valeur vraie	44
2.4.2	Comportement asymptotique à régularisation quasi-nulle	45
2.5	Conclusion	48
3	Estimation des aberrations de la caméra CONICA	53
3.1	Présentation de NAOS+CONICA	53
3.1.1	CONICA	53
3.1.2	NAOS	55
3.2	Mise en œuvre	56
3.3	Validations expérimentales	58
3.3.1	Influence de la quantité de défocalisation introduite	58
3.3.2	Forme et position de la pupille	59

3.3.3	Fond résiduel	60
3.3.4	Incertitude sur le facteur d'échantillonnage	60
3.3.5	Incertitude sur la distance de défocalisation	61
3.3.6	Influence de la translation dans le plan focal d'entrée	63
3.3.7	Forme des trous	63
3.3.8	Décalage latéral entre les deux trous	64
3.3.9	Synthèse de l'étude des différentes sources de perturbations	64
3.4	Premiers résultats et comparaisons	64
3.4.1	Comparaison des résultats obtenus à partir de différentes paires de trous	65
3.4.2	Présentation des résultats des différents filtres	65
3.4.3	Comparaison entre les RI estimées et mesurées	69
3.4.4	Comparaison avec des calculs de Strehl sur l'image plan focal	69
3.5	Conclusion	71
4	Estimation marginale des aberrations	73
4.1	Introduction	73
4.2	Construction du critère	74
4.2.1	Principe	74
4.2.2	Expression de R_I^{-1} et \mathbf{m}_I	74
4.2.3	Calcul de $(\mathbf{I} - \mathbf{m}_I)^T R_I^{-1} (\mathbf{I} - \mathbf{m}_I)$	75
4.2.4	Calcul du déterminant de R_I	76
4.3	Expression du critère marginal	77
4.4	Relation conjoint/marginal	77
4.5	Expression du gradient	77
4.6	Validation à hyperparamètres connus	78
4.7	Propriétés asymptotiques	79
4.8	Estimation non supervisée des hyperparamètres	81
4.8.1	Principe	81
4.8.2	Minimisation du critère	82
4.8.3	Résultats de l'estimation non supervisée	84
4.9	Validation sur données expérimentales	86
4.9.1	Objet étendu : Traitement des effets de bords	86
4.9.2	Montage expérimental	89
4.9.3	Résultats	90
4.10	Conclusion	91
5	Comparaison joint/marginal	93
5.1	Introduction	93
5.2	Sur données simulées	94
5.2.1	Conditions de simulation et d'estimation	94
5.2.2	Résultats de la comparaison	94
5.2.3	En résumé	96
5.2.4	Interprétations des résultats	99
5.3	Sur données expérimentales	102
5.3.1	Acquisition des données	102
5.3.2	Calibration des aberrations du banc à fort flux	102
5.3.3	Résultats : évolution en fonction du niveau du flux	103
5.4	Conclusion	103

6	Restauration de l'objet : méthodes possibles	107
6.1	Introduction	107
6.2	Restauration de l'objet avec l'estimateur joint	107
6.3	Restauration de l'objet avec l'estimateur marginal	108
6.3.1	Principe	108
6.3.2	Exemples de restaurations	109
6.3.3	Adéquation entre $\hat{\mathbf{o}}_{\text{optimal}}$ et $\hat{\mathbf{a}}_{\text{optimal}}$	109
6.4	Par estimation "mixte"	112
6.4.1	Principe	112
6.4.2	Les trois étapes	112
6.4.3	Résultats	113
6.5	Conclusion	113
7	Estimation des fortes aberrations	115
7.1	Introduction	115
7.2	Indétermination de 2π sur la phase	115
7.3	Méthodes d'estimation possibles	116
7.3.1	Estimation de la phase dépliée	116
7.3.2	Estimation de $\phi[2\pi]$ (puis déroulage)	118
7.3.3	Méthodes employées dans la littérature	119
7.4	Construction du critère MAP	120
7.4.1	Choix d'un terme de régularisation	120
7.4.2	Critère MAP	121
7.4.3	Réglage des hyperparamètres	122
7.4.4	Minimisation du critère	122
7.4.5	Choix d'un critère de qualité	123
7.5	Résultats de simulations	123
7.5.1	Exemple montrant l'intérêt de la reconstruction point-à-point	123
7.5.2	Phases turbulentes	125
7.5.3	En résumé	132
7.6	Conclusion	132
A	Domaine de validité de la RI d'un système imageur	137
A.1	Introduction	137
A.2	Domaine de validité de la RI d'un système imageur	137
A.2.1	Au foyer	137
A.2.2	Validité le long de l'axe optique	138
B	Indétermination sur les aberrations	141
B.1	Décomposition de la phase	141
B.2	Propriété de l'autocorrélation	141
B.3	Indétermination sur la phase	142
B.4	Indétermination sur la décomposition de la phase sur les Zernike	142
B.5	Utilisation d'une deuxième image défocalisée	142
C	Quelques filtres de CONICA	145

D Propriétés matricielles	147
D.1 Lemme d'inversion des matrices partitionnées	147
D.2 Le déterminant d'une matrice bloc	147
E Gradients du critère marginal	149
E.1 A partir du critère marginal J_{MAP}	149
E.1.1 Expression du gradient par rapport à μ	149
E.1.2 Expression du gradient par rapport à v_o	150
E.1.3 Expression du gradient par rapport à p	150
E.2 A partir du critère marginal alterné $J_{\text{MAP}}^{\text{alt}}$	150
E.2.1 Expression de \hat{k}	151
E.2.2 Expression du gradient par rapport à μ	151
E.2.3 Expression du gradient par rapport à v_o	151
E.2.4 Expression du gradient par rapport à p	151
F Extension du critère marginal à trois images	153
F.1 Le critère	153
F.2 Le gradient par rapport à ϕ	154
F.3 Les hyperparamètres	154
G Article A&A	157
H Article GRETSI	183
I Article SPIE	189

Remerciements

J'adresse mes remerciements à Henri Maître pour avoir accepté de présider mon jury de thèse. Je remercie vivement Guy Demoment qui m'a accueillie au L2S et qui a accepté d'être examinateur de ce travail. Je tiens à remercier chaleureusement Sylvie Roques et Eric Thiébaud pour avoir accepté le lourd travail de rapporteur et ce dans un temps très court. Merci pour leurs remarques constructives.

C'est bien évidemment Jérôme Idier et Laurent Mugnier que je souhaite remercier avant tout ici. C'est un privilège pour moi d'avoir pu travailler avec eux. Ils ont en commun de grandes compétences et une grande rigueur scientifique. Merci à Jérôme pour avoir toujours pris le temps de répondre longuement à mes questions et pour ses conseils avisés et merci à Laurent pour sa patience, ses encouragements et pour sa rigueur impitoyable (notamment lors de la rédaction de ce manuscrit).

Un grand merci à Vincent Michau pour s'être intéressé de près à mes travaux tout au long de ma thèse, pour n'avoir jamais hésité à me consacrer du temps et enfin pour m'avoir aidée à mettre un point final à ces travaux.

Cette thèse comporte une partie expérimentale dont le succès doit beaucoup à Marie-Thérèse Velluet. Ce fut un plaisir de travailler en sa compagnie. Elle m'a offert du temps, de précieux conseils et sa bonne humeur communicative. Je ne saurais oublier Francis Mendez qui a effectué les réglages optiques du banc d'optique active et qui a effectué un grand nombre d'acquisitions expérimentales pour moi. Qu'ils trouvent ici, tous les deux, l'expression de ma profonde gratitude.

Merci à Gérard Rousset de m'avoir proposé de travailler sur CONICA, ce qui m'a donné l'occasion de rencontrer Markus Hartung. Ce travail de "trois semaines" s'est transformé en collaboration de plusieurs mois... Markus et moi avons connu des moments difficiles lorsque tout s'obstinait à ne pas marcher et de longues heures passées à essayer de comprendre ce qui n'allait pas. Je le remercie chaleureusement pour son optimisme, sa gentillesse et son excellent français. Un grand merci également à Thierry Fusco qui a repris les mesures de diversité lors de la fin de l'intégration de NAOS-CONICA à Bellevue et à Paranal. Il nous a apporté son pragmatisme et sa très grande efficacité, il est pour beaucoup dans la fin heureuse de cette aventure.

J'ai eu la chance de pouvoir effectuer ma thèse à la fois à l'ONERA et au L2S et de rencontrer dans chacun de ces deux labos de nombreuses personnes qui n'ont jamais manqué de m'apporter une aide précieuse: Gio, Ali, Jean-Marc, Fred, Christophe, Clélia, Laurent, Bruno et Christelle. A tous un grand merci. Les bons moments vécus pendant cette thèse doivent beaucoup aux thésards et stagiaires que j'ai côtoyés (et pardon à ceux que j'aurais oubliés): Brice, Chuck, Cyril, Dolo, Emmanuel, Fred, Hicham, Marc, Myriam, Philippe et Vincent. Je vous souhaite à tous le meilleur des avenir.

Je remercie toute ma famille pour leur constant soutien au cours de mes longues études et notamment mes parents pour m'avoir donné la chance de pouvoir en faire, eux qui ne

l'ont pas eue. Une mention toute spéciale à ma maman pour son aide (logistique) plus que précieuse tout au long de ces années et sans qui je n'aurai jamais pu mener à bien ce travail. Que ces remerciements soit l'occasion pour moi d'exprimer enfin ma gratitude à Alain-Philippe à qui je dois beaucoup d'explications de maths et surtout trois précieux neurones! Enfin, merci à Thierry et Mathias d'avoir été là.

Introduction

La limite de résolution théorique d'un télescope est fixée par la diffraction et est donc inversement proportionnelle au diamètre de l'instrument. En pratique cette résolution théorique est rarement atteinte à cause de la présence d'aberrations sur le trajet optique de l'onde provenant de l'objet observé, qui viennent dégrader la qualité de l'image qui se forme au foyer de l'instrument. Dans le cas de télescopes spatiaux, ces dégradations proviennent des défauts propres de l'optique ainsi que de déformations de l'instrument dues à des perturbations extérieures. Pour les télescopes au sol, à ces défauts viennent s'ajouter les perturbations subies par l'onde à la traversée de l'atmosphère terrestre. Lors de la formation des images, ces aberrations donnent naissance à des variations de la phase optique dans la pupille de l'instrument qui atténuent la fonction de transfert. Pour s'affranchir de ces perturbations, deux méthodes peuvent être envisagées : soit une correction temps réel, soit par un traitement *a posteriori*. La correction temps réel requiert, à chaque instant, la mesure des aberrations dégradant les images et leur correction par l'intermédiaire d'un système dédié. Si l'on ne dispose pas d'un tel système, le traitement peut être effectué *a posteriori* (c'est-à-dire une fois l'image acquise), il faut, dans ce cas restaurer l'objet observé à partir de l'image enregistrée.

La diversité de phase est une extension de la technique de la « phase retrieval ». Elle permet à la fois de mesurer le front d'onde et de restaurer l'image. Cette technique repose sur l'acquisition simultanée d'images focalisée et défocalisée(s). L'avantage principal de cette méthode est d'être optiquement très simple à mettre en œuvre, de voir toutes les aberrations affectant l'image et de fonctionner sur objet très étendu. Sa difficulté majeure réside dans la résolution du problème inverse qui consiste à estimer, à partir des images, les inconnues que sont les aberrations et l'objet. Le travail proposé dans cette thèse porte sur l'amélioration de l'estimation des aberrations et de l'objet à partir des données (les images).

Le **chapitre 1** est consacré à des rappels sur la formation d'image, l'origine des aberrations et à la présentation du principe de la diversité de phase. Le **chapitre 2** propose une analyse critique et originale de la seule méthode d'estimation utilisée en diversité de phase qui repose sur la reconstruction conjointe des aberrations et de l'objet observé. En particulier, ses propriétés asymptotiques ainsi que l'influence du réglage des termes de régularisation sont étudiées. Une application expérimentale de la méthode jointe portant sur l'estimation des aberrations d'un système astronomique opérationnel est présentée **chapitre 3**. Afin de résoudre les problèmes de l'estimation jointe mise en évidence dans le chapitre 2, nous avons développé une méthode d'estimation novatrice basée sur la restauration marginale des seules aberrations : elle est présentée dans le **chapitre 4**. La qualité d'estimation des aberrations à partir des deux méthodes est comparée dans le **chapitre 5** sur données simulées et données expérimentales. Le **chapitre 6** présente et compare les méthodes de restauration possible de l'objet. Enfin dans le **chapitre 7**, une méthode originale d'estimation des phases fortement aberrantes est développée; elle est basée sur une paramétrisation pixel-à-pixel de la phase dans la pupille. Ses performances sont comparées à celles d'une méthode, plus classique, utilisant une décomposition modale de la phase.

Chapitre 1

Introduction : Imagerie et Analyse de surface d'onde

1.1 La formation d'image

1.1.1 Réponse impulsionnelle d'un instrument

Avant de rappeler les principes de la formation d'images étendues, nous allons à l'aide de la théorie de la diffraction, rapidement redémontrer l'expression de l'image que donne un système optique d'un point source, qui est par définition la réponse impulsionnelle du système (pour un développement plus complet, on pourra utilement se référer à (Goodman, 1968; Born et Wolf, 1983)).

1.1.1.1 Théorie scalaire de la diffraction

Une onde électromagnétique monochromatique s'écrit :

$$U(A,t) = u(A)e^{j[\phi(A)-\omega t]} \quad (1.1)$$

$u(A)$ est l'amplitude de l'onde à la position A et $\phi(A)$ sa phase. La partie spatiale de l'onde $U(A) = u(A)e^{j\phi(A)}$ vérifie l'équation d'Helmoltz :

$$(\nabla^2 + k^2)U = 0 \quad (1.2)$$

où $k = 2\pi/\lambda$ est le nombre d'onde (λ est la longueur d'onde) et ∇^2 le laplacien. En utilisant cette équation et le théorème de Green, Rayleigh et Sommerfeld ont développé un formalisme permettant de calculer la diffraction induite par un écran plan. Étant donné une source monochromatique située en A à gauche d'un écran plan opaque S possédant une ouverture Σ , la formule de Rayleigh-Sommerfeld permet de calculer l'amplitude complexe en un point C situé à droite du plan (voir figure 1.1).

$$U(C) = U(x,y) = \frac{1}{j\lambda} \int \int_S U_s(X,Y) \frac{\exp(-jkr)}{r} \cos(\mathbf{n},\mathbf{r}) dXdY \quad (1.3)$$

$U_s(X,Y)$ est l'amplitude complexe sur l'écran S , $\cos(\mathbf{n},\mathbf{r})$ représente le cosinus de l'angle entre la normale au plan S dirigée vers la source et le vecteur $\mathbf{r} = \mathbf{CB}$ de norme r qui s'écrit :

$$r = \sqrt{z^2 + (x - X)^2 + (y - Y)^2}. \quad (1.4)$$

Les bornes des intégrales peuvent être étendues à $\pm\infty$ en introduisant la fonction $P(X,Y)$ qui vaut 1 sur l'ouverture et 0 en dehors. Dans la suite, on pose $U'_s(X,Y) = P(X,Y)U_s(X,Y)$.

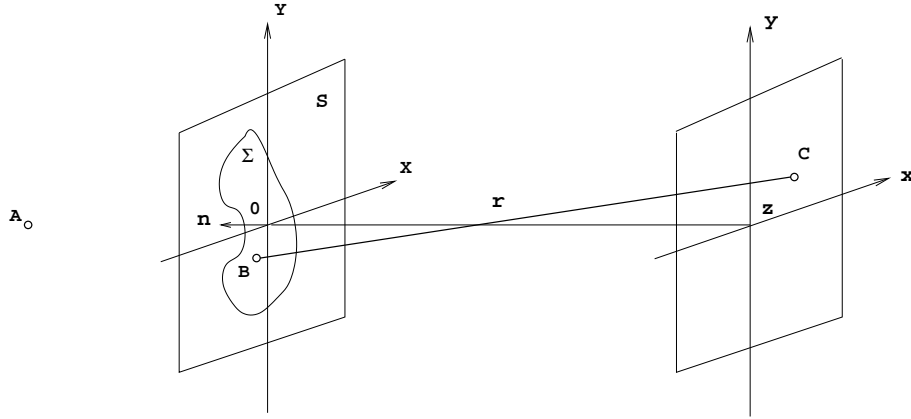


FIG. 1.1 – Diffraction à travers un écran plan

Approximation paraxiale Dans le cas où la dimension maximale de l'ouverture Σ est faible par rapport à la distance r et où la zone d'observation dans le plan (zxy) est également faible par rapport à r , on peut adopter l'approximation $\cos(\mathbf{n}, \mathbf{r}) \simeq 1$. Elle introduit une erreur inférieure à 5% pour un angle de 18° .

Dans ces mêmes conditions, r qui se trouve au dénominateur de l'expression (1.3), n'est pas très différent de z .

Approximation de Fresnel (diffraction à distance finie) Lorsque la distance d'observation z est suffisamment grande par rapport à x , X , y et Y , on peut faire l'approximation suivante :

$$r = \sqrt{z^2 \left(1 + \frac{(x - X)^2 + (y - Y)^2}{z^2} \right)} \simeq z \left(1 + \frac{(x - X)^2 + (y - Y)^2}{2z^2} \right). \quad (1.5)$$

On obtient l'approximation de Fresnel,

$$U(x, y) = \frac{\exp(-jkz)}{j\lambda z} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U'_s(X, Y) \exp \left[-j \frac{\pi}{\lambda z} ((x - X)^2 + (y - Y)^2) \right] dX dY \quad (1.6)$$

$U(x, y)$ apparaît alors, au facteur $\frac{1}{j\lambda z} \exp(-jkz)$ près, comme la convolution de $U'_s(X, Y)$ et de $\exp \left[-j \frac{\pi}{\lambda z} (X^2 + Y^2) \right]$. En remarquant que le terme $\frac{1}{j\lambda z} \exp(-jkz) \exp \left[-j \frac{\pi}{\lambda z} (X^2 + Y^2) \right]$ est l'expression d'une onde sphérique en optique paraxiale, on retrouve le principe d'Huyghens-Fresnel : chaque point du plan (OXY) émet une onde sphérique d'amplitude proportionnelle à $U'_s(X, Y)$. Ces ondes sphériques se propagent et l'amplitude en un point du plan (zxy) est la somme des amplitudes de toutes ces ondes.

En développant l'argument de l'exponentielle, l'expression (1.3) s'écrit

$$U(x, y) = \frac{\exp(-jkz)}{j\lambda z} \exp \left[-j \frac{\pi}{\lambda z} (x^2 + y^2) \right] \times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U'_s(X, Y) \exp \left[-j \frac{\pi}{\lambda z} (X^2 + Y^2) \right] \exp \left[j \frac{2\pi}{\lambda z} (xX + yY) \right] dX dY \quad (1.7)$$

On reconnaît la transformée de Fourier inverse de $U'_s(X, Y) \exp \left[-j \frac{\pi}{\lambda z} (X^2 + Y^2) \right]$ à un facteur multiplicatif près.

Approximation de Fraunhofer (diffraction à l'infini) L'expression précédente se simplifie encore lorsque que les deux plans sont suffisamment éloignés par rapport aux dimensions de l'ouverture pour que $z \gg \frac{\pi(X^2+Y^2)}{\lambda}$: le terme de phase $\exp\left[-j\frac{\pi}{\lambda z}(X^2+Y^2)\right]$ peut être pris égal à 1 et dans ce cas, l'amplitude est simplement proportionnelle à l'inverse de la transformée de Fourier de $U_s(X,Y)$ calculée aux fréquences particulières $x' = x/\lambda z$ et $y' = y/\lambda z$:

$$U(x,y) = \frac{\exp(-jkz)}{j\lambda z} \exp\left[-j\frac{\pi}{\lambda z}(x^2+y^2)\right] \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U'_s(X,Y) \exp\left[j\frac{2\pi}{\lambda z}(xX+yY)\right] dXdY \quad (1.8)$$

L'erreur commise par cette approximation est inférieure à 1% pour une ouverture circulaire de diamètre 1mm et une longueur d'onde de 0,6 μm , si la distance z est supérieure à 520 m. Les conditions de la diffraction à l'infini sont donc très sévères mais on peut ramener la distance z à une dimension plus raisonnable en observant dans le plan focal d'une lentille convergente. En effet, la transmission d'une lentille convergente parfaite de focale f positionnée dans le plan (OXY) est donnée par (Goodman, 1968) :

$$T_l(X,Y) \propto \exp\left[j\frac{\pi}{\lambda f}(X^2+Y^2)\right] \quad (1.9)$$

L'approximation de Fresnel donne l'expression de l'onde dans le plan (z,x,y) après traversée de l'ensemble pupille+lentille :

$$U(x,y) = \frac{\exp(-jkz)}{j\lambda z} \exp\left[-j\frac{\pi}{\lambda z}(x^2+y^2)\right] TF^{-1}\left[U'_s(X,Y) \exp\left(j\frac{\pi}{\lambda}(X^2+Y^2)\left(\frac{1}{f}-\frac{1}{z}\right)\right)\right] \quad (1.10)$$

avec TF^{-1} la transformée de Fourier inverse. Pour $z = f$, on retrouve la diffraction de Fraunhofer.

1.1.1.2 Application à un instrument imageur

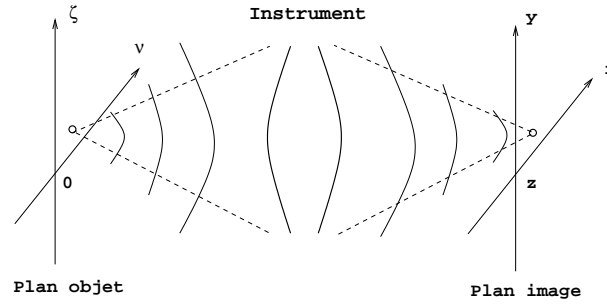
Le rôle d'un instrument imageur est de fournir des images les plus fidèles possibles d'objets. Un schéma de principe est proposé figure 1.2. Le système optique composant un instrument parfait forme à partir des ondes divergentes ou planes issues de l'objet et arrivant sur la pupille d'entrée des ondes convergentes à la sortie de la pupille. Un détecteur sensible à l'énergie de l'onde est placé dans le plan (zxy) afin d'acquérir l'image.

L'onde incidente arrivant sur l'instrument est diffractée par la pupille. Cette pupille peut être schématisée par une lentille convergente d'ouverture $P(X,Y)$ (appelée fonction pupille). Dans le cas d'instruments imageurs, l'ouverture du système est généralement assez faible pour que l'approximation paraxiale soit bien vérifiée. La diffraction de l'onde observée dans le plan focal du système est donnée par l'approximation de Fraunhofer. L'image mesurée au foyer du système est égale au module carré de l'amplitude complexe $U(x,y)$:

$$i(x',y') \propto \left| \{TF^{-1}(PU_s)\}(x',y') \right|^2 \quad (1.11)$$

L'annexe A propose une étude de la validité de cette expression le long de l'axe optique.

Cas d'un instrument parfait On dit qu'un instrument est parfait ou encore limité par la diffraction si un point source donne naissance à une onde sphérique parfaite en sortie de l'instrument qui vient converger exactement au point prévu par l'optique géométrique.

FIG. 1.2 – *Instrument imageur*

Réponse impulsionnelle (RI) La réponse impulsionnelle $h_{optique}$ d'un tel instrument est, par définition, l'image que donne cet instrument d'un point source (objet situé à l'infini et donc onde plane arrivant sur la pupille $U_s = 1$). $h_{optique}$ est simplement égale au module carré de la transformée de Fourier inverse de la fonction pupille :

$$h_{optique}(x',y') = |\{TF^{-1}(P)\}(x',y')|^2 \quad (1.12)$$

Le domaine où, en bonne approximation, la réponse impulsionnelle de l'instrument est invariante par translation est appelé domaine isoplanétique.

Fonction de transfert optique (FTO) La FTO notée $\tilde{h}_{optique}$ associée à l'instrument est donnée par la transformée de Fourier de la réponse impulsionnelle, ou encore en utilisant le théorème de Wiener-Kintchine par l'autocorrélation (notée \otimes) de la fonction pupille :

$$\tilde{h}_{optique}(u,v) = (P \otimes P)(u,v) \quad (1.13)$$

avec (u,v) les variables conjuguées de (x',y') au sens de la transformée de Fourier.

En présence d'aberrations Les aberrations entraînent des déformations de la surface d'onde en sortie de l'instrument par rapport à une forme sphérique idéale (l'origine physique de ses aberrations sera discutée au paragraphe 1.2). En présence d'aberrations, l'amplitude complexe de l'onde provenant d'un point source à l'infini sur la pupille n'est plus $U_s(X,Y) = 1$ mais $U_s(X,Y) = \Psi(X,Y) \exp j\phi(X,Y)$. On peut généraliser la fonction pupille introduite précédemment, en introduisant une fonction pupille complexe :

$$\mathbf{P}(X,Y) = P(X,Y)\Psi(X,Y) \exp j\phi(X,Y) \quad (1.14)$$

En remplaçant $P(X,Y)$ par $\mathbf{P}(X,Y)$ dans les équations (1.12) et (1.13), on obtient l'expression de la réponse impulsionnelle et de la FTO d'un système imageur en présence d'aberrations.

1.1.2 Relation entre objet et image

Dans le domaine isoplanétique, lorsque les différents points d'un objet étendu émettent des vibrations incohérentes (cas des objets astronomiques et de l'observation de la Terre depuis l'espace), l'image de l'objet est simplement donnée par la superposition des images de ses différents points (Goodman, 1968; Maréchal et Françon, 1970). Ainsi, si on note $o(\nu,\zeta)$

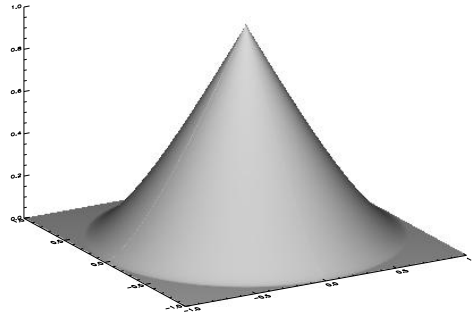


FIG. 1.3 – FTO d'un instrument parfait à pupille circulaire sans occultation (les fréquences spatiales sont exprimées en unité de D/λ).

la répartition d'éclairement de l'objet, la répartition d'éclairement de son image $i(x',y')$ est donnée par :

$$i(x',y') = \int_x \int_y o(\nu,\zeta) h_{optique}(x' - \nu, y' - \zeta) d\nu d\zeta \quad (1.15)$$

On reconnaît une relation de convolution. Dans l'espace de Fourier, la formulation de cette relation est donc particulièrement simple :

$$\tilde{i}(u,v) = \tilde{o}(u,v) \tilde{h}_{optique}(u,v) \quad (1.16)$$

avec \tilde{x} la transformée de Fourier de x et (u,v) les coordonnées dans l'espace de Fourier. Cette expression permet d'interpréter la formation d'une image comme le résultat du filtrage fréquentiel par l'instrument des composantes de Fourier de l'objet \tilde{o} . Chaque composante $\tilde{o}(u,v)$ est atténuée, lors de son passage à travers l'instrument, par le facteur de transmission $\tilde{h}_{optique}(u,v)$, qui n'est autre que la fonction de transfert optique de l'instrument. Dans le cas d'un instrument parfait, à ouverture circulaire, la réponse impulsionnelle a la forme bien connue d'une tâche d'Airy et sa FTO est représentée figure 1.3. Cette FTO est à support limité, ainsi le système se comporte comme un filtre « passe-bas » qui transmet les fréquences plus ou moins atténuées jusqu'à la fréquence de coupure $f_c = D/\lambda$ radian^{-1} (où D est le diamètre de l'ouverture et λ la longueur d'onde) et rien au-delà.

La relation (1.15) est chromatique (implicitement, elle a été donnée ici dans un cas monochromatique). Lorsque l'éclairement est polychromatique, une sommation supplémentaire, sur la largeur de la bande spectrale, est introduite :

$$i(x',y') = \int_\lambda \int_\nu \int_\zeta o(\nu,\zeta; \lambda) h_{optique}(x' - \nu, y' - \zeta; \lambda) d\nu d\zeta d\lambda \quad (1.17)$$

1.1.3 Les dégradations de l'image

En pratique, l'image qui nous est accessible est la réalisation d'une « mesure » de l'image continue $i(x',y')$. Cette mesure est une version échantillonnée et bruitée de l'image théorique.

1.1.3.1 Échantillonnage et filtrage spatial

L'image numérique est enregistrée à l'aide d'un capteur qui filtre et échantillonne spatialement l'image. Afin de prendre en compte la « réponse impulsionnelle » du détecteur, on définit la réponse impulsionnelle globale de l'instrument qui est donnée par la convolution (notée $*$) de la réponse impulsionnelle de l'optique $h_{optique}$ et de la réponse impulsionnelle du détecteur $h_{capteur}$:

$$h(x',y') = (h_{optique} * h_{capteur})(x',y') \quad (1.18)$$

On ne mesure qu'une version échantillonnée i de l'image continue i :

$$i = (o * h) \Pi_p \quad (1.19)$$

avec Π_p un peigne de Dirac de pas p . Cette discrétisation ne fera pas perdre d'information sur l'image continue si la fréquence d'échantillonnage du détecteur $f_p = 1/p$ est supérieure ou égale à 2 fois la fréquence de coupure de l'optique f_c (théorème de Shannon).

En notant o la forme discrète de l'objet et H la matrice correspondant à la convolution par la réponse impulsionnelle h (matrice de Toeplitz à blocs Toeplitz) également discrétisée, le modèle d'imagerie discret peut s'écrire sous la forme matricielle $i = Ho$.

1.1.3.2 Le bruit

L'image enregistrée est, en outre, dégradée par un bruit noté b qui est dû, d'une part, au bruit de photon (qui suit une statistique de Poisson), et d'autre part, au bruit de mesure généré par le capteur. On fera l'hypothèse (bien vérifiée dans le cas d'un CCD) que le bruit de lecture suit une loi gaussienne de moyenne nulle et de variance identique sur chaque pixel. Ces deux bruits sont décorrélés d'un pixel à l'autre. Le modèle d'imagerie bruité s'écrit :

$$i = Ho + b \quad (1.20)$$

1.2 La RI : origine des dégradations

1.2.1 Introduction

Les aberrations dégradent la qualité des images obtenues par l'instrument. En effet, elles entraînent des fluctuations de la phase et de l'amplitude, dans le plan pupille, de l'onde issue de l'objet observé. Cependant, comme nous le verrons dans le paragraphe suivant, l'effet des fluctuations d'amplitude peut être négligé face aux fluctuations de phase. Dans le cas où les aberrations sont faibles, elles induisent une atténuation de la FTO mais préservent la fréquence de coupure du système. Par contre, lorsque les aberrations deviennent importantes, elles réduisent la résolution du système et entraînent la perte totale de l'information haute fréquence. Afin de quantifier l'importance des aberrations, on utilise souvent le rapport de Strehl. Il est défini comme l'intégrale de la fonction de transfert optique $\tilde{h}_{optique}$ du système aberrant sur l'intégrale de la FTO d'un système parfait noté $\tilde{h}_{parfait}$ de même pupille.

$$SR = \frac{\int \tilde{h}_{optique}(u,v) du dv}{\int \tilde{h}_{parfait}(u,v) du dv} \quad (1.21)$$

Il est égal à un en absence d'aberrations et il est inférieur pour une image dégradée.

Les deux paragraphes suivants s'intéressent à l'origine physique des aberrations.

1.2.2 La turbulence atmosphérique

L'atmosphère terrestre est constituée de masses d'air de températures variables animées de mouvements turbulents (Kolmogorov, 1941). Ces fluctuations de température engendrent des variations de l'indice de réfraction de l'air qui entraînent des variations aléatoires du chemin optique et donc des perturbations sur une onde électromagnétique traversant ce milieu turbulent. L'étude des perturbations subies par l'onde nécessite la connaissance des fluctuations spatiales de l'indice de réfraction. La constante de structure des fluctuations d'indice $C_n^2(l)$ (Obukhov, 1949; Tatarski, 1961) caractérise la force de la turbulence à l'altitude l . Bien que les valeurs de cette constante dépendent de nombreux paramètres (topographie du terrain, conditions météorologiques, etc.), plusieurs mesures *in situ* ont permis de dégager une tendance générale des profils de turbulence. Ces études montrent que la turbulence est principalement concentrée à basse altitude (du fait des forts échanges thermiques avec le sol); la valeur du C_n^2 est importante près du sol et décroît rapidement avec l'altitude (la quasi-totalité de la turbulence est concentrée dans les 20 premiers kilomètres).

L'atmosphère terrestre engendre donc des déformations du front d'onde dues aux différences de chemin optique parcouru. De plus au cours de la propagation de l'onde, les effets de la diffraction entraînent également des fluctuations d'amplitude de l'onde (phénomène appelé scintillation). Le front d'onde aberrant dans la pupille s'écrit :

$$U(X,Y) = \Psi(X,Y) \exp j\phi(X,Y) \quad (1.22)$$

Il est possible de modéliser les variations du front d'onde comme un phénomène aléatoire et de caractériser ses propriétés asymptotiques en faisant l'hypothèse que les fluctuations de l'indice de réfraction suivent la loi de Kolmogorov.

Cas de l'observation astronomique Dans le cas d'observations d'objets astronomiques non résolus depuis un télescope au sol, l'onde avant traversée de l'atmosphère est une onde plane. On montre (Roddiér, 1981) que les fluctuations d'amplitude de l'onde peuvent être négligées devant les fluctuations de phase. La phase turbulente sur la pupille est gaussienne car le front d'onde qui arrive sur le télescope peut être vu comme la superposition (de fluctuations aléatoires) de retards subis par l'onde à la traversée d'un grand nombre d'écrans de phase statistiquement indépendants (théorème de type central limite). Le phénomène est stationnaire. On peut caractériser la phase dans la pupille d'un télescope par sa densité spectrale de puissance (DSP) :

$$DSP_\phi(f) = 0,023r_o^{-5/3} f^{-11/3} \quad (1.23)$$

où f est le module de la fréquence spatiale. La variance spatiale de la phase turbulente sur une pupille de diamètre D est donnée, quant à elle, par la relation (Noll, 1976) :

$$\sigma_\phi^2 = 1,030 (D/r_o)^{5/3} \quad (1.24)$$

Le paramètre r_o appelé paramètre de Fried (Fried, 1966) correspond au diamètre d'un télescope hors atmosphère qui aurait la même résolution qu'un télescope de diamètre infini uniquement limité par la turbulence atmosphérique. Il est lié au profil de C_n^2 par la relation :

$$r_o = \left(\frac{0,42}{\cos \gamma} \left(\frac{2\pi}{\lambda} \right)^2 \int C_n^2(l) dl \right)^{-3/5} \quad (1.25)$$

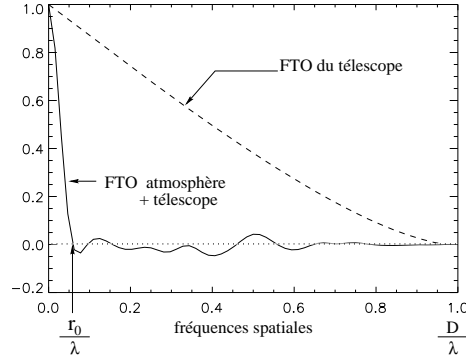


FIG. 1.4 – *En tireté, partie réelle de la FTO d'un télescope parfait à pupille circulaire sans occultation, en trait plein partie réelle de la FTO courte pose du système en présence de turbulence*

γ est l'angle zénithal et l l'altitude. L'intégrale est effectuée le long du trajet optique. La valeur de r_o est liée à la longueur d'onde, au profil de la turbulence et à la nature de l'onde qui traverse les couches turbulentes (l'expression correspond ici à une onde plane). Le temps d'évolution typique de la phase turbulente est donné par le rapport entre r_o et la vitesse moyenne du vent V (hypothèse de Taylor). Dans des conditions typiques $r_o = 10$ cm et $V = 10$ m.s⁻¹, le temps d'évolution est de 10 ms.

Le phénomène limitant la résolution d'un système dépend du rapport D/r_o : lorsque $D \ll r_o$, c'est le diamètre du télescope qui fixe la résolution maximale ($f_c = D/\lambda$), mais lorsque $D \gg r_o$, la turbulence devient le facteur limitant prépondérant. La comparaison entre la FTO d'un système en présence et en l'absence de turbulence est présentée figure 1.4 : la FTO courte pose (temps de pose < temps d'évolution de la turbulence) est caractérisée par une chute rapide des fréquences autour de r_o/λ puis pour les hautes fréquences par des oscillations autour de zéro jusqu'à la fréquence de coupure D/λ . Lorsque le temps de pose augmente, ces oscillations aléatoires se moyennent et toute information au-delà de la fréquence r_o/λ disparaît.

Cas de l'observation de la Terre depuis l'espace Dans le cas de l'observation de la surface terrestre, depuis l'espace, c'est une onde sphérique qui traverse la zone fortement turbulente de l'atmosphère. Dans ce type d'observation, les effets de la scintillation sont importants mais restent toutefois négligeables dans le cas de la résolution des instruments existants. Pour une onde sphérique, l'expression 1.25 donnant l'expression du paramètre de Fried r_o devient :

$$r_o = \left(\frac{0,42}{\cos \gamma} \left(\frac{2\pi}{\lambda} \right)^2 \int_0^L \left(\frac{l}{L} \right)^{5/3} C_n^2(l) dl \right)^{-3/5} \quad (1.26)$$

avec L l'altitude d'observation. Le profil du C_n^2 et la dépendance de r_o en $\left(\frac{l}{L} \right)^{5/3}$ montre que l'influence de la turbulence est bien moindre dans ce cas que dans le cas d'observation astronomique. Ici, r_o vaut typiquement quelques dizaines de mètres. Ainsi l'influence de la turbulence n'est pas limitante dans le cas d'un télescope spatial.

1.2.3 Les défauts propres de l'instrument

Les systèmes optiques d'observation de la Terre ou de l'espace sont des instruments de très grande qualité optique dont la résolution est limitée par la diffraction. Cependant des défauts et des déformations de l'instrument peuvent venir dégrader leurs performances. Ces défauts peuvent avoir différentes origines :

- Ils peuvent être dus à des contraintes de réalisation optique : par exemple défauts de polissage des miroirs. Ces défauts sont fixes et de fréquences spatiales élevées.
- Les déformations peuvent également être introduites par des erreurs d'alignements des composants optiques entre eux. Par exemple, dans le cas d'instruments spatiaux, l'environnement spatial (phase de lancement et rotation du télescope autour de la Terre) impose des contraintes fortes de rigidité sur le système afin d'assurer sa stabilité mécanique. Ces contraintes ne peuvent pas toutefois être assurées (afin d'alléger l'instrument notamment) et des désalignements du système peuvent apparaître. Ces défauts sont lentement variables et de fréquences spatiales faibles. S'ils peuvent être d'amplitude importante lors de la phase d'alignement, leur amplitude reste faible en cours de fonctionnement.
- Enfin des déformations peuvent apparaître à l'endroit du supportage des composants. Ces déformations évoluent lentement mais peuvent être de fréquences spatiales variables.

Même si elles ne peuvent être décrites ici car spécifiques à chaque instrument, il existe en général des informations d'ordre statistique sur les aberrations susceptibles d'être mesurées sur un instrument pour chacun des défauts évoqués.

La spécificité des aberrations engendrées par les défauts propres de l'instrument par rapport aux aberrations dues à la turbulence atmosphérique est, d'une part, leur faible amplitude et, d'autre part, leur évolution lente dans le temps (allant de quelques secondes à quelques jours).

1.2.4 Base des Zernike et décomposition de la phase

Il est souvent intéressant de décomposer la phase aberrante sur une base modale. Une base bien adaptée à la décomposition de la phase sur la pupille circulaire d'un télescope est la base des polynômes de Zernike qui est une base orthonormée sur un tel support. Ces polynômes s'expriment comme le produit d'une fonction trigonométrique et d'une fonction polynomiale (Noll, 1976).

$$Z_i(\mathbf{r}) = R_n^m(r)\Theta_n^m(\theta) \quad (1.27)$$

La fonction trigonométrique est définie par :

$$\Theta_n^m(\theta) = \begin{cases} \sqrt{n+1} & \text{si } m = 0 \\ \sqrt{2(n+1)} \cos(m\theta) & \text{si } m \neq 0 \text{ et } i \text{ pair} \\ \sqrt{2(n+1)} \sin(m\theta) & \text{si } m \neq 0 \text{ et } i \text{ impair} \end{cases} \quad (1.28)$$

et la fonction polynomiale :

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! [(n+m)/2 - s]! [(n-m)/2 - s]!} r^{n-2s} \quad (1.29)$$

polynômes de Zernike

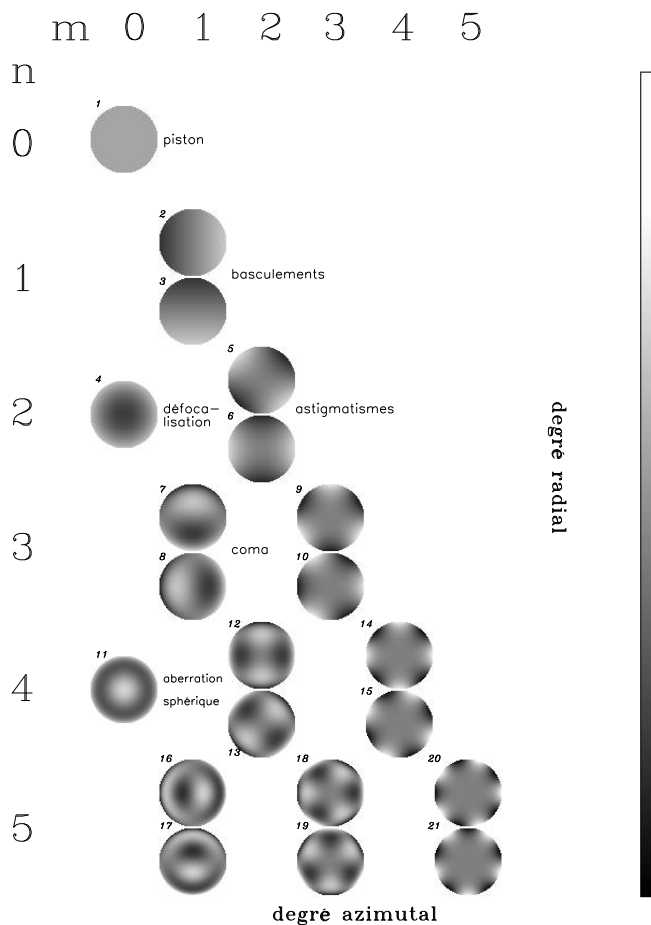


FIG. 1.5 – Représentation des premiers polynômes de Zernike

A chaque polynôme Z_i est donc associé un degré radial n et un degré azimutal m . Nécessairement, on doit avoir m et n de même parité et $0 \leq m \leq n$. Une représentation des premiers ordres est proposée figure 1.5. Plusieurs propriétés de ces polynômes, outre l'orthonormalité sur un support circulaire sans occultation, conduisent souvent à choisir cette base pour la décomposition modale de la phase :

- les bas ordre radiaux correspondent aux aberrations optiques bien connues (défocalisation, astigmatisme, coma ...)
- leur classement par ordre radial et azimutal croissant correspond au classement en fréquences spatiales croissantes.

La phase se décompose sur la base des Zernike sous la forme :

$$\phi(\mathbf{r}) = \sum_{i=1}^{+\infty} a_i Z_i(\mathbf{r}) \quad (1.30)$$

Du fait de l'orthonormalité de la « base des Zernike », la variance spatiale de la phase, hors piston, sur la pupille S s'exprime très simplement en fonction des coefficients a_i (appelés

également modes):

$$\sigma_{\phi}^2 = \frac{1}{S} \int_S \phi^2(\mathbf{r}) d\mathbf{r} = \sum_{i=2}^{+\infty} a_i^2 \quad (1.31)$$

1.3 L'analyse de surface d'onde

La présence d'aberrations, quelle que soit leur origine, entraîne une dégradation de la qualité des images enregistrées à cause de l'atténuation voire même de la disparition totale d'information sur les hautes fréquences spatiales de l'objet observé. Afin d'atteindre la résolution théorique de l'instrument, on doit estimer ces aberrations, c'est-à-dire la surface d'onde, pour en corriger les images. Ce qui rend difficile l'estimation de la phase aberrante est que les détecteurs optiques ne sont sensibles qu'à l'intensité d'une onde électromagnétique et pas à sa phase. On ne peut donc avoir qu'une mesure indirecte du front d'onde. Plusieurs méthodes ont été proposées afin d'effectuer cette mesure. Certaines reposent sur l'approximation de l'optique géométrique (localement les rayons lumineux sont perpendiculaires à la surface d'onde). Ces méthodes sont donc achromatiques. Elles permettent d'avoir accès aux gradients de la surface d'onde (strioscopie, Shack-Hartmann, senseur pyramidal). D'autres reposent sur l'analyse de figures d'interférences. Elles permettent en général d'avoir directement accès à la phase mais sont chromatiques par nature (interféromètre de Smart). La méthode de la diversité de phase peut être classée parmi ces dernières en considérant l'image fournie au foyer ou autour du foyer comme le résultat des interférences entre les ondes issues de chacun des points de la pupille. Parmi toutes les méthodes développées, très peu sont susceptibles de fonctionner avec une source infiniment étendue. Les trois principales techniques fonctionnant sur objet étendu sont le Shack-Hartmann, la Courbure et la diversité de phase¹.

1.3.1 Shack-Hartmann

Le principe de l'analyseur de surface d'onde (ASO) de Shack-Hartmann est de mesurer le gradient de la phase aberrante (Shack et Platt, 1971). Pour ce faire, le front d'onde arrivant sur la pupille est échantillonné par un réseau de micro-lentilles positionné dans un plan conjugué de la pupille d'entrée du télescope. Chaque micro-lentille forme une image de l'objet qui est enregistrée par une mosaïque de détecteurs placée au foyer des micro-lentilles. Les déformations du front d'onde induisent un déplacement relatif de l'image par rapport à une position de référence correspondant à une onde plane. Le schéma de principe est représenté figure 1.6. Une mesure de ce déplacement x donne directement la pente locale du front d'onde (Fontanella, 1985):

$$\frac{2\pi x}{\lambda f} = \frac{1}{S} \int_S \frac{\partial \phi(x, y)}{\partial x} dx dy \quad (1.32)$$

où S est la surface d'une sous-pupille et f la focale des micro-lentilles. Sur objet ponctuel, le déplacement x est estimé par calcul du centre de gravité. C'est à partir de ces mesures discrètes de pentes que la phase peut être reconstruite à un piston près. Un compromis est à trouver quant au nombre de sous-pupilles à utiliser; en effet le bruit de mesure augmente

1. On trouvera dans (Rodier, 1999), une présentation exhaustive des différents analyseurs de surface d'onde.

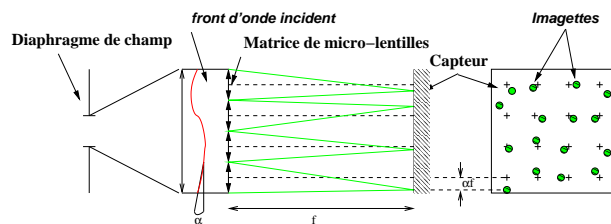


FIG. 1.6 – Principe de l'analyseur Shack-Hartmann

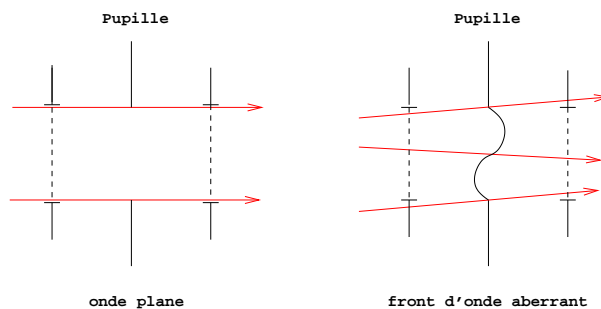


FIG. 1.7 – Principe de l'analyseur à courbure

avec le nombre de sous-pupilles mais la résolution spatiale de l'analyseur est également liée à ce nombre de sous-pupilles. On voit, d'après l'équation (1.32), que c'est une relation linéaire qui lie les mesures à la phase ce qui rend relativement simple et rapide la résolution du problème inverse qui reste à résoudre. Le Shack-Hartmann (SH) est achromatique (son fonctionnement est décrit par l'optique géométrique). Dans le cas d'objet étendu, chaque micro-lentille forme une image de la scène à son plan focal. Le déplacement de chaque imagette peut être estimé par corrélation si la scène est suffisamment texturée (Michau et al., 1992).

Le Shack-Hartmann est largement employé notamment pour l'optique adaptative en astronomie stellaire (source ponctuelle) et en astronomie solaire (source infiniment étendue) du fait de sa sensibilité et du traitement simple des données. Il peut être également utilisé pour le contrôle optique. Il est d'une grande précision sur source ponctuelle.

1.3.2 Courbure

L'analyseur à courbure est basé sur la mesure de la courbure du front d'onde (Roddiér, 1988; Roddiér et Roddiér, 1993). Son principe repose sur l'approximation de l'optique géométrique (tracé de rayon), il est rappelé figure 1.7 : une déformation de la phase ayant une courbure positive entraîne une convergence des rayons perpendiculaires à la surface d'onde en amont de la pupille créant une surintensité et une divergence en aval de la pupille entraînant une sous-intensité. La mesure de différence d'intensité entre deux plans situés de part et d'autre de la pupille fournit donc une mesure de la courbure du front d'onde. Le schéma d'une implémentation optique possible est présenté figure 1.8 : une lentille L de focale f , simulant la pupille, focalise le front d'onde au foyer F du télescope. Deux images défocalisées sont enregistrées : l'une en amont du foyer dans le plan P_1 situé à la distance l du foyer, l'autre en aval dans le plan P_2 symétrique de P_1 par rapport au foyer. Une lentille de champ L_c de focale $f/2$ assure la symétrie en réimagent la pupille L en L' . En soustrayant

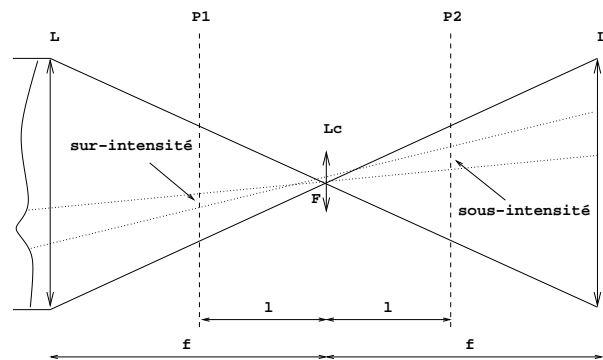


FIG. 1.8 – Schéma de l'analyseur à courbure

les valeurs d'intensité associées à chacun des plans de mesure et en normalisant par leur somme, on obtient une mesure locale de la convergence ou divergence du faisceau. A l'aide de l'optique géométrique, on montre que cette valeur est directement liée au laplacien de la phase dans la pupille.

$$\frac{I_1 - I_2}{I_1 + I_2} = \frac{f(f-l)}{2l} \left[\frac{\partial}{\partial \rho} \phi(\rho, \theta) \delta_c - \nabla^2 \phi(\rho, \theta) \right] \quad (1.33)$$

où δ_c est une distribution continue de Dirac sur le pourtour de la pupille, I_1 et I_2 sont les intensités dans les deux plans de mesures, l la distance de défocalisation, f la focale de l'instrument et ∇^2 le laplacien. La méthode est applicable à des objets d'étendue limitée. L'équation (1.33) montre que la sensibilité de l'analyseur dépend de la distance de défocalisation l : la sensibilité diminue lorsque l augmente (dépendance en l^{-1}). Le choix de l est dicté par plusieurs considérations: tout d'abord la distance de défocalisation doit être suffisante pour que l'approximation de l'optique géométrique soit vérifiée c'est-à-dire qu'on puisse négliger les effets liés à la diffraction, puis les conditions d'observations permettent de fixer la distance optimale: l devra être d'autant plus grand que l'objet est large et/ou que les aberrations recherchées sont de hauts ordres (Roddiier, 1999). La reconstruction de ϕ (résolution de l'équation (1.33)) simple comme pour le SH s'apparente à la résolution d'une équation de Poisson. Tout comme le SH, l'analyseur à courbure est achromatique. Des travaux sont en cours pour valider son fonctionnement avec une source infiniment étendue (Kupke et al., 1998).

L'analyseur à courbure a été développé et utilisé pour l'optique adaptative en astronomie stellaire. Il est particulièrement bien adapté aux systèmes ayant un nombre de degré de liberté réduit (typiquement inférieur à 100) du fait de son comportement en présence de bruit. Il a été également utilisé pour la caractérisation des aberrations des grands télescopes.

1.3.3 La diversité de phase

L'analyseur à courbure et le SH font partie des techniques « plan pupille »: leur fonctionnement est décrit par l'optique géométrique. La mesure effectuée est liée aux dérivées de la surface d'onde par une relation linéaire. L'estimation de la surface d'onde est donc relativement simple mais nécessite pour acquérir les mesures un dispositif optique assez important, annexe par rapport à la voie imagerie. On peut penser, plus simplement, à utiliser directement l'image enregistrée pour déduire la phase aberrante.

1.3.3.1 Principe

Les techniques « plan focal » sont basées sur la résolution d'un problème inverse qui consiste à estimer la phase aberrante à partir de l'image plan focal. Cette estimation (connue aussi sous le nom « phase retrieval ») a tout d'abord été développée dans le cas où l'objet observé était ponctuel (l'image est alors directement une version bruitée de la réponse impulsionnelle du système) puis adaptée au cas d'objets étendus. Une des difficultés majeures rencontrée lors de l'estimation de la phase à partir de l'image est la non-unicité de la solution trouvée. Cette indétermination est due à la relation particulière qui lie la FTO \tilde{h} à la phase dans la pupille ϕ :

$$\tilde{h}(u,v) = (Pe^{j\phi} \otimes Pe^{j\phi})(u,v) \quad (1.34)$$

Si la pupille P est à symétrie de révolution, cette relation d'autocorrélation est telle que $\phi(X,Y)$ et $\phi'(X,Y) = -\phi(-X,-Y)$ donnent la même FTO : $\tilde{h}(\phi) = \tilde{h}(\phi')$. Plus précisément, en décomposant la phase ϕ comme la somme d'une fonction paire et d'une fonction impaire, on montre que l'indétermination porte sur le signe de la partie paire de la phase ϕ . Lorsque l'on décompose la phase sur la base des polynômes de Zernike, cette indétermination conduit à ne pas pouvoir déterminer le signe global des modes d'ordre radial pair. Le détail de ces démonstrations se trouve dans l'annexe B. Ce problème d'indétermination intrinsèque de la phase a conduit Gonsalves (Gonsalves, 1976; Gonsalves, 1982) à proposer d'utiliser une deuxième image lors du processus d'inversion. L'idée est de coder de l'information supplémentaire sur la phase de cette image afin de lever les indéterminations. Pour cela, on ajoute une déformation connue ϕ_d à la surface d'onde aberrante ϕ :

$$\begin{aligned} \text{image plan focal} \quad \mathbf{i}_f &= (o * h_f) \Pi_p + \mathbf{b}_f \quad \text{avec} \quad h_f = |TF^{-1}(P \exp j\phi)|^2 \\ \text{deuxième image} \quad \mathbf{i}_d &= (o * h_d) \Pi_p + \mathbf{b}_d \quad \text{avec} \quad h_d = |TF^{-1}[P \exp j(\phi + \phi_d)]|^2 \end{aligned} \quad (1.35)$$

Les deux images \mathbf{i}_f et \mathbf{i}_d sont ainsi reliées par une relation non linéaire; c'est cette relation connue qui permet de lever les incertitudes. Cette technique s'appelle la **diversité de phase**. La diversité entre les images peut être introduite, par exemple, en défocalisant la deuxième image d'une distance connue. Cette implémentation sera utilisée par la suite (c'est celle qui est également employée dans la littérature), du fait de sa simplicité de mise en œuvre optique (il suffit de translater le détecteur). Le terme de phase ajouté est alors quadratique et permet bien de lever les indéterminations (voir Annexe B). La figure 1.9 présente le schéma de principe de la diversité de phase : une première image est acquise dans le plan focal du système et une deuxième est enregistrée simultanément dans un plan défocalisé, par ajout, dans le montage optique, d'une lame séparatrice et d'un deuxième détecteur.

La diversité de phase, contrairement aux méthodes précédentes, est sensible au chromatisme (car basée sur la diffraction). De plus, la relation non linéaire qui lie les mesures (les images) à la phase dépend également de l'objet; cette double dépendance rend le problème inverse à résoudre plus difficile que pour les analyseurs plan pupille mais permet à la diversité de phase d'être utilisable aussi bien comme analyseur de surface d'onde que comme méthode de restauration d'images. Cette technique fonctionne donc de façon très naturelle sur objet très étendu. Enfin, sa simplicité de mise en œuvre (très peu voire aucun élément optique² spécifiques à ajouter au montage optique imageur) est un avantage majeur.

2. Si le temps d'évolution de l'objet et des aberrations est assez grand pour acquérir les images successivement.

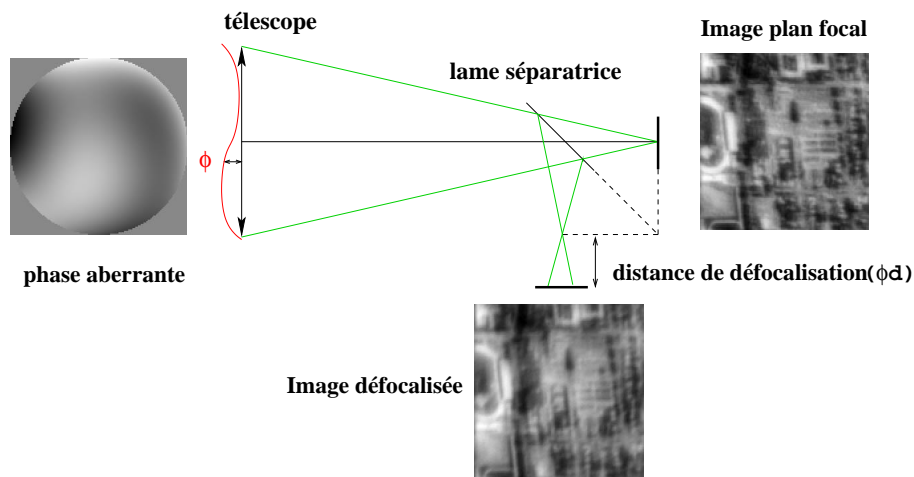


FIG. 1.9 – Schéma de principe de la diversité de phase

Bien que le concept de la diversité de phase ait été proposé il y a plus de vingt ans par Gonsalves, cette méthode n'a vraiment pris son essor qu'au cours des dix dernières années où elle a été mise en œuvre dans différentes applications utilisant directement les avantages de la diversité de phase sur les autres méthodes. En temps qu'ASO, la diversité de phase a été utilisée dans des cas où il n'était pas possible de rajouter des éléments optiques spécifiques pour effectuer les mesures : estimations des aberrations propres de systèmes optiques, soit seuls (Löfdahl et Duncan, 1998; Mugnier et al., 2001b; Mugnier et al., 2001a) (et voir chapitre 3), soit à travers la turbulence atmosphérique (Kendrick et al., 1994; Lee et al., 1997b; Thelen et al., 1999; Baba et Mutoh, 2001) ainsi que pour estimer les aberrations résiduelles d'une boucle d'optique adaptative (Lee et al., 1997a; Löfdahl et al., 2000). L'image résultant de l'interférence de tous les points de la pupille, la diversité de phase est également parfaitement adaptée au cophasage des miroirs segmentés (Löfdahl et al., 1998; Carrara et al., 2000; Seldin et Paxman, 2000). La diversité de phase a permis de faire de la restauration d'images perturbées par la turbulence atmosphérique (en utilisant un grand nombre de paires d'images courtes poses, technique appelée *phase-diverse speckle reconstruction*) (PDS); l'application principale en est l'imagerie solaire, le phénomène d'anisoplanétisme étant traité soit en « vignettant » l'image en sous-domaines isoplanétiques (Löfdahl et Scharmer, 1994; Seldin et Paxman, 1994; Acton et al., 1996), soit en reconstruisant plusieurs écrans de phase placés à différentes altitudes (Paxman et al., 1994; Thelen et al., 2000).

1.3.3.2 Comparaison avec le Shack-Hartmann et la Courbure

Les différences et les ressemblances qui existent entre les trois analyseurs qui viennent d'être présentées sont résumées dans le tableau 1.1. On pourra trouver une comparaison entre l'analyseur à courbure et le SH dans (Rousset, 1999), une comparaison de l'analyseur à courbure et la diversité de phase dans l'article (Fienup et al., 1998), cette comparaison étant faite dans le cas de la correction de la turbulence sur objet étendu, ainsi qu'une comparaison, dans un cas particulier, entre le Shack-Hartmann et la diversité de phase dans l'article (Ellerbroek et al., 1997)³.

3. Cet article s'intéresse au cas particulier d'un système d'optique adaptative utilisant une étoile laser et une étoile naturelle (la diversité et le Shack-Hartmann étant mis en concurrence pour l'estimation des aberrations de bas ordres -à partir de l'étoile naturelle- non vues par l'étoile laser).

Analyseur	Shack-Hartmann	Diversité de phase	Courbure
Mesure	$\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}$	image i_f et i_d	$\nabla^2 \phi$
Relation mesure/phase	linéaire	non linéaire	linéaire
Calibration	oui	non	non
Mise en œuvre	volumineux	très simple	volumineux
Bande spectrale	achromatique	chromatique	achromatique
Support de l'objet	pas de limitation	pas de limitation	objet compact
Grandeur estimée	surface d'onde	surface d'onde et objet	surface d'onde
Restauration	optim. linéaire	optim. non-linéaire	optim. linéaire
Temps réel OA	oui	pas encore !	oui

TAB. 1.1 – Comparaison entre le Shack-Hartmann, la diversité de phase et l'analyseur à courbure

1.3.3.3 Comparaison avec d'autres méthodes de restauration d'image

On peut citer deux articles comparant la méthode de diversité de phase avec d'autres techniques de restauration d'images. L'article de (Van Kampen et Paxman, 1998) propose une comparaison, dans le cas de l'imagerie solaire, de la diversité de phase (PDS) avec une méthode de déconvolution aveugle multi-trame. Cette méthode est basée sur l'estimation jointe de l'objet et des RI associées à une série d'images courtes poses (les données sont proches de celle utilisées en diversité de phase speckle excepté que seule l'image focalisée est utilisée dans l'inversion). Malgré le fait que le flux par image soit deux fois plus faible pour les images PDS, les cas étudiés dans cet article ont montré que la diversité de phase nécessite un moins grand nombre de trames courtes poses pour donner de bons résultats. Un autre article (Tyler et al., 1998) présente une comparaison de différentes méthodes de traitement *a posteriori* d'images corrigées ou non par un système d'optique adaptative. Ces méthodes sont, outre la méthode de PDS, une technique speckle, la déconvolution de l'objet à partir des mesures du front d'onde par AO, la déconvolution aveugle multi-trame et une méthode itérative de super-résolution utilisant les mesures AO. Pour plus de détail sur chacune de ces méthodes on se rapportera utilement à l'article et aux références qui y sont citées. En résumé, cette étude montre la supériorité de la diversité de phase dans les différents scénarios étudiés excepté dans le cas de faibles flux et de fortes aberrations non compensées par OA.

1.3.3.4 Problème inverse : choix d'un estimateur

Tout le problème de la diversité de phase réside dans la résolution du problème inverse qui consiste à estimer, à partir de l'image plan focal, de l'image défocalisée et de la distance de défocalisation, les inconnues que sont la phase et l'objet. La solution qui a été utilisée par Gonsalves (Gonsalves, 1982), lorsqu'il a proposé le concept de diversité de phase (et qui est d'ailleurs la solution la plus « naturelle ») est basée sur l'estimation par moindres carrés. Cette méthode repose sur la minimisation de l'erreur quadratique sur les données :

$$J(\mathbf{o}, \phi) = \|\mathbf{i}_f - \mathbf{H}_f(\phi)\mathbf{o}\|^2 + \|\mathbf{i}_d - \mathbf{H}_d(\phi)\mathbf{o}\|^2 \quad (1.36)$$

et donc

$$(\hat{\mathbf{o}}, \hat{\phi})_{\text{MC}} = \arg \min_{\mathbf{o}, \phi} J(\mathbf{o}, \phi) \quad (1.37)$$

Cependant, en pratique, l'utilisation de cet estimateur pour restaurer conjointement l'objet et la phase conduit à une amplification du bruit (inévitablement présent) et ne fournit donc pas une solution acceptable. Le problème inverse qui nous intéresse est donc un problème mal posé (au sens de Hadamard) (Tikhonov et Arsenin, 1977). On ne peut pas à partir des seules données, même très faiblement bruitées, estimer correctement les inconnues que sont l'objet et la phase aberrante. Il faut donc régulariser le problème inverse en introduisant des connaissances *a priori* sur les grandeurs que l'on recherche (Demoment, 1989). Les différentes méthodes de régularisation qui ont été proposées dans la littérature de la diversité de phase sont résumées dans l'introduction du chapitre suivant.

Le travail proposé dans cette thèse porte sur l'amélioration de l'estimation de la surface d'onde et de l'objet à partir des données (les images) qui comme il vient d'être rappeler est le point essentiel et difficile de la méthode de diversité de phase.

La diversité de phase peut être utilisée comme analyseur de surface d'onde (la grandeur à reconstruire est alors la phase) et comme technique de restauration d'images (ici la grandeur d'intérêt est l'objet). Ainsi suivant l'application, on ne s'intéresse qu'à l'une des deux grandeurs inconnues : objet ou phase aberrante. Cependant dans la littérature, quel que soit le problème étudié, c'est toujours une estimation conjointe objet et phase qui est effectuée. Cette méthode est présentée dans le chapitre 2 : une interprétation probabiliste de cette estimation conjointe est présentée et nous proposons une étude originale des propriétés de cet estimateur. Nous montrons en particulier qu'il n'a pas de bonnes propriétés asymptotiques et que le réglage des hyperparamètres est difficile, surtout pour les faibles rapports signal à bruit (RSB).

Une application expérimentale de cette méthode jointe est proposée dans le chapitre 3, elle porte sur l'estimation des aberrations de la caméra infrarouge CONICA.

Afin de résoudre les problèmes asymptotiques de l'estimation jointe ainsi que ceux liés au réglage des hyperparamètres, on propose une nouvelle méthode totalement novatrice dans le cadre de la diversité de phase, d'estimation des seules aberrations qui s'affranchit de la restauration de l'objet en l'intégrant hors du problème : l'estimateur marginal ainsi obtenu est présenté chapitre 4. On établit la relation qui lie le critère marginal au critère joint et on étudie ses propriétés asymptotiques. Une méthode d'estimation non supervisée des hyperparamètres est proposée et validée sur simulations. Enfin, on valide sur données expérimentales le fonctionnement de ce nouvel estimateur.

Le chapitre 5 est consacré à une comparaison sur simulations et données expérimentales de la restauration d'aberrations obtenues par chacune des deux méthodes pour différentes conditions (objet observé, RSB, etc.).

Après le problème de l'estimation des aberrations, le chapitre 6 aborde le cas de la restauration de l'objet. Les solutions proposées par la méthode jointe et marginale sont présentées : on montre notamment que l'estimateur marginal permet très simplement de restaurer également l'objet.

Le dernier chapitre aborde le problème spécifique des fortes aberrations (d'amplitude supérieure à 2π). Leur estimation est rendue difficile par l'absence d'information dans les images autre que sur la phase modulo 2π . Après avoir passé en revue les différentes façons d'aborder ce problème, une méthode originale de reconstruction point-à-point de la phase modulo 2π est développée. Ces performances sont comparées, sur données simulées, avec celles d'une méthode classique basée sur la décomposition modale de la phase aberrante sur la base des Zernike.

Chapitre 2

Estimation conjointe des aberrations et de l'objet

2.1 Introduction

Les données traitées en diversité de phase (image focalisée et défocalisée) contiennent à la fois de l'information sur l'objet et sur la surface d'onde. Mais suivant l'application étudiée, seule une des deux grandeurs phase ou objet est le paramètre d'intérêt : la phase pour l'ASO (correction temps réel), l'objet pour la restauration d'image (traitement *a posteriori*) mais les deux paramètres étant inconnus, toute la littérature de la diversité de phase repose sur une estimation conjointe de l'objet et de la phase aberrante. L'estimation conjointe à laquelle on s'intéresse est un problème mal-posé et doit donc être régularisée (Bucci et al., 1999). Différentes méthodes de régularisation ont été proposées. Vis-à-vis de la phase, une première régularisation est introduite en recherchant la phase via sa décomposition tronquée sur la base des Zernike. Dans le cas d'imagerie à travers la turbulence (Thelen et al., 1999; Lee et al., 1997b) ont introduit un terme de régularisation explicite sur les coefficients de Zernike en utilisant des connaissances statistiques sur la phase turbulente. (Löfdahl et Scharmer, 1994; Lee et al., 1997b) utilisent un filtre passe-bas pour éviter l'amplification du bruit. Seldin (Seldin et Paxman, 1994) propose de paramétrer l'objet par une fonction convoluée avec une gaussienne. Une régularisation quadratique sur l'objet est utilisée dans les articles (Vogel et al., 1998; Bucci et al., 1999). Enfin on trouve des méthodes très ad-hoc de régularisation par arrêt de la minimisation avant convergence (Thelen et al., 1999).

Nous proposons, dans ce chapitre, une interprétation probabiliste de la reconstruction conjointe de l'objet et de la phase aberrante. Cette approche est basée sur la recherche du Maximum *A posteriori* Joint (MAPJ). Après avoir présenté cet estimateur, on montre, à l'aide de simulations, ses propriétés asymptotiques et enfin on s'intéresse à l'influence de la valeur des hyperparamètres sur la qualité de restauration de la phase et de l'objet. La présentation de l'estimateur joint qui est faite dans ce chapitre est très générale. Pour les simulations, nous nous sommes placés dans le cas particulier de l'observation de la Terre depuis l'espace.

2.2 L'estimateur MAPJ

2.2.1 Construction du critère MAPJ

L'approche bayésienne permet de fusionner données et connaissances *a priori* sur les grandeurs recherchées : l'objet \mathbf{o} et la phase ϕ . Ces connaissances sont introduites sous la forme de densités de probabilité *a priori*. L'estimation au sens du MAPJ consiste à choisir l'objet et la phase qui maximisent la densité de probabilité jointe des données (les images \mathbf{i}_f et \mathbf{i}_d), de l'objet \mathbf{o} , et de la phase aberrante ϕ . Cette densité dépend d'hyperparamètres $\boldsymbol{\theta}$ ($\boldsymbol{\theta}$ représente tous les paramètres autres que \mathbf{o} et ϕ qui sont liés à l'estimation).

$$(\hat{\mathbf{o}}, \hat{\phi})_{\text{MAPJ}} = \arg \max_{\mathbf{o}, \phi} f(\mathbf{i}_d, \mathbf{i}_f, \mathbf{o}, \phi; \boldsymbol{\theta}) \quad (2.1)$$

La règle de Bayes lie données et *a priori* en exprimant la densité de probabilité jointe comme étant égale au produit :

$$f(\mathbf{i}_f, \mathbf{i}_d, \mathbf{o}, \phi; \boldsymbol{\theta}) = f(\mathbf{i}_f, \mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta}) f(\mathbf{o}, \phi; \boldsymbol{\theta}) = f(\mathbf{i}_f, \mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta}) f(\mathbf{o}; \boldsymbol{\theta}) f(\phi; \boldsymbol{\theta}) \quad (2.2)$$

avec $f(\mathbf{i}_f, \mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta})$ la densité de probabilité conditionnelle d'avoir les images \mathbf{i}_f et \mathbf{i}_d sachant l'objet \mathbf{o} , la phase ϕ et les hyperparamètres $\boldsymbol{\theta}$ (c'est le terme de vraisemblance, il contient le modèle de dégradation et de bruit), $f(\mathbf{o}; \boldsymbol{\theta})$ la densité de probabilité *a priori* de l'objet et $f(\phi; \boldsymbol{\theta})$ la densité de probabilité *a priori* de la phase. Les images \mathbf{i}_f et \mathbf{i}_d sont statistiquement indépendantes à \mathbf{o} et ϕ fixés; la vraisemblance de l'ensemble des deux images peut donc s'écrire comme le produit des vraisemblances individuelles de chaque image : $f(\mathbf{i}_f, \mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta}) = f(\mathbf{i}_f | \mathbf{o}, \phi; \boldsymbol{\theta}) f(\mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta})$. Les estimées ($\hat{\mathbf{o}}$ et $\hat{\phi}$) au sens du MAPJ sont donc celles qui minimisent le critère :

$$J(\mathbf{o}, \phi) = \underbrace{-\ln f(\mathbf{i}_f | \mathbf{o}, \phi; \boldsymbol{\theta}) - \ln f(\mathbf{i}_d | \mathbf{o}, \phi; \boldsymbol{\theta})}_{\text{termes d'attache aux données}} - \underbrace{\ln f(\phi; \boldsymbol{\theta})}_{\text{a priori sur la phase}} - \underbrace{\ln f(\mathbf{o}; \boldsymbol{\theta})}_{\text{a priori sur l'objet}} \quad (2.3)$$

2.2.1.1 Terme d'attache aux données

Le bruit de mesure a deux origines : le bruit de détecteur qui suit une loi de probabilité gaussienne blanche et le bruit de photon qui suit une loi de probabilité poissonnienne. Avec une assez bonne approximation, on peut modéliser ces deux sources de bruits par un bruit blanc gaussien homogène. En effet à faible flux, c'est le bruit de détecteur qui domine (il est alors effectivement gaussien) et à fort flux le bruit dominant est le bruit de photon. En première approximation, dans le cas d'objets brillants et étendus, ce bruit de photon peut être modélisé par un bruit blanc gaussien homogène. Cette approximation est, en particulier, assez légitime pour l'imagerie de la Terre depuis l'espace.

$$f(\mathbf{i}_k | \mathbf{o}, \phi; \boldsymbol{\theta}) = \frac{1}{(2\pi)^{N^2/2} \sigma^{N^2}} \exp\left(-\frac{1}{2\sigma^2} (\mathbf{i}_k - H_k \mathbf{o})^t (\mathbf{i}_k - H_k \mathbf{o})\right) \quad k = \{f, d\} \quad (2.4)$$

N^2 est le nombre de pixels de l'image et σ l'écart type du bruit (on suppose que l'écart type du bruit est le même pour les deux images). Le critère d'attache aux données, noté J_b , est obtenu en prenant l'opposé du logarithme de $f(\mathbf{i}_k | \mathbf{o}, \phi; \boldsymbol{\theta})$.

2.2.1.2 Terme de régularisation sur la phase

Une première façon de régulariser le problème vis-à-vis de la phase est de réduire la dimension de l'espace des inconnus $\boldsymbol{\phi}$ c'est-à-dire de reconstruire la phase via sa décomposition tronquée sur la base de Zernike. On écrit :

$$\boldsymbol{\phi}(\mathbf{r}) = \sum_{i=2}^k a_i Z_i(\mathbf{r}) \quad (2.5)$$

On rappelle que ces polynômes sont rangés dans un ordre correspondant à des fréquences spatiales croissantes. Les aberrations propres d'un instrument étant plutôt des défauts basses fréquences, une vingtaine de modes suffisent généralement à décrire correctement la phase aberrante. Dans le cas d'aberrations induites par la turbulence atmosphérique, le choix du nombre de modes à reconstruire est moins évident. On utilise alors les connaissances statistiques que l'on possède sur la turbulence (qui ont été présentées au paragraphe 1.2.2) pour introduire un terme d'*a priori* sur les coefficients de Zernike. Dans ce cas, les coefficients turbulents suivent une loi gaussienne de moyenne nulle et de matrice de covariance R_a :

$$f(\boldsymbol{\phi}(\mathbf{a}); \boldsymbol{\theta}) = \frac{1}{(2\pi)^{(k-1)/2} \det(R_a)^{1/2}} \exp\left(-\frac{1}{2} \mathbf{a}^t R_a^{-1} \mathbf{a}\right) \quad (2.6)$$

où $\det(R_a)$ représente le déterminant de la matrice R_a . Dans la suite désormais (excepté au chapitre 7), on notera \mathbf{a} la phase inconnue pour rappeler que $\boldsymbol{\phi}(\mathbf{r})$ est décomposée sur la base des Zernike.

2.2.1.3 Terme de régularisation sur l'objet

On choisit d'associer à l'objet \mathbf{o} une loi de probabilité gaussienne de moyenne \mathbf{o}_m et de matrice de covariance R_o .

$$f(\mathbf{o}; \boldsymbol{\theta}) = \frac{1}{(2\pi)^{N^2/2} \det(R_o)^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{o} - \mathbf{o}_m)^t R_o^{-1} (\mathbf{o} - \mathbf{o}_m)\right) \quad (2.7)$$

On peut, de plus, imposer une contrainte de positivité sur l'objet (justifiée puisque l'objet est une carte d'intensité) par reparamétrisation (Biraud, 1969; Thiébaud et Conan, 1995) ou par projection (Commenges, 1984; Nakamura et al., 1988) par exemple. Cependant lorsque l'objet observé est très étendu (i.e. occupe toute l'image), cette contrainte ne modifie pas les estimées. Sa prise en compte n'est pas difficile mais alourdit de façon importante le calcul pratique des quantités estimées (voir fin du paragraphe 2.2.1.4). C'est pourquoi dans la suite où nous nous sommes intéressés aux cas de scènes terrestres, cette contrainte n'a pas été imposée.

2.2.1.4 Critère MAPJ

Le critère MAPJ est donné par l'opposé du logarithme du produit des quatre densités de probabilité présentées précédemment :

$$\begin{aligned} J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}) &= -\ln f(\mathbf{i}_f, \mathbf{i}_d, \mathbf{o}, \mathbf{a}; \boldsymbol{\theta}) \\ &= N^2 \ln \sigma^2 + \frac{1}{2} \ln \det(R_o) + \frac{1}{2} \ln \det(R_a) + \frac{1}{2\sigma^2} (\mathbf{i}_f - H_f \mathbf{o})^t (\mathbf{i}_f - H_f \mathbf{o}) \\ &\quad + \frac{1}{2\sigma^2} (\mathbf{i}_d - H_d \mathbf{o})^t (\mathbf{i}_d - H_d \mathbf{o}) + \frac{1}{2} (\mathbf{o} - \mathbf{o}_m)^t R_o^{-1} (\mathbf{o} - \mathbf{o}_m) + \frac{1}{2} \mathbf{a}^t R_a^{-1} \mathbf{a} + A \end{aligned} \quad (2.8)$$

Où A est une constante, que nous ne noterons plus dans la suite. Ce critère doit être minimisé vis-à-vis des coefficients aberrants \mathbf{a} et de l'objet \mathbf{o} . L'annulation de la dérivée partielle de J_{MAPJ} par rapport à l'objet fournit une expression analytique de l'objet qui minimise le critère à \mathbf{a} fixé :

$$\hat{\mathbf{o}}(\mathbf{a}) = (H_f^t H_f + H_d^t H_d + \sigma^2 R_o^{-1})^{-1} (H_f^t \mathbf{i}_f + H_d^t \mathbf{i}_d + \sigma^2 R_o^{-1} \mathbf{o}_m) \quad (2.9)$$

En introduisant $\hat{\mathbf{o}}(\mathbf{a})$ dans le critère J_{MAPJ} , on obtient un « nouveau » critère J'_{MAPJ} qui ne dépend plus explicitement de l'objet.

Il n'existe malheureusement pas d'expression analytique pour les coefficients aberrants \mathbf{a} . La minimisation de J'_{MAPJ} devra alors être faite de façon itérative. Notons que dans le cas de la prise en compte d'une contrainte de positivité sur l'objet, $\hat{\mathbf{o}}(\mathbf{a})$ et donc J'_{MAPJ} n'ont plus d'expression explicite et donc que J doit être minimisé numériquement en \mathbf{o} et \mathbf{a} .

2.2.2 Écriture du critère dans le domaine de Fourier

Les matrices H_f et H_d représentant un opérateur de convolution sont des matrices de Toeplitz à blocs Toeplitz. En supposant que $\mathbf{o} - \mathbf{o}_m$ est stationnaire, la matrice de covariance de l'objet R_o est également une matrice de Toeplitz à blocs Toeplitz. Ces matrices peuvent être approchées par des matrices circulantes à blocs circulants (l'approximation correspondant à une périodisation). Les matrices carrées circulantes ont des propriétés intéressantes :

- Une matrice circulante est diagonalisable,
- la base de ses vecteurs propres est la base des exponentielles de Fourier discrètes,
- les valeurs propres de H_f et H_d sont les valeurs de la fonction de transfert du système obtenue par Transformée de Fourier discrète de la RI.

Ces propriétés sont généralisables aux matrices circulantes à blocs circulants, on peut donc écrire :

$$\begin{aligned} R_o &= F^{-1} \text{diag}[S_o] F \\ H_d &= F^{-1} \text{diag}[\tilde{h}_d] F \\ H_f &= F^{-1} \text{diag}[\tilde{h}_f] F \end{aligned}$$

avec F la matrice de la transformée de Fourier 2D, $\text{diag}[x]$ représente une matrice diagonale ayant x sur sa diagonale, \tilde{x} la transformée de Fourier 2D de x et S_o la densité spectrale de puissance (DSP) de l'objet.

Le critère J_{MAPJ} ainsi que $\hat{\mathbf{o}}(\mathbf{a})$ peuvent donc s'écrire dans le domaine de Fourier discret moyennant l'approximation de périodisation :

$$\begin{aligned} J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}) &= N^2 \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) + \frac{1}{2} \ln \det(R_a) + \frac{1}{2} \mathbf{a}^t R_a^{-1} \mathbf{a} + \sum_v \frac{|\tilde{o}(v) - \tilde{o}_m(v)|^2}{2S_o(v)} \\ &+ \sum_v \frac{1}{2\sigma^2} |\tilde{i}_f(v) - \tilde{h}_f(\mathbf{a}, v) \tilde{o}(v)|^2 + \sum_v \frac{1}{2\sigma^2} |\tilde{i}_d(v) - \tilde{h}_d(\mathbf{a}, v) \tilde{o}(v)|^2 \end{aligned} \quad (2.10)$$

$$\text{et } \hat{\mathbf{o}}(\mathbf{a}, v) = \frac{\tilde{h}_f^*(\mathbf{a}, v) \tilde{i}_f(v) + \tilde{h}_d^*(\mathbf{a}, v) \tilde{i}_d(v) + \frac{\sigma^2 \tilde{o}_m(v)}{S_o(v)}}{|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\sigma^2}{S_o(v)}}. \quad (2.11)$$

avec v la fréquence spatiale 2D. Le critère J'_{MAPJ} s'écrit alors :

$$\begin{aligned}
J'_{\text{MAPJ}}(\mathbf{a}) &= J_{\text{MAPJ}}(\hat{o}(\mathbf{a}), \mathbf{a}) \\
&= N^2 \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) + \frac{1}{2} \ln \det(R_a) + \frac{1}{2} \mathbf{a}^t R_a^{-1} \mathbf{a} \\
&+ \frac{1}{2} \sum_v \frac{\left(|\tilde{i}_d(v) \tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v) \tilde{h}_d(\mathbf{a}, v)|^2 \right) + \frac{\sigma^2}{S_o(v)} \left(|\tilde{h}_f(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right)}{\sigma^2 (|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\sigma^2}{S_o(v)})}
\end{aligned} \tag{2.12}$$

La méthode jointe dans la littérature La méthode jointe a fait l'objet de différentes interprétations. Historiquement, Gonsalves (Gonsalves, 1982) a proposé une estimation de \mathbf{o} et \mathbf{a} par la méthode des moindres carrés. L'estimation au sens du Maximum de Vraisemblance a été développée par Paxman (Paxman et al., 1992) d'une part pour un bruit gaussien (on retrouve alors l'expression donnée par les moindres carrés) d'autre part pour un bruit Poissonnien (c'est dans cet article que l'on peut trouver le calcul analytique du gradient du critère MV par rapport aux coefficients aberrants \mathbf{a}). Bucci (Bucci et al., 1999) a adopté une approche déterministe pour régulariser le problème. Dans l'article de Thelen et Paxman (Thelen et al., 1999) apparaît une interprétation probabiliste de cette méthode jointe. Il est à noter que contrairement à ce qui est appelé estimation par MAP dans ce dernier article, ce n'en est en réalité pas puisque si il y a bien un terme d'*a priori* associé aux aberrations, aucun n'est introduit pour l'objet, il s'agit donc plutôt de Maximum de Vraisemblance généralisé. A notre connaissance, aucun article ne propose une vraie estimation au sens du MAPJ.

2.2.3 Choix et réglage des hyperparamètres

2.2.3.1 Hyperparamètres liés à l'*a priori* sur la phase

Les hyperparamètres θ_ϕ liés au terme J_a (*a priori* sur les coefficients turbulents) peuvent être facilement déduits de la statistique de Kolmogorov (voir paragraphe 1.2.2). Noll (Noll, 1976) a montré que la matrice de covariance des coefficients turbulents s'écrit :

$$\begin{aligned}
\langle a_i a_j \rangle &= 3,90 \left(\frac{D}{r_o} \right)^{\frac{5}{3}} [(n_i + 1)(n_j + 1)]^{\frac{1}{2}} (-1)^{(n_i + n_j - 2m_i)/2} \delta_{m_i m_j} \\
&\times \frac{2^{-\frac{14}{3}} \Gamma[\frac{14}{3}] \Gamma[(n_i + n_j - \frac{14}{3} + 3)/2]}{\Gamma[(-n_i + n_j + \frac{14}{3} + 1)/2] \Gamma[(n_i - n_j + \frac{14}{3} + 1)/2] \Gamma[(n_i + n_j + \frac{14}{3} + 3)/2]}
\end{aligned} \tag{2.13}$$

où D est le diamètre de la pupille, n_i et m_i (resp. n_j et m_j) les ordres radiaux et azimutaux du polynôme i (resp. j), $\delta_{m_i m_j}$ le symbole de Kronecker et $\Gamma(x)$ représente la fonction gamma. Cette matrice est représentée figure 2.1 (gauche) pour un $D/r_o = 1$ (l'échelle est logarithmique). La diagonale de cette matrice correspond à la variance des coefficients (voir figure 2.1 à droite); on remarque qu'il existe également des liens statistiques entre certains coefficients de la décomposition, mais la valeur de ces corrélations est faible par rapport à la variance. Le calcul de cette matrice nécessite uniquement de connaître la valeur de D/r_o . Il introduit donc un hyperparamètre r_o mais dont la valeur peut être estimée. En effet, r_o

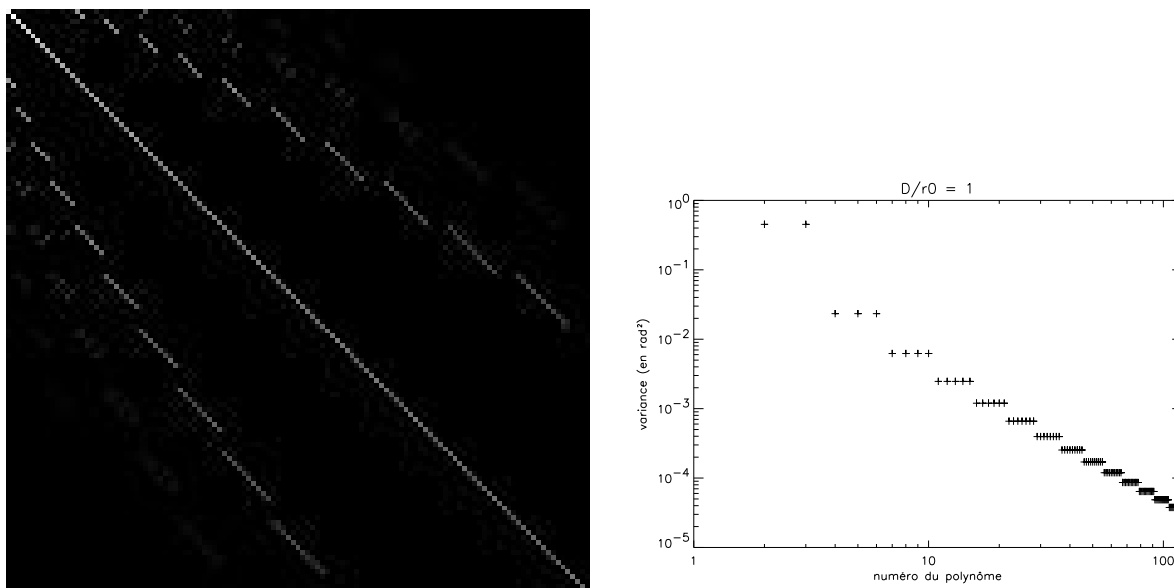


FIG. 2.1 – La matrice de covariance des 120 premiers polynômes de Zernike à gauche et leur variance à droite (échelle logarithmique pour les deux graphes).

peut être connu, au cours d'observation, en effectuant une mesure par un *seeing monitor* ou par SCIDAR si un de ces instruments est disponible. Ou, dans le cas d'un système d'optique adaptative, il peut être déduit des mesures des résidus de turbulence et des commandes.

2.2.3.2 Hyperparamètres liés aux termes d'attache aux données

Les hyperparamètres θ_b liés au terme J_b se limitent à l'écart type du bruit σ . Celui-ci peut être estimé à partir du flux total dans l'image dans l'hypothèse d'un bruit essentiellement photonique.

2.2.3.3 Hyperparamètres liés à l'*a priori* objet

On note θ_o les hyperparamètres introduits par J_o .

Choix du modèle de DSP objet On choisit le modèle paramétrique suivant pour la densité spectrale de puissance de l'objet S_o (Kattinig et Primot, 1997) :

$$\begin{aligned} S_o(v) &= E[|\tilde{o}(v) - \tilde{o}_m(v)|^2] \\ &= k/[v_o^p + v^p] - |\tilde{o}_m(v)|^2 \end{aligned} \quad (2.14)$$

Ce modèle fait intervenir quatre hyperparamètres k , p , v_o et o_m .

Réglage empirique des hyperparamètres objet Le réglage de ces hyperparamètres est bien plus difficile à effectuer que celui des hyperparamètres liés à la phase puisque, bien sûr, il sera différent pour chaque objet observé. Sur un objet compact (planète, galaxie,...), le réglage supervisé des hyperparamètres peut être approché en utilisant des considérations physiques sur l'objet : k est approximativement donné par le carré du nombre de photons contenus dans l'image, v_o est égal à l'inverse de la taille caractéristique de l'objet et la

puissance p caractérise la régularité de l'objet (généralement comprise entre 1 et 4). En faisant l'hypothèse d'ergodicité, l'objet moyen o_m peut être estimé en faisant une moyenne locale sur l'image. Cette estimation « à la main » est forcément approximative mais a donné d'assez bons résultats dans le cas d'objets à support fini (Conan et al., 1998). Elle s'avère bien plus difficile à effectuer sur objet très étendu. La solution idéale serait d'estimer ces hyperparamètres de façon non-supervisée, conjointement avec les aberrations et l'objet.

Estimation des hyperparamètres objet : dégénérescence du critère L'estimation conjointe des hyperparamètres θ_o avec l'objet et les aberrations entraîne une dégénérescence du critère. En effet, en reprenant l'expression (2.8) du critère MAPJ et en posant $R_o = \sigma_o^2 R'_o$ avec $\sigma_o \in \mathfrak{R}$, on remarque que le couple $\{\sigma_o = 0, o = o_m\}$ est solution (le critère tend alors vers $-\infty$) mais cette solution est aberrante puisqu'elle ne dépend pas des données. L'estimation non-supervisée de la valeur de ces hyperparamètres n'est donc pas possible de cette façon.

Le réglage des hyperparamètres liés à l'objet reste donc un problème majeur lors de l'utilisation de cet estimateur. Nous verrons au paragraphe 2.4 l'influence de la valeur de ces hyperparamètres sur la qualité de restauration de l'objet et des coefficients aberrants.

2.2.4 Minimisation du critère

L'objet et les coefficients sont estimés par minimisation du critère J_{MAPJ} . Dans le cas où l'on peut utiliser l'expression analytique (2.11) pour l'objet, seule la minimisation vis-à-vis des coefficients \mathbf{a} est faite numériquement. Mais lorsque cette expression n'est pas utilisable : ajout d'une contrainte de positivité sur l'objet ou objet très étendu (l'approximation circulante n'est alors plus vérifiée), la minimisation vis-à-vis des deux grandeurs est effectuée itérativement. Cette minimisation numérique est faite en utilisant une méthode de type gradient conjugué. Le critère à minimiser n'étant pas quadratique en \mathbf{a} , on prend soin régulièrement de réinitialiser la direction de descente (en utilisant alors comme direction de descente le gradient et non le gradient conjugué) ce qui a pour effet d'accélérer la convergence.

De façon pratique, même dans le cas de l'existence de \hat{o} et donc de la forme J'_{MAPJ} du critère joint, c'est la forme non intégrée J_{MAPJ} ainsi que son gradient par rapport aux aberrations qui sont utilisés en se plaçant en $\mathbf{o} = \hat{o}(\mathbf{a})$. En effet, en utilisant le gradient d'une fonction composée, on peut écrire :

$$\frac{\partial J'_{\text{MAP}}(\mathbf{a})}{\partial a_i} = \frac{\partial J_{\text{MAP}}(\hat{o}(\mathbf{a}), \mathbf{a})}{\partial a_i} = \sum_{k,l} \frac{\partial J_{\text{MAP}}(\mathbf{o}, \mathbf{a})}{\partial o_{k,l}} \cdot \frac{\partial o_{k,l}}{\partial a_i} \Big|_{\mathbf{o}=\hat{o}} + \sum_j \frac{\partial J_{\text{MAP}}(\mathbf{o}, \mathbf{a})}{\partial a_j} \Big|_{\mathbf{o}=\hat{o}} \frac{\partial a_j}{\partial a_i} \quad (2.15)$$

Par définition, le premier membre de la somme est nulle et donc :

$$\frac{\partial J'_{\text{MAP}}(\mathbf{a})}{\partial a_i} = \frac{\partial J_{\text{MAP}}(\mathbf{o}, \mathbf{a})}{\partial a_i} \Big|_{\mathbf{o}=\hat{o}} \quad (2.16)$$

Ainsi à chaque étape de la minimisation, on dispose à la fois des estimées des aberrations et de l'estimée de l'objet.

Lorsque la minimisation des deux inconnues est effectuée itérativement, une minimisation globale du critère vis-à-vis des aberrations et de l'objet (i.e. \mathbf{a} et \mathbf{o} sont empilés dans un même vecteur et le critère minimisé vis-à-vis de ce vecteur) pose des problèmes de différence

d'échelle entre les gradients des deux grandeurs à reconstruire. Ce problème est dû à la nature très différente des deux inconnues recherchées : coefficients de Zernike exprimés en radian pour les aberrations et carte d'intensité en photons pour l'objet. On peut remédier à cela en reparamétrant l'une des deux grandeurs (en posant, par exemple, $o = mo'$ avec m une constante) ce qui a pour effet de multiplier le gradient par une constante de telle sorte que le poids des deux gradients soit similaire. Mais une méthode plus simple et plus rapide en pratique est de minimiser alternativement le critère vis-à-vis de l'objet à aberrations fixées, puis de le minimiser vis-à-vis des aberrations à objet fixé. Chacune de ces deux étapes ne nécessite pas le calcul de tous les termes composant le critère J_{MAPJ} : lorsque la minimisation est effectuée vis-à-vis de \mathbf{a} , le terme $J_o(\mathbf{o}; \boldsymbol{\theta}_o)$ ne change pas et lors de la minimisation objet à \mathbf{a} fixé, le terme $J_a(\mathbf{a}; \boldsymbol{\theta}_o)$ demeure inchangé. A l'étape k , cette méthode de minimisation « alternée » peut se résumer ainsi :

$$\begin{cases} \hat{\mathbf{o}}_k = \arg \min_{\mathbf{o}} J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}_{k-1}; \boldsymbol{\theta}) = \arg \min_{\mathbf{o}} [J_b(\mathbf{o}, \mathbf{a}_{k-1}; \boldsymbol{\theta}) + J_o(\mathbf{o}; \boldsymbol{\theta})] \\ \hat{\mathbf{a}}_k = \arg \min_{\mathbf{a}} J_{\text{MAPJ}}(\mathbf{o}_k, \mathbf{a}; \boldsymbol{\theta}) = \arg \min_{\mathbf{a}} [J_b(\mathbf{o}_k, \mathbf{a}; \boldsymbol{\theta}) + J_a(\mathbf{a}; \boldsymbol{\theta})] \end{cases} \quad (2.17)$$

Les méthodes de minimisation envisagées ici sont des méthodes locales : la convergence sera atteinte dès lors qu'un minimum est trouvé mais sans garantie aucune que ce minimum soit le minimum global. Notre critère n'étant *a priori* pas convexe (voir paragraphe 2.3.2.2), le choix du point de départ de la minimisation est donc primordial : ce point doit être le plus proche possible de la vraie solution. En pratique, on part d'un jeu de coefficients nuls et pour l'objet, quand celui-ci est également estimé par minimisation numérique, de l'image plan focal.

2.3 Propriétés asymptotiques

Les propriétés de l'estimateur MAPJ n'ont jamais été étudiées dans le cadre de la diversité de phase. Nous nous proposons ici de montrer pour différents niveaux de bruit, l'évolution de la qualité de restauration de l'objet et des aberrations lorsque le nombre de données traitées augmente.

2.3.1 Définitions : biais, variance et écart quadratique moyen

Deux paramètres permettent de caractériser l'écart d'un estimateur \hat{o} à la valeur vraie o_{vrai} de la grandeur recherchée : le biais et la variance.

Biais Le biais associé à l'estimateur \hat{o} est l'écart moyen entre valeur vraie o_{vraie} et valeur estimée :

$$b_{i|o}(\hat{o}) = E_{i|o}[\hat{o} - o_{\text{vrai}}] = E_{i|o}[\hat{o}] - o_{\text{vrai}} \quad (2.18)$$

$E_{i|o}[\cdot]$ représente l'espérance mathématique selon la loi de probabilité des mesures i sachant o c'est-à-dire la loi du bruit.

Variance La matrice de covariance indique la dispersion des valeurs estimées autour de la valeur moyenne.

$$C_{i|o}(\hat{o}) = E_{i|o}[(\hat{o} - E_{i|o}[\hat{o}])(\hat{o} - E_{i|o}[\hat{o}])^t] \quad (2.19)$$

Les éléments de la variance sont les éléments diagonaux de cette matrice :

$$V_{i|o}(\hat{o}) = \text{diagonale}(C_{i|o}(\hat{o})) \quad (2.20)$$

Écart quadratique moyen L'écart quadratique moyen (EQM) est défini comme la moyenne du carré de l'écart entre \hat{o} et la valeur vraie du paramètre :

$$EQM = E_{i|o}[(\hat{o} - o_{vraie})^2] \quad (2.21)$$

$$= V_{i|o}(\hat{o}) + (b_{i|o}(\hat{o}))^2 \quad (2.22)$$

2.3.2 Comportement asymptotique

L'influence du niveau de bruit ainsi que du nombre de données traitées sur la qualité des estimées objet et aberration va être étudiée à l'aide de simulations. Pour cela, on utilise 50 paires d'images (i_f, i_d) qui correspondent à la réalisation de 50 tirages de bruit. A l'aide de ces 50 images, on estime par MAPJ, 50 paires objet-aberrations que l'on utilise afin de déduire biais, variance et EQM. Le paragraphe qui suit détaille les conditions de simulations de ces images. Pour ces simulations, nous nous sommes placés dans le cadre de l'observation de la Terre depuis l'espace.

2.3.2.1 Conditions de simulation

La phase Les aberrations qui dégradent la qualité des images acquises par un télescope spatial proviennent des défauts propres de l'instrument. La phase aberrante simulée a été calculée à partir de la distribution de Zernike donnée table 2.1. Cette distribution correspond à une décroissance fréquentielle représentative des aberrations qui peuvent être associées à un tel système optique¹. L'amplitude de la phase ainsi simulée est de $\lambda/7$ rms.

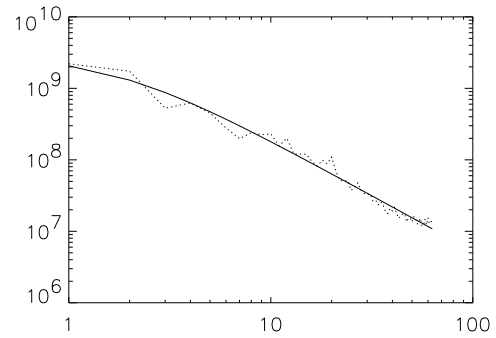
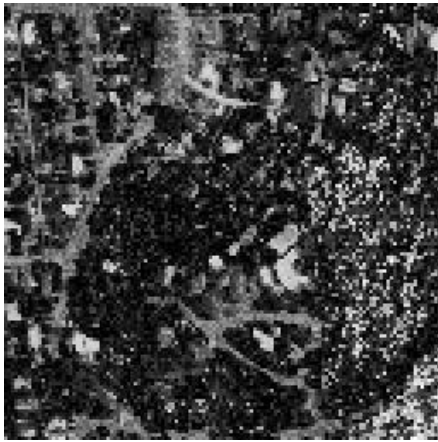
a_i	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
(radian)	-0,2	0,3	-0,45	0,4	0,3	-0,25	0,35	0,2	0,1
a_i	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}
(radian)	0,05	-0,05	0,05	0,02	0,01	-0,01	-0,02	0,01	0,01

TAB. 2.1 – Valeur des coefficients de Zernike utilisés lors des simulations.

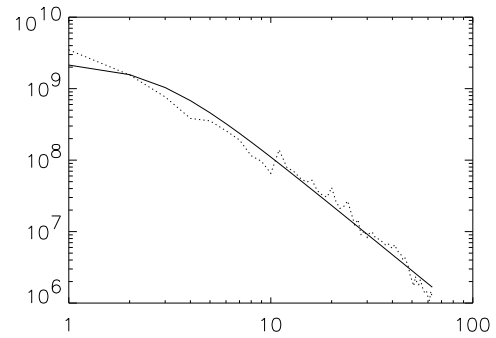
Les objets Comme le biais, la variance et l'EQM dépendent de la scène observée, trois objets ont été utilisés lors des simulations. Ces objets ont été choisis pour leur contenu spectral différent. Ils sont présentés figure 2.3.2.1 accompagnés de leur « DSP » (on a supposé que la DSP était simplement égale à la moyenne circulaire du module carré de la transformée de Fourier de l'objet).

Les images Les images simulées à partir de la phase et des objets présentés précédemment sont monochromatiques et échantillonnées en respectant le critère de Shannon. La distance de défocalisation de la deuxième image est choisie telle que l'amplitude crête-à-crête de la défocalisation soit de 2π radians. Les images sont obtenues par la convolution (effectuée dans le domaine de Fourier) de l'objet et de la RI optique du système aberrant (il n'a pas été modélisé de RI détecteur). Cette convolution est ensuite bruitée par un bruit blanc gaussien stationnaire. Le fait de réaliser la convolution dans le domaine de Fourier génère des images périodiques. Cette situation est bien entendu irréaliste mais permet de rendre les matrices H_f, H_d et R_o circulantes à blocs circulants, et donc d'utiliser l'expression de $\hat{o}(\mathbf{a})$, ceci à

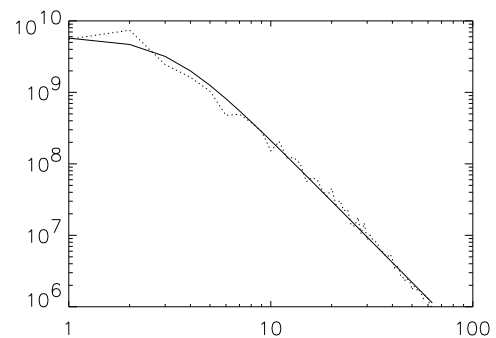
1. Cette même distribution avait déjà été utilisée dans (Meynadier, 1997)



- Scène vue aérienne haute -



- Scène entrepôt -



- Scène maison -

FIG. 2.2 – Objets et leur densité spectrale de puissance vraie (en pointillé) et ajustée sur le modèle paramétrique (trait plein) - voir le tableau 2.2 pour la valeur des paramètres ajustés.

seule fin de réduire le temps de traitement des 50 paires d'images (le problème spécifique du traitement d'images très étendues sera étudié au chapitre 4 page 86)².

2.3.2.2 Résultats

Comme nous l'avons montré au paragraphe 2.2.3.3, un problème majeur de l'estimation par MAPJ est le réglage de la valeur des hyperparamètres. Les résultats présentés ici ont été obtenus en utilisant les hyperparamètres θ_o ajustés sur l'objet vrai. Ils ont été obtenus en ajustant le modèle paramétrique de DSP sur la DSP de l'objet (obtenue par moyenne circulaire du module carré de la TF de l'objet). Ces hyperparamètres seront, dans toute la suite, appelés hyperparamètres « vrais ». L'objet moyen o_m est pris égal à zéro. La vraie valeur de l'écart type du bruit a également été utilisée. Le tableau 2.2 rassemble la valeur vraie des hyperparamètres associée à chaque objet (de taille 128×128 pixels). Afin de faire varier le nombre de données traitées, on utilise une taille d'image plus ou moins grande (de 32×32 , 64×64 et 128×128 pixels). Les résultats concernant la taille image 32×32 correspondent au moyennage des résultats obtenus à partir des 16 médaillons 32×32 contenus dans l'image 128×128 pixels, de même pour la taille image 64×64 les résultats présentés proviennent du moyennage des 4 médaillons 64×64 constituant l'image 128×128 pixels. La taille de l'objet utilisée variant, on a pris soin d'ajuster les hyperparamètres sur chaque taille d'objet observé afin de s'affranchir de toute variation éventuelle de contenu spectral entre les zones observées.

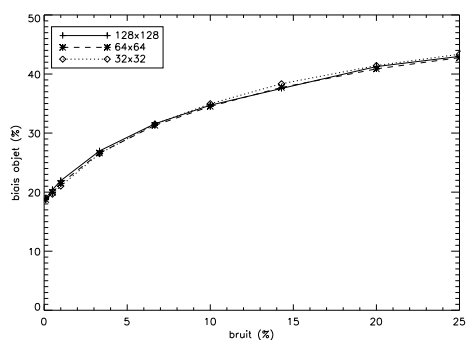
On reconstruit les coefficients de Zernike de a_4 à a_{21} (la vraie carte de phase ayant été simulée à partir de ces mêmes polynômes). En effet, le piston a_1 correspondant à l'ajout d'un terme constant sur la phase n'a pas d'influence sur la RI et les tilts a_2 et a_3 introduisant simplement un décalage dans l'image ne sont donc pas identifiables dans le cas d'objet étendu. D'autre part, dans le cas où l'on recherche aussi peu d'aberrations, la régularisation vis-à-vis de la phase induite par la troncature de la décomposition sur la base des Zernike suffit. On n'utilise donc pas le terme d'*a priori* J_a .

Les résultats pour les trois objets sont présentés sur les figures 2.3, 2.4 et 2.5. Pour l'objet nous avons tracé le biais, l'écart type et la racine carrée de l'EQM moyen par pixel (appelé également erreur), normalisé par le nombre total de photons par pixel dans l'image. Pour les aberrations, il s'agit du biais, écart type et racine carrée de l'EQM total (également dénommé erreur). Le niveau de bruit est exprimé en pourcentage, c'est le rapport entre l'écart type du bruit et le flux moyen par pixel. Les courbes montrent à la fois l'évolution de l'erreur en fonction du nombre de données traitées mais aussi l'évolution avec le niveau de bruit.

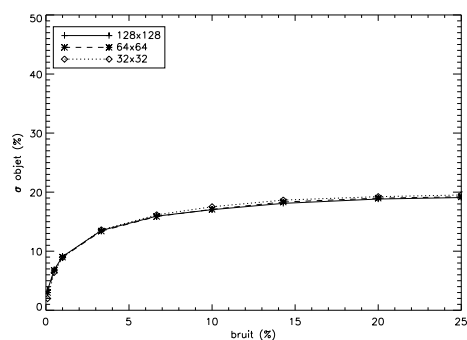
objet	σ^2	k	p	v_o
vue aérienne	1.e4	6,86e14	1,6	1,7
entrepôt	1.e4	2,33e14	2,3	2,7
maison	1.e4	1,64e14	2,9	3,1

TAB. 2.2 – Valeurs vraies des hyperparamètres utilisées lors des simulations pour les images 128×128 pixels.

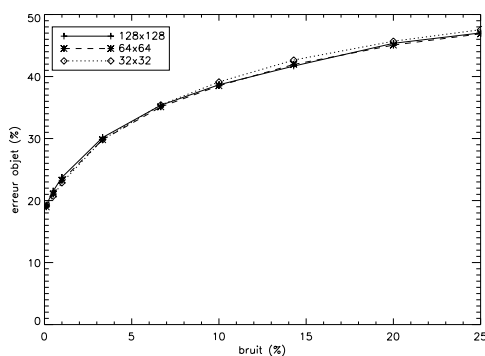
2. Nous avons vérifié sur l'objet vue aérienne que les comportements asymptotiques étaient identiques dans le cas d'images non périodiques.



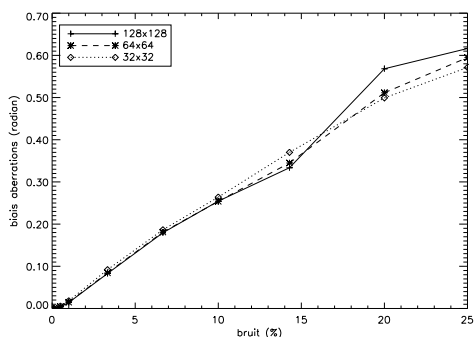
- Biais sur l'objet -



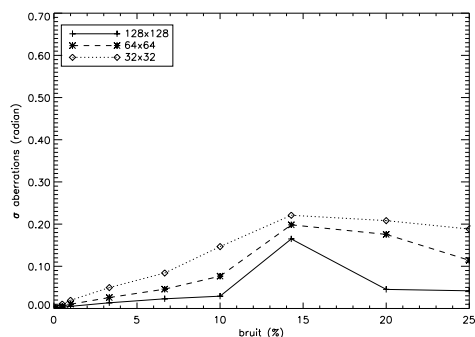
- Ecart type sur l'objet -



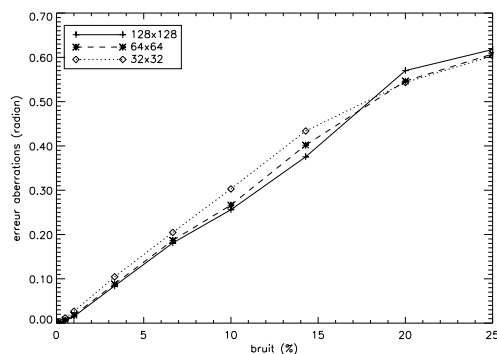
- Erreur sur l'objet -



- Biais sur la phase -

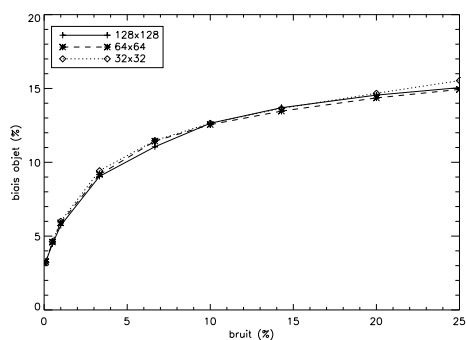


- Écart type sur la phase -

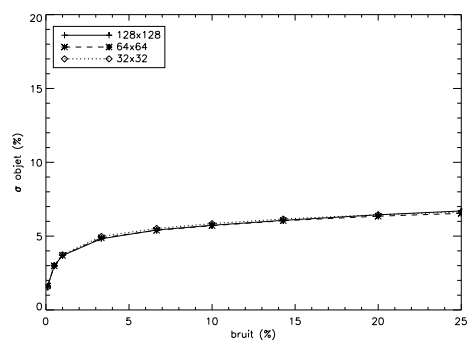


- Erreur sur la phase -

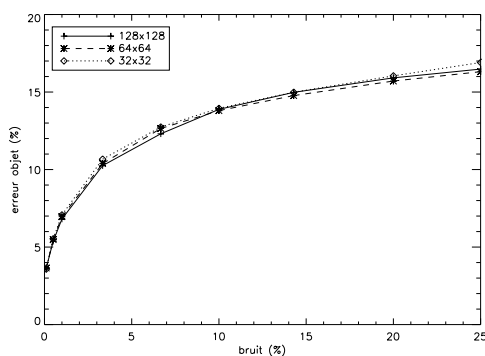
FIG. 2.3 – Propriétés asymptotiques - Scène vue aérienne.



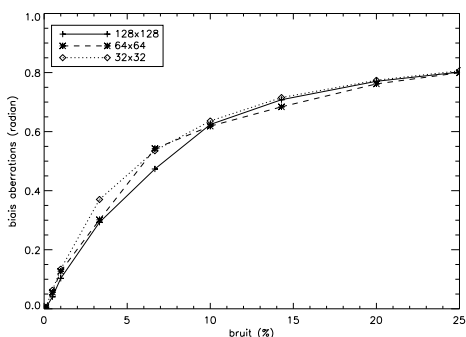
- Biais sur l'objet -



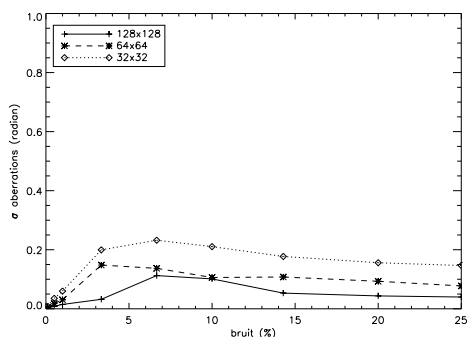
- Écart type sur l'objet -



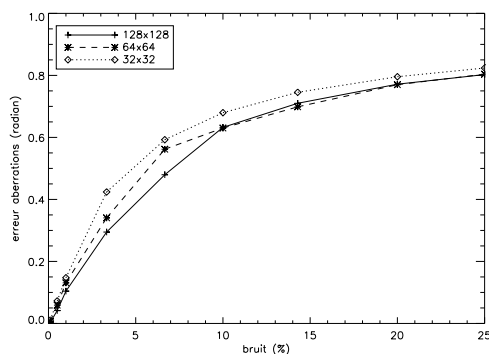
- Erreur sur l'objet -



- Biais sur la phase -



- Écart type sur la phase -



- Erreur sur la phase -

FIG. 2.4 – Propriétés asymptotiques - Scène entrepôt.

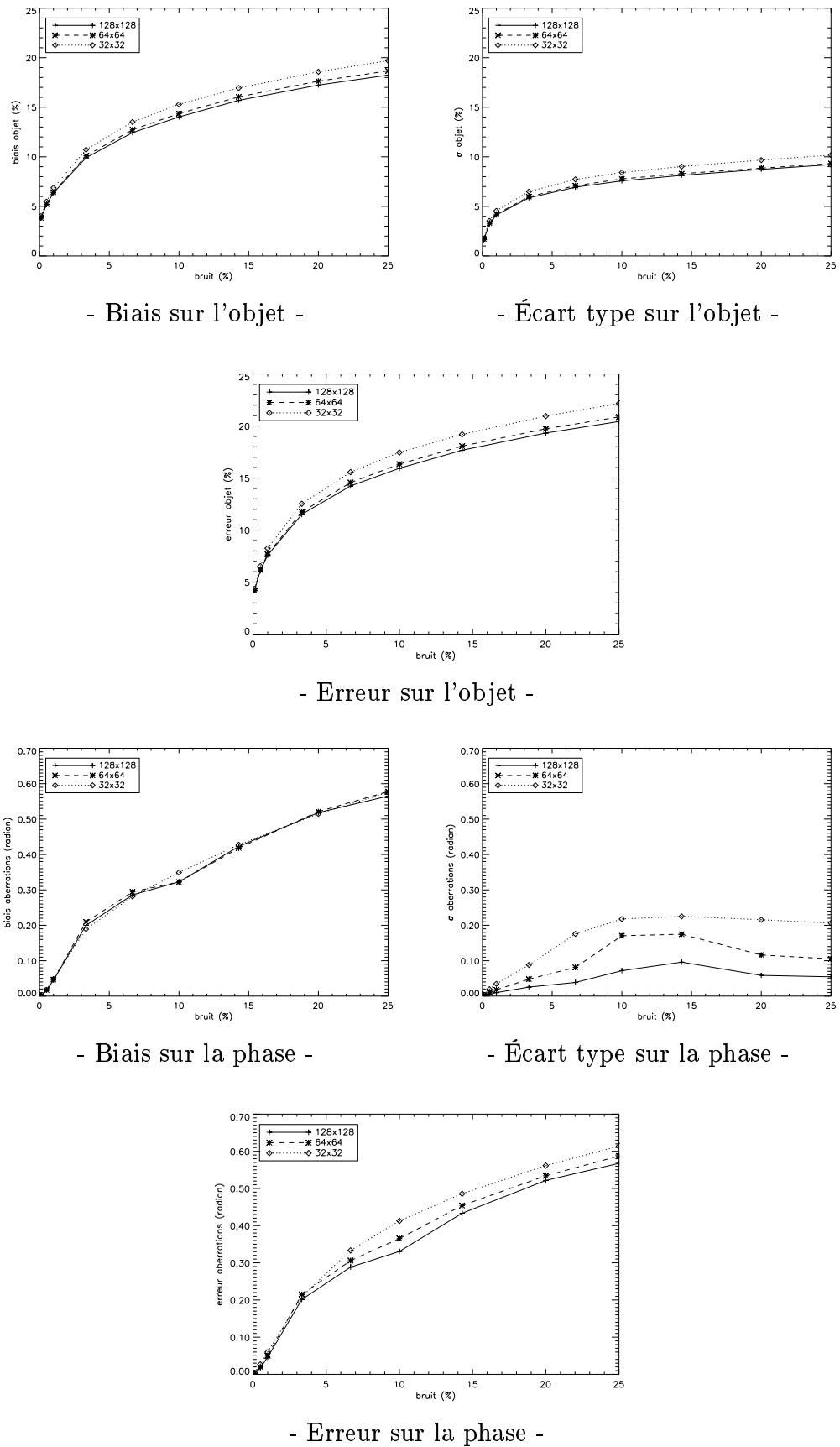


FIG. 2.5 – Propriétés asymptotiques - Scène Maison.

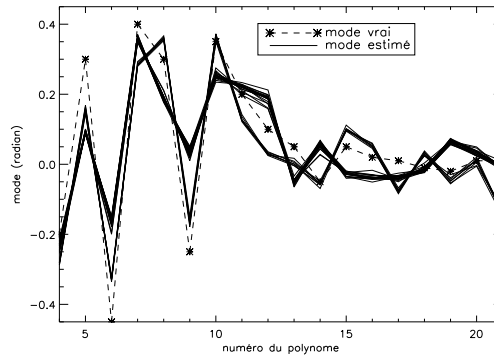


FIG. 2.6 – Les 50 estimations des aberrations obtenues pour l’image 128×128 pixels bruit de 14% et l’objet « vue aérienne » ; en trait plein les aberrations estimées, en tireté les aberrations vraies.

objet Pour les trois objets, la comparaison des courbes de biais et écart type à bruit donné montre que ces deux grandeurs suivent un comportement identique : il n’y a pas d’amélioration de la qualité de l’objet estimé lorsque le nombre de données augmente. Dans le cas de l’objet « maison », il y a même une légère dégradation de l’estimation lorsque le nombre de données augmente, pour les objets « vue aérienne » et « entrepôt », la qualité de l’estimation demeure quasiment identique. Ce comportement asymptotique assez peu satisfaisant, peut s’expliquer par le fait que lorsque l’on augmente le nombre de données traitées (c’est-à-dire la taille de l’image) le nombre d’inconnues « objet » à reconstruire augmente d’autant, ainsi le rapport données/inconnues reste quasiment constant.

Le comportement en fonction du niveau de bruit est le même pour les trois objets : il y a une dégradation rapide lorsque le bruit est faible puis une relative stabilité dans le domaine des faibles flux.

La séparation de l’erreur totale (donnée par l’EQM) en ces deux composantes que sont le biais et la variance permet de montrer que l’erreur est dominée par le terme de biais.

phase Dans le cas des aberrations, l’évolution du biais et de l’écart type avec le nombre de données diffère. Le biais reste stable ou augmente légèrement lorsque la taille de l’image traitée augmente alors que l’écart type lui diminue pour les trois objets. On observe donc ici, tout comme dans le cas de l’objet, un biais asymptotique. Par contre, dans le cas de la phase, l’écart type diminue. En effet, le nombre d’inconnues « phase », très faible par rapport au nombre de données, reste constant lorsque la taille image augmente. Il y a donc une amélioration du rapport données/inconnues qui entraîne une diminution de la dispersion des résultats et donc une amélioration asymptotique de l’écart type.

Quel que soit l’objet, on observe une augmentation du biais et de l’écart type lorsque le niveau de bruit croît. La courbe de variance liée à l’objet « vue aérienne » présente un comportement irrégulier : un point semble « aberrant » ; image 128×128 pixels et un bruit de 14%. Nous avons tracé l’ensemble des résultats obtenus à partir des cinquante paires d’images de l’objet « vue aérienne » 128×128 pixels et un bruit de 14%, figure 2.6. On constate que la minimisation du critère joint a conduit à deux faisceaux de solutions $\{a_i\}_1$ et $\{a_i\}_2$, c’est pourquoi la variance (qui quantifie la dispersion des valeurs estimées) est très forte pour ces points là. La minimisation a conduit à deux solutions qui sont deux minima du critère, ce qui permet d’affirmer que le critère joint n’est pas convexe.

Pour la phase tout comme pour l'objet, c'est le terme de biais qui est prépondérant dans l'erreur totale. On remarque, de plus, que la qualité de l'estimée des aberrations dépend fortement de l'objet observé. L'EQM sur l'estimée des aberrations obtenue avec l'objet « entrepôt » est près de deux fois moins bon que celui obtenu par l'objet « maison » et par l'objet « vue aérienne », ce dernier conduisant à la meilleure estimée des aberrations.

2.3.2.3 Conclusion

Nous venons de montrer que, quel que soit l'objet, l'erreur d'estimation sur les aberrations, tout comme celle sur l'objet ne diminue pas asymptotiquement : si on augmente le nombre de données traitées, la qualité d'estimation des grandeurs recherchées ne s'améliore pas. On résume ce comportement en disant que cet estimateur joint est asymptotiquement biaisé. Ces mauvaises propriétés asymptotiques d'estimateurs joints sont avérées en statistiques (Little et Rubin, 1983).

2.4 Influence des hyperparamètres

Après avoir étudié les propriétés asymptotiques de l'estimateur joint, on revient au problème des hyperparamètres. Comme il a été montré au paragraphe 2.2.3, les hyperparamètres ne peuvent pas être estimés conjointement avec l'objet et la phase. Il faut régler leur valeur au mieux empiriquement (méthode dite supervisée). Dans ce paragraphe, on s'intéresse à l'influence de la valeur de ces hyperparamètres sur la qualité de restauration de l'objet et de la phase.

2.4.1 Variation des hyperparamètres autour de leur valeur vraie

On s'intéresse à la qualité d'estimation de l'objet et des aberrations lorsque les hyperparamètres varient autour de leur valeur ajustée à l'aide de l'objet vrai. Cette étude est effectuée à partir de simulations prises dans les mêmes conditions que dans le paragraphe précédent (notamment les trois mêmes objets sont utilisés) et un médaillon image de 32×32 pixels. Bien que les résultats soient présentés pour une unique image de 32×32 pixels, il a été vérifié sur plusieurs médaillons que les résultats détaillés ici dans ce cas particulier sont représentatifs du comportement général des différents médaillons constituant les images 128×128 pixels. Nous avons étudié l'influence de l'hyperparamètre global k , c'est-à-dire celui qui règle l'équilibre entre les termes d'attache aux données et le terme d'*a priori* sur l'objet. Les courbes sont tracées pour deux bruits différents : fort bruit (14%) figure 2.7 et faible bruit (1%) figure 2.8 (en abscisse, la valeur 1 correspond à la valeur vraie de l'hyperparamètre k).

2.4.1.1 bruit 14%

Les courbes concernant l'erreur d'estimation objet présentent un minimum centré sur la valeur vraie de l'hyperparamètre. Lorsque celui-ci s'éloigne de cette valeur optimale, l'erreur croît rapidement. Cette sensibilité importante de la valeur des hyperparamètres sur la qualité de l'estimée objet montre que ce réglage est un point critique de l'estimation jointe puisque, comme nous l'avons montré, il n'est pas possible d'estimer simplement ces hyperparamètres de façon non supervisée.

Les courbes concernant les aberrations présentent un minimum moins marqué que pour l'objet. Ce minimum est obtenu pour une valeur de l'hyperparamètre qui ne coïncide pas,

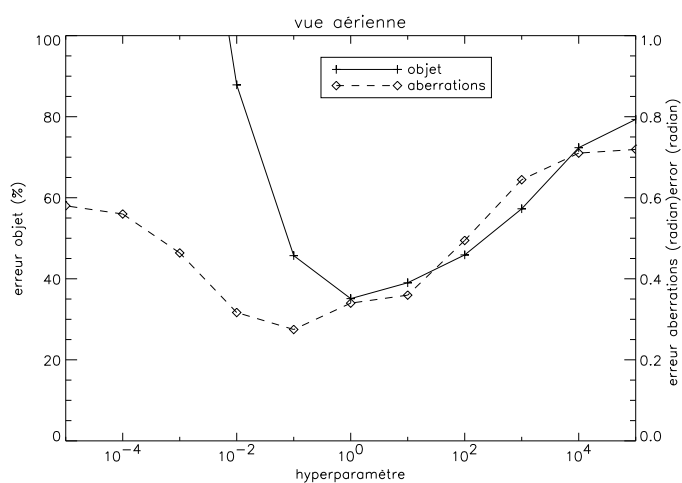
dans le cas des objets « vue aérienne » et « entrepôt », avec la valeur de l'hyperparamètre qui minimise l'erreur d'estimation sur l'objet. Cela signifie que, pour ces objets, même dans un cas d'école où les hyperparamètres objets optimaux seraient connus, il n'est pas possible d'estimer de façon optimale objet et aberrations conjointement : en se plaçant à la valeur des hyperparamètres optimale pour l'estimation de l'objet, l'erreur sur l'estimée des aberrations ne serait pas minimale et bien sûr réciproquement si l'on se place aux hyperparamètres optimaux pour les aberrations.

2.4.1.2 bruit 1%

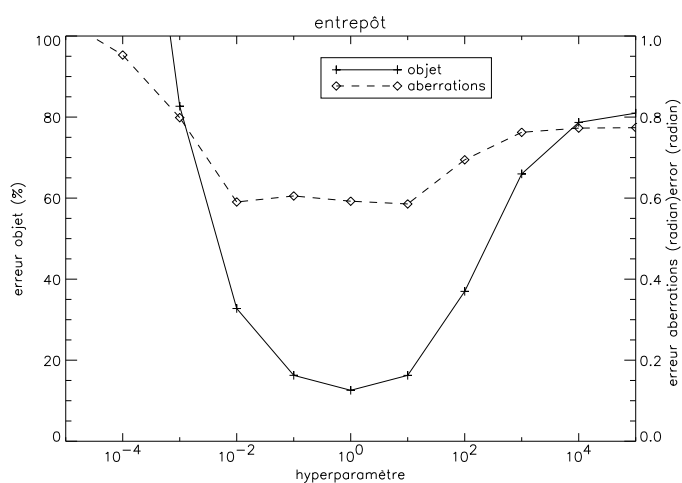
Le comportement des courbes portant sur l'objet reste identique au cas bruit de 14% mais pas celui des courbes concernant les aberrations : lorsque le niveau de bruit diminue, un plateau de plus en plus important de minima quasiment équivalents apparaît (ce plateau n'apparaît pas de part et d'autre de la valeur optimale mais uniquement dans le sens d'une sous-régularisation objet). Finalement, pour de faibles bruits (ici 1%), quelle que soit la valeur des hyperparamètres inférieure à la valeur optimale, l'erreur sur l'estimée des aberrations est minimale (seulement pour l'objet « entrepôt », on obtient une faible dégradation lorsque l'on s'éloigne de la valeur vraie) alors que l'estimée objet sous-régularisé est de très mauvaise qualité. Le « quel que soit » va jusqu'à une régularisation quasi-nulle sur l'objet (voir équation 2.11) c'est-à-dire pour un rapport σ^2/S_o qui tend vers zéro ! On prend soin, toutefois, de ne pas mettre ce rapport à zéro exactement mais égal à une constante ϵ . Cette constante n'a rien à voir avec un quelconque *a priori* sur l'objet, sa valeur, très faible (ici 10^{-16}), est fixée en fonction de la précision du calculateur afin d'éviter d'éventuelles problèmes numériques lors de l'évaluation de \hat{o} . Un exemple de restauration conjointe à régularisation « quasi-nulle » sur l'objet est présentée figure 2.9. Les aberrations sont parfaitement estimées (l'erreur est de 0,009 radian) alors que l'objet reconstruit est quant à lui totalement bruité.

2.4.2 Comportement asymptotique à régularisation quasi-nulle

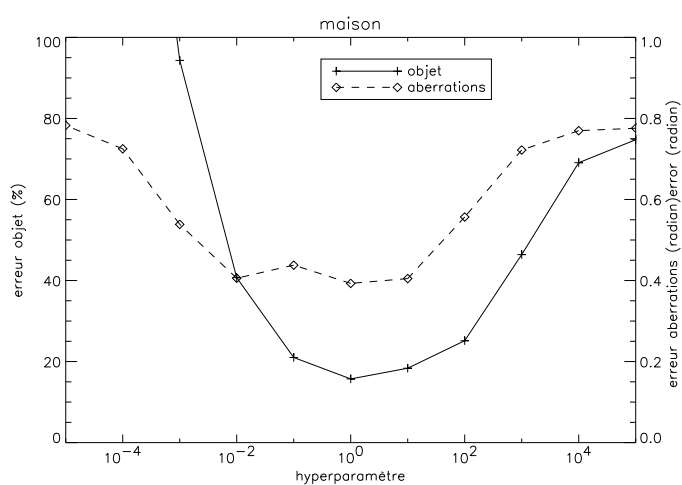
La bonne estimation des aberrations à régularisation quasi-nulle sur l'objet nous a conduit à étudier l'évolution de l'erreur sur l'estimation des aberrations à régularisation quasi-nulle sur l'objet lorsque la taille de l'image traitée augmente. Les résultats sont rassemblés sur les figures 2.10, 2.11 et 2.12. Pour cet estimateur dérégulé, le biais, la variance (et donc l'EQM) diminuent lorsque le nombre de données traitées augmente. On observe ici la prédominance de la composante variance dans le terme d'erreur. Les courbes concernant l'objet n'ont pas été tracées, celui-ci étant, quel que soit le bruit et la taille image utilisée, très dégradé. En résumé, pour d'assez forts flux, quel que soit l'objet, l'estimateur joint sans régularisation sur l'objet présente un bon comportement asymptotique vis-à-vis des aberrations lorsque le nombre de données augmente. L'évolution du rapport données/inconnues est le même que dans le cas d'une régularisation sur l'objet mais ici tout se passe comme si l'objet n'était pas estimé ! Une confirmation de ce comportement surprenant peut être trouvée dans la littérature de la diversité de phase. En effet, différents articles (Kendrick et al., 1994; Lee et al., 1997b; Löfdahl et Duncan, 1998; Meynadier et al., 1999; Thelen et al., 1999; Seldin et Paxman, 2000) qui portent sur l'analyse de surface d'onde par diversité de phase effectuent avec succès, dans le cas de bruits faibles, l'estimation des aberrations sans terme de régularisation sur l'objet.



- objet vue aérienne -

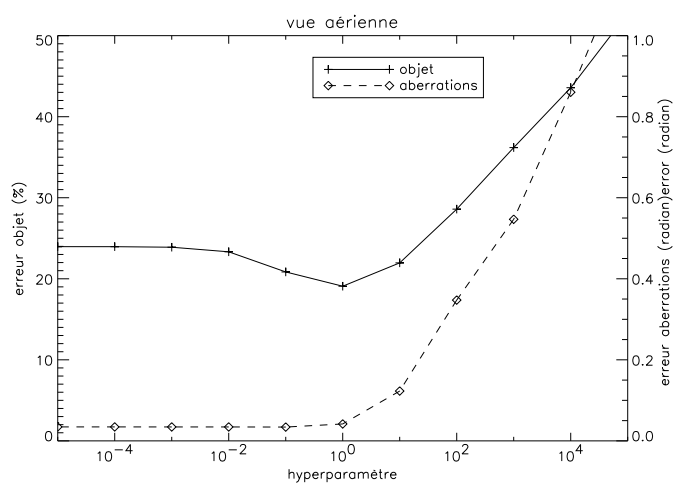


- objet entrepôt -

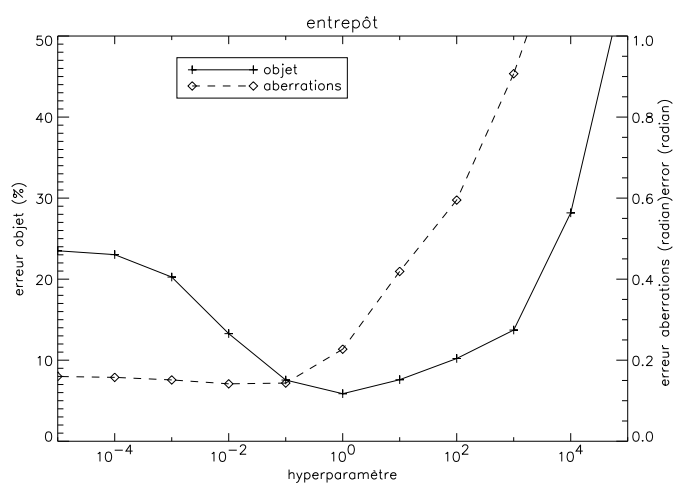


- objet maison -

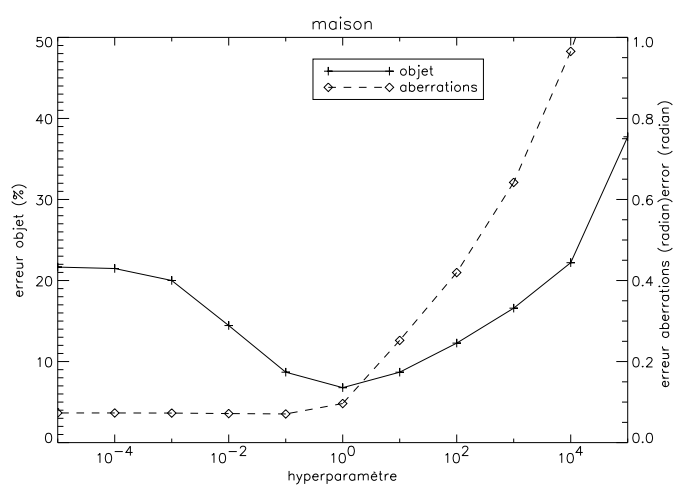
FIG. 2.7 – Influence de la valeur de l'hyperparamètre k sur l'estimation des aberrations et de l'objet image 32×32 pixels et un bruit de 14%, pour les trois objets.



- objet vue aérienne -



- objet entrepôt -



- objet maison -

FIG. 2.8 – Influence de la valeur de l'hyperparamètre k sur l'estimation des aberrations et de l'objet image 32×32 pixels et un bruit de 1%, pour les trois objets.

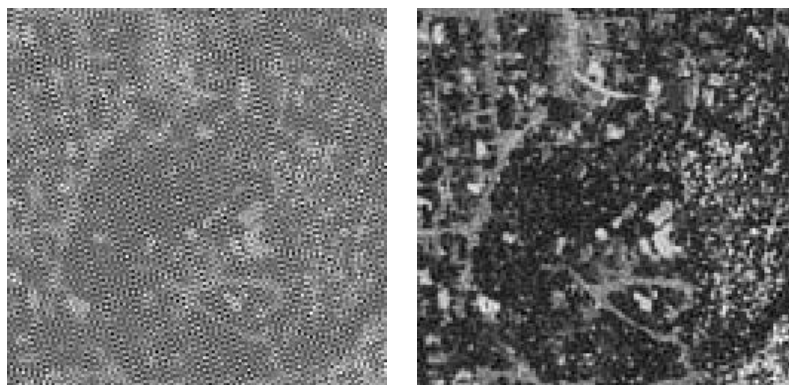


FIG. 2.9 – *Objet vue aérienne reconstruit, à gauche avec une régularisation « quasi-nulle » sur l'objet et à droite (pour comparaison) avec une régularisation à hyperparamètres vrais sur l'objet, bruit 1%.*

2.5 Conclusion

Dans ce chapitre, nous avons proposé une interprétation probabiliste au sens du MAPJ de l'estimateur joint puis nous avons mené une étude critique et originale de ces propriétés.

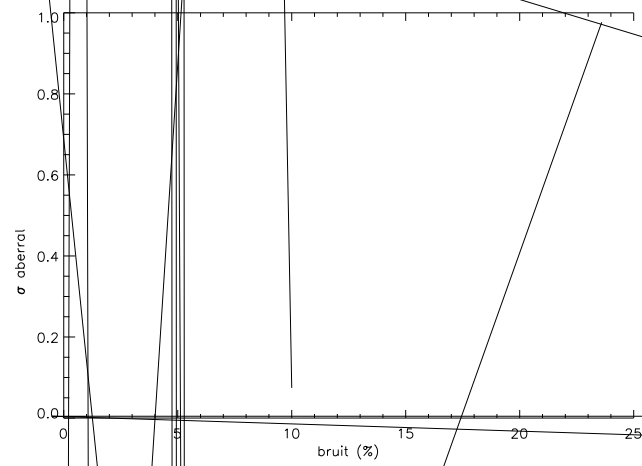
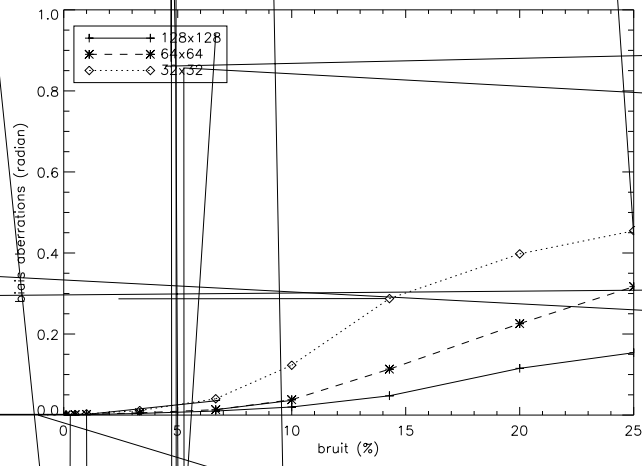
Nous avons montré que cet estimateur ne possède pas de bonnes propriétés asymptotiques : l'utilisation d'un plus grand nombre de données ne permet pas de réduire l'erreur d'estimation des inconnues que sont l'objet et les aberrations. L'introduction de termes d'*a priori* nécessite le choix de la valeur des hyperparamètres. Nous avons montré que leur réglage non supervisé ne peut pas être effectué conjointement avec l'objet et les aberrations pour des raisons de dégénérescence du critère.

L'influence de la valeur des hyperparamètres sur la qualité de restauration de l'objet et des aberrations a été étudiée. Nous avons mis en évidence qu'il n'y a pas adéquation entre les hyperparamètres optimaux vis-à-vis de l'objet (i.e ceux donnant l'objet estimé le plus proche de l'objet vrai) et les hyperparamètres optimaux vis-à-vis des aberrations. Cela a permis également de montrer qu'à fort flux et en enlevant le terme de régularisation lié à l'objet (i.e. en dérégulant l'estimateur), bien que l'objet estimé soit très dégradé, les aberrations peuvent être correctement reconstruites.

Dans le chapitre suivant, nous proposons une application expérimentale de la diversité de phase qui utilise le bon comportement à fort flux de cet estimateur dérégulé (qui correspond alors à une estimation par simples moindres carrés « quasi non-régularisés »³) pour estimer les aberrations d'une caméra d'imagerie.

Afin de s'affranchir des mauvaises propriétés de l'estimateur joint et de résoudre le problème de l'estimation non-supervisée des hyperparamètres qui est impossible dans la méthode jointe, nous avons développé une nouvelle méthode d'estimation qui est présentée dans le chapitre 4.

3. le terme « quasi non-régularisé » indique la présence de la constante ϵ au dénominateur de l'équation 2.11, voir paragraphe 2.4.1.2



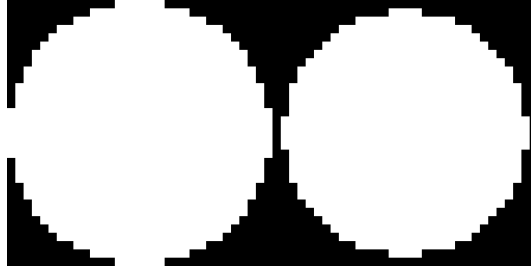


FIG. 3.6 – Influence de l'échantillonnage : masques utilisés pour simuler la pupille de CO-NICA, à droite $f_c/f_p = 4,06$ et à gauche $f_c/f_p = 4,38$.

positionnée le masque, on enregistre une deuxième image avec le trou situé à 4 mm. A partir de cette paire d'images, on effectue plusieurs estimations des aberrations en faisant varier le pas pixel en entrée du programme d'estimation. De façon pratique, ces variations sont effectuées en adaptant la taille de la pupille simulée à la taille des images. Ici les variations du pas pixel entraînent une variation de l'échantillonnage de 4,06 à 4,38, le diamètre de la pupille varie alors de 31,51 à 29,22 pixels (les deux masques correspondant à ces valeurs extrêmes sont représentés figure 3.6).

Les courbes 3.7 ont été obtenues en fixant la distance de défocalisation à sa valeur théorique (4 mm) et en faisant varier le pas pixel autour de sa valeur théorique (13,25 marcsec/pixel). On constate que ces variations ont peu d'influence sur l'estimation, excepté pour a_4 et plus faiblement pour a_{11} et a_{14} . L'influence de l'échantillonnage sur ces aberrations peut s'expliquer par le fait que la modification du grandissement induit une modification de la focale du système ce qui affecte l'estimation du terme de défocalisation a_4 . La variation du facteur d'échantillonnage entraîne également une variation du support de la RI qui fait apparaître des termes d'aberrations d'ordres élevés : a_{11} et a_{14} . Les variations observées restent très faibles dans le cas de a_{11} et a_{14} , en simulation on trouve un EQM maximal (rms) de l'ordre de $\lambda/300$, plus importantes dans le cas du terme de défocalisation a_4 , avec une erreur maximale de $\lambda/100$.

3.3.5 Incertitude sur la distance de défocalisation

Les distances longitudinales entre les trous du masque ne sont pas connues exactement. L'incertitude sur la distance exacte est au plus de 200 μm . Or la diversité de phase suppose que la distance de défocalisation entre les deux plans de mesure est parfaitement déterminée. On s'intéresse donc à l'influence sur l'estimation des aberrations d'une erreur sur cette distance. Comme on l'a fait précédemment pour le pas pixel, on montre l'influence d'une erreur sur la distance de défocalisation sur l'estimation des aberrations. Les courbes 3.8 ont été obtenues dans les mêmes conditions de simulation et expérimentales que dans le paragraphe précédent, en fixant la valeur du pas pixel à sa valeur théorique (13,25 marcsec/pixel) et en faisant varier la distance de défocalisation autour de sa valeur théorique (4 mm). Les deux courbes, qui sont très semblables, montrent que l'estimation de la défocalisation et dans une bien moindre mesure celle de l'aberration sphérique sont affectées par l'incertitude sur la distance de défocalisation. Les différents résultats obtenus en simulation sont rassemblés dans le tableau 3.2: il a été reporté l'erreur entre la vraie distance de défocalisation et celle supposée (δd vraie), l'écart exprimé en nm entre la valeur estimée de \hat{a}_4 et sa valeur vraie a_4^{vraie} et ce même écart exprimé en distance de défocalisation dans le

Erreur sur le pa

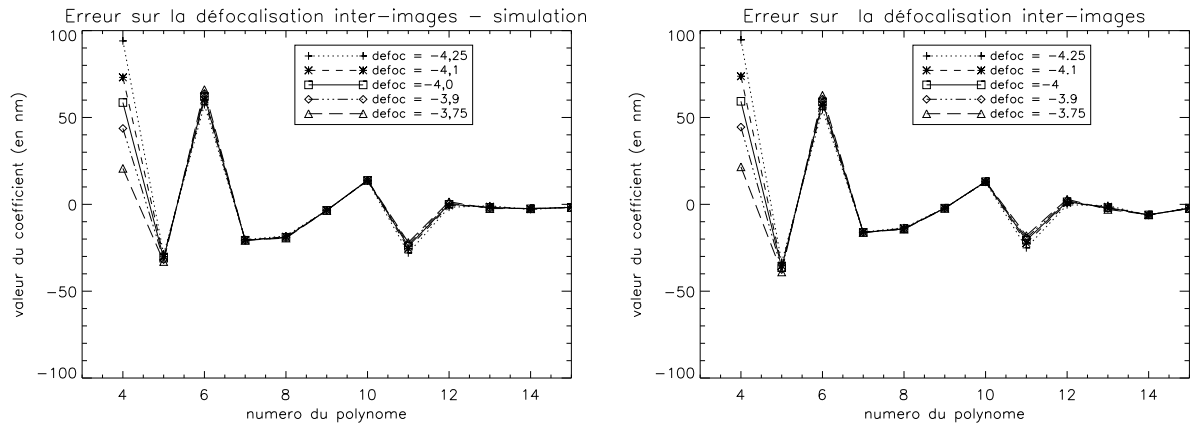


FIG. 3.8 – Influence d'une erreur sur la distance de défocalisation - Caméra C50S, filtre $2,166 \mu\text{m}$, trous 0 et 4mm . La courbe de gauche est obtenue sur données simulées (objet quasi-ponctuel, RSB de 200, image 128×128 pixels, défocalisation vraie de 2π), la courbe de droite sur données expérimentales.

3.3.6 Influence de la translation dans le plan focal d'entrée

Il n'a pas été possible d'utiliser la translation du détecteur pour introduire la défocalisation nécessaire à la deuxième image (voir explication donnée au paragraphe 3.2). Dans le cas de CONICA, la diversité est générée par la translation de « l'objet » dans le plan focal d'entrée. Cette translation n'introduit pas, contrairement à la translation du détecteur, une défocalisation pure mais également (au premier ordre) de l'aberration sphérique. Afin de quantifier l'importance de l'écart à une défocalisation pure, des simulations d'une translation dans le plan focal d'entrée ont été effectuées par l'équipe du MPIA à l'aide d'un logiciel de calcul optique. Ils ont montré que dans le cas de la combinaison optique de CONICA et des distances de translation introduites, le terme d'aberration sphérique introduit était négligeable. Dans notre cas, la translation de l'objet peut donc être considérée comme équivalente à la translation du détecteur et donc la diversité introduite dans la deuxième image uniquement composée d'une défocalisation.

3.3.7 Forme des trous

Une hypothèse de la diversité de phase est que l'objet observé par les deux plans de mesure est exactement le même. Dans notre montage expérimental, il n'a pas été possible d'utiliser le même trou pour acquérir les deux images, on en utilise en fait deux trous physiquement distincts. Ces trous ont été percés mécaniquement et l'on ne peut pas être sûr qu'ils soient strictement identiques. Théoriquement, ce sont des trous circulaires de $10 \mu\text{m}$ de diamètre mais un trou peut avoir été mal percé et ne pas être circulaire mais plus oblong que les autres. Afin de s'assurer qu'il n'y a pas de trou différent des autres, deux masques portant 4 trous circulaires ont été fabriqués et les résultats obtenus à partir de chacun d'eux comparés. Cette comparaison n'a pas montré d'écart significatif (écart inférieur à $\lambda/100$) entre les aberrations reconstruites. Ce résultat permet de réduire l'incertitude quant à la présence éventuelle d'un trou différent des autres.

3.3.8 Décalage latéral entre les deux trous

De même que l'objet, les aberrations dégradant chacune des images sont supposées identiques. Or la rotation de la roue portant le masque des trous ne permet pas de positionner tous les trous au même endroit du champ. Le déplacement des trous dans le champ ne peut se faire que suivant la direction horizontale (direction x), le décalage porte sur la position verticale (direction y). Si on prend comme position de référence la position du trou 0 mm, on a pour les trous 1 et 2 mm un décalage de 100 pixels en y et pas de décalage pour le trou à 4 mm. Certaines aberrations comme l'astigmatisme et la coma varient dans le champ donc dans le cas d'une paire d'images obtenue avec le trou 0 et 2 mm, les aberrations « vues » par les deux plans ne sont pas exactement les mêmes. Ce décalage ne devrait pas poser de problème lors de l'estimation si les aberrations sont suffisamment lentement variables dans le champ. Nous ne possédons pas de connaissance de cette évolution (mais on peut légitimement espérer qu'elle est effectivement assez faible), on a donc privilégié l'utilisation de la paire de trous 0 mm et 4 mm lorsque c'est possible (distance de défocalisation comprise entre π et 2π), afin d'être sûr de ne pas être gêné par cette éventuelle différence. L'annexe C propose un tableau récapitulatif présentant pour les différents filtres utilisés, leur longueur d'onde centrale, la défocalisation introduite par chacun des trous ainsi que l'échantillonnage théorique.

3.3.9 Synthèse de l'étude des différentes sources de perturbations

Avant de s'intéresser réellement aux différentes sources de perturbations, nous avons montré par simulations (dans des conditions proches des conditions expérimentales), que la plage de défocalisation optimale à appliquer entre les images était comprise entre π et 2π radians crête-à-crête. Cette étude nous permettra, dans la suite, de choisir la ou les paire(s) de trous à utiliser suivant la configuration de CONICA étudiée.

Puis il a été rappelé la nécessité de parfaitement connaître la forme de la pupille, celle-ci n'étant pas estimée par la méthode employée ici. Un fond résiduel présent sur les images a été mise en évidence, il doit être soustrait avant de traiter les données.

Différentes sources d'incertitudes liées à la mise en oeuvre de la diversité de phase à CONICA ont été étudiées afin de déterminer leur influence sur l'estimation des aberrations. On a montré que l'incertitude sur le facteur d'échantillonnage des images ainsi celle portant sur la distance exacte entre les trous dégrade la qualité d'estimation de la défocalisation, les autres aberrations étant quasiment insensibles. Il a été vérifié par calcul optique que la translation de l'objet dans le plan d'entrée induit bien uniquement de la défocalisation (le terme d'aberration sphérique également introduit étant négligeable). Les objets observés par les deux plans de mesure n'étant physiquement pas les mêmes, on a vérifié à l'aide d'un deuxième masque qu'il n'y avait pas sur le masque utilisé un trou mal percé et donc différent des autres. Tous les trous ne pouvant pas être placés à la même position du champ, suivant les paires de trous utilisées, les aberrations liées à chaque plan de mesure peuvent être différentes (pour les aberrations variant dans le champ). On privilégiera donc l'utilisation des paires pouvant être positionnées au même endroit (0-4 par exemple).

3.4 Premiers résultats et comparaisons

Après avoir passé en revue les différentes sources d'erreurs induites par le montage expérimental utilisé pour acquérir les images, on présente à présent les résultats pour la caméra

C50S et pour les différents filtres de la bande JH et K. Tout d'abord, la comparaison des résultats obtenus pour un même filtre et différentes paires de trous est commentée puis on présente les résultats obtenus avec tous les filtres des deux bandes. En bande K, quelques exemples de RI reconstruites à partir des aberrations estimées par diversité de phase sont présentés et comparés aux images plan focal correspondantes. Enfin, on compare le rapport de Strehl calculé par diversité de phase avec celui calculé à partir de l'image plan focal.

3.4.1 Comparaison des résultats obtenus à partir de différentes paires de trous

On estime à partir des différentes combinaisons de paires de trous, les aberrations associées à différentes combinaisons de la caméra C50S et de filtres étroits de la bande JH ainsi que de la bande K. Le pas pixel est fixé à sa valeur théorique (13,25 marcsec/pixel) et il en est de même pour la distance de défocalisation entre les images. Les résultats sont présentés figure 3.9 pour deux filtres de la bande JH (HeI centré sur 1,083 μm et FeII centré sur 1,644 μm) et figure 3.10 pour ceux de la bande K (centré sur 2,03 μm et 2,39 μm). Ces courbes présentent, pour chaque filtre, la valeur moyenne des aberrations estimées à partir des différentes paires de trous ainsi que la valeur maximale et minimale estimées (représentées sous forme de barres d'erreur). Pour les estimations où l'image dite « plan focal » n'a pas été obtenue avec le trou 0 mm, la défocalisation théorique du trou utilisé a été soustraite à la valeur de défocalisation estimée afin de pouvoir être comparée aux autres. Pour les filtres de la bande J, les moyennes ont été effectuées à partir de toutes les combinaisons possibles de trous (c'est à dire six paires). Pour les filtres de la bande K, les paires (0mm-1mm) et (1mm-2mm) n'ont pas été utilisées du fait de la trop faible défocalisation introduite (de l'ordre de $\lambda/4$). La courbe 3.4 a en effet montré qu'avec une telle valeur de défocalisation différentielle entre les deux images, l'erreur sur l'estimée des aberrations était près de deux fois plus importante que celle obtenue avec la défocalisation optimale. Ceci a été confirmé par les résultats expérimentaux : les estimations obtenues avec ces paires de trous étaient effectivement fortement différentes de celles obtenues avec les autres paires. C'est pourquoi les résultats obtenus à partir de ces différentes associations de trous n'ont pas été retenus.

C'est l'estimation de la défocalisation qui présentent la plus grande variabilité. Ce n'est pas surprenant compte tenu des incertitudes dont nous avons parlé précédemment (voir paragraphes 3.3.4 et 3.3.5). De plus la correction qui a été effectuée afin de supprimer la défocalisation introduite par la position du trou utilisé pour acquérir l'image « plan focal » est approximative là encore à cause de l'incertitude sur la position longitudinale exacte des différents trous. Pour les autres aberrations, il y a une très bonne adéquation entre les résultats obtenus avec les différentes paires de trous : les fluctuations les plus importantes constatées sur l'estimation d'un coefficient sont inférieures à $\lambda/40$. Compte tenu des différentes sources d'incertitude du système (discutées au paragraphe précédent), ces résultats sont très satisfaisants.

3.4.2 Présentation des résultats des différents filtres

On compare, à présent, les résultats obtenus pour tous les filtres étroits de la bande JH et pour tous ceux de la bande K. La figure 3.11 présente les valeurs moyennes des aberrations des différentes associations caméra C50S + filtres avec les fluctuations à $+/-\sigma$ (écart-type) associées (fluctuations dues aux différentes valeurs estimées à partir des différentes paires de trous). Les aberrations estimées correspondent à celles de la caméra elle-même (ici la

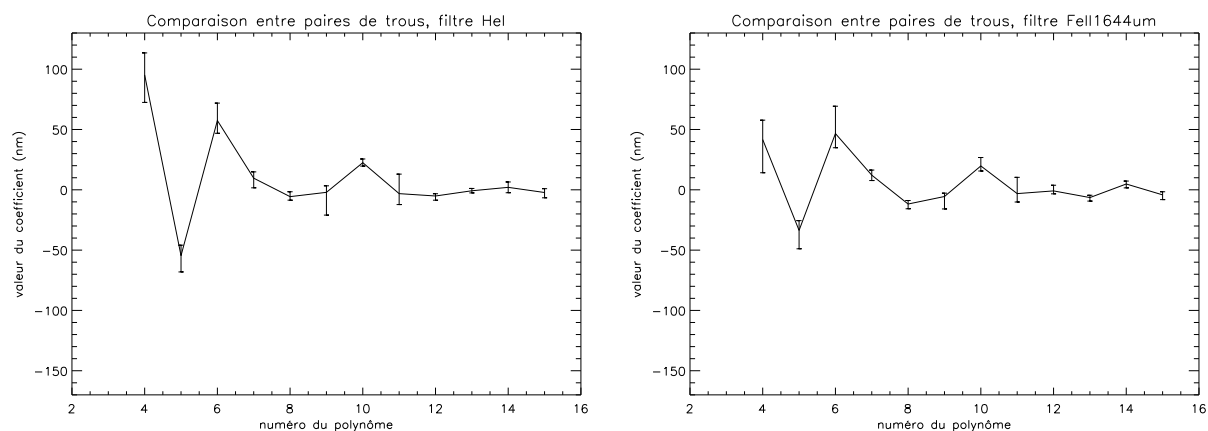


FIG. 3.9 – Comparaison des résultats obtenus avec différentes paires de trous - pour deux filtres de la bande JH, caméra C50S, à gauche filtre centré sur $1,083 \mu\text{m}$ et à droite pour un filtre centré sur $1,644 \mu\text{m}$. La valeur moyenne des estimées est tracée et les barres d'erreur correspondent à la valeur minimum et maximum estimée.

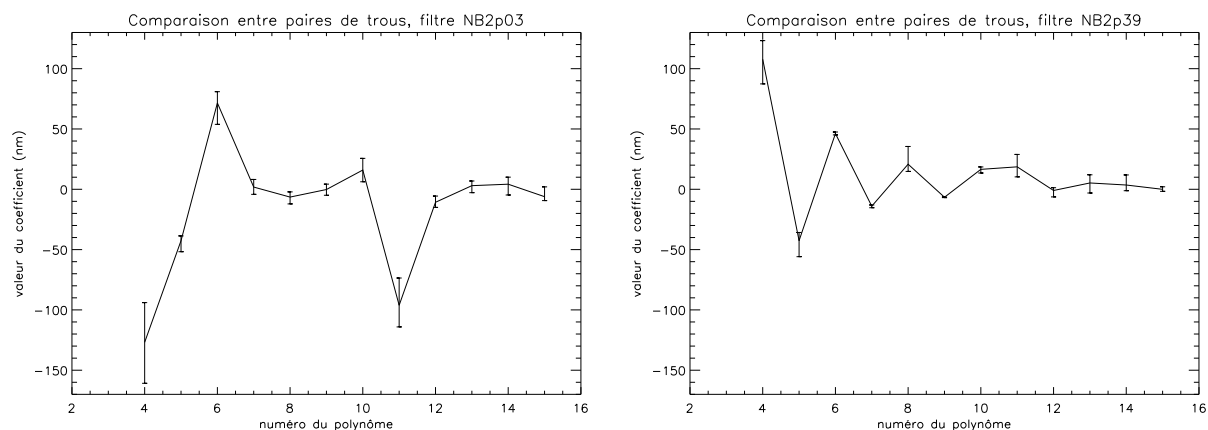


FIG. 3.10 – Comparaison des résultats obtenus avec différentes paires de trous - pour deux filtres de la bande K, caméra C50S, à gauche filtre centré sur $2,03 \mu\text{m}$ et à droite pour un filtre centré sur $2,39 \mu\text{m}$. La valeur moyenne des estimées est tracée et les barres d'erreur correspondent à la valeur minimum et maximum estimée.

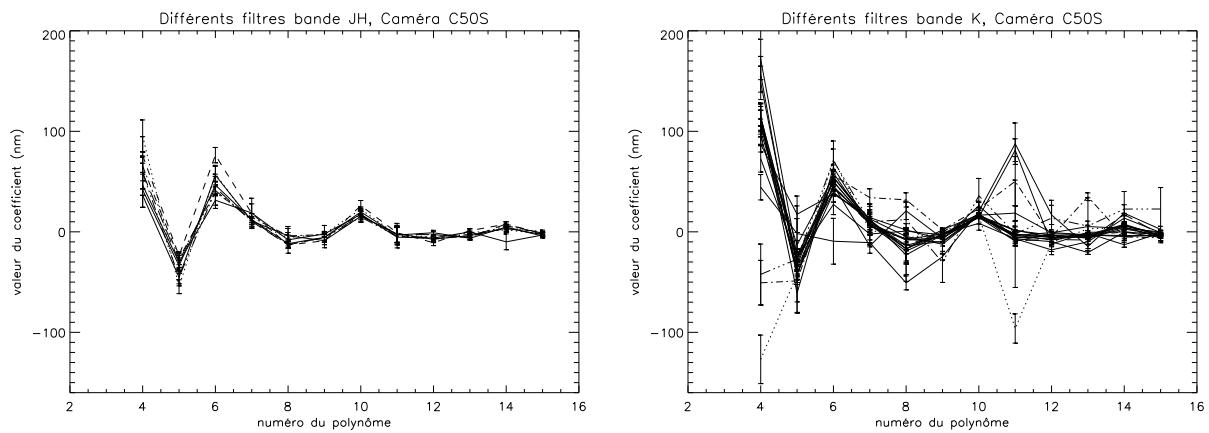


FIG. 3.11 – *Courbe de gauche : valeurs moyennes des aberrations de différentes associations caméra C50S +filtres de la bande J, obtenues à partir des différentes paires de trous avec variation à $\pm\sigma$, à droite la même chose en K.*

C50S) auxquelles s'ajoutent celles introduites par le filtre. On souhaite cependant pouvoir séparer les aberrations introduites par la caméra de celles introduites par les filtres, afin de constituer une table d'aberrations associant à chaque composant de CONICA les aberrations qu'il introduit et permettant ainsi de connaître les aberrations à corriger pour toutes les configurations de CONICA sans avoir à refaire la mesure à chaque fois. Comme nous l'avons déjà évoqué, on propose d'associer la valeur médiane² des aberrations obtenues sur l'ensemble des filtres à la caméra elle-même (on rappelle que les aberrations de la caméra sont proches d'être achromatiques, voir discussion page 55) et l'écart entre cette valeur et les aberrations de la caméra + le filtre, au filtre. Afin de valider cette méthode de séparation, on compare la médiane des aberrations obtenue dans chacune des bandes (figure 3.13). L'adéquation entre les deux courbes est très bonne, excepté pour la défocalisation. Comme nous l'avons évoqué précédemment, l'estimation de ce coefficient est entachée d'une grande incertitude et ne sera donc pas estimé par cette méthode (voir paragraphe 3.8). Cette très bonne concordance des résultats obtenus dans les deux bandes, valide ainsi l'implémentation pratique que nous proposons. Cette méthode permet donc de connaître les aberrations introduites par l'association de deux filtres et d'une caméra, en ajoutant simplement, aux aberrations introduites par la caméra, les valeurs associées à chacun des filtres.

La figure 3.13 présente les aberrations associés à chacun des filtres (c'est la valeur moyenne obtenue à partir des différentes paires de trous qui est reportée sur les courbes). Les résultats obtenus avec les filtres de la bande JH sont très homogènes. Les aberrations obtenues avec les filtres en K sont plus disparates : on retrouve une majorité de filtres assez semblables et présentant des aberrations proches de celles des filtres en JH mais quelques filtres introduisent des aberrations plus fortes (notamment de la défocalisation et de l'aberration sphérique et pour d'autres plus faiblement de l'astigmatisme ou de la coma). Ces variations de qualité optique observées pour certains filtres proviennent des différents procédés utilisés lors de leur fabrication.

2. En pratique, on choisit la médiane et non la valeur moyenne afin de minimiser l'influence de quelques filtres très aberrants (en K)

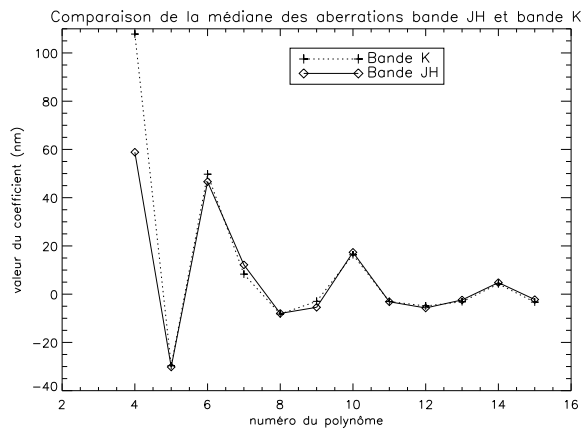


FIG. 3.12 – Comparaison de la médiane des aberrations obtenue à partir des différentes associations caméra C50S + filtres de la bande J et filtres de la bande K.

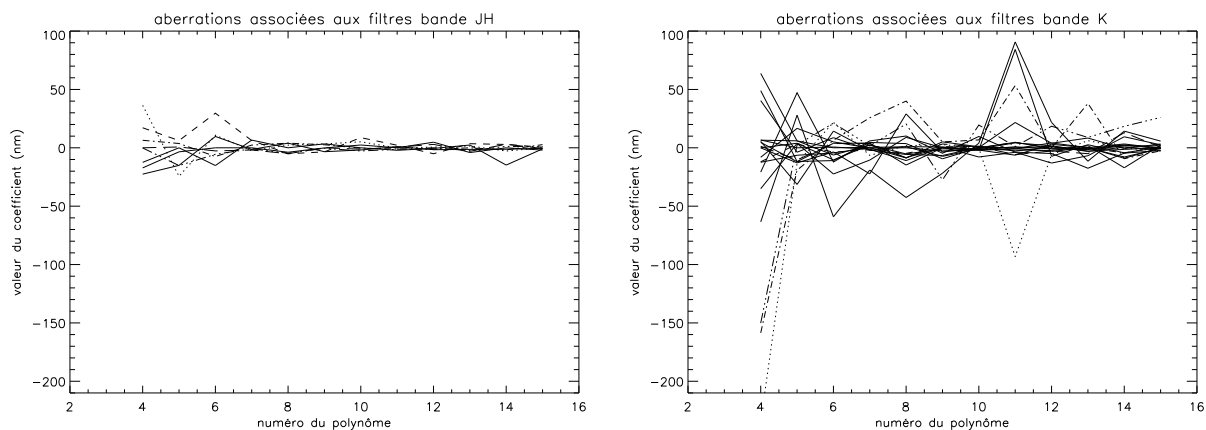


FIG. 3.13 – Valeurs des aberrations associées aux différents filtres de la bande JH (à gauche) et de la bande K (à droite). Il s'agit des valeurs moyennes obtenues à partir des différentes paires de trous.

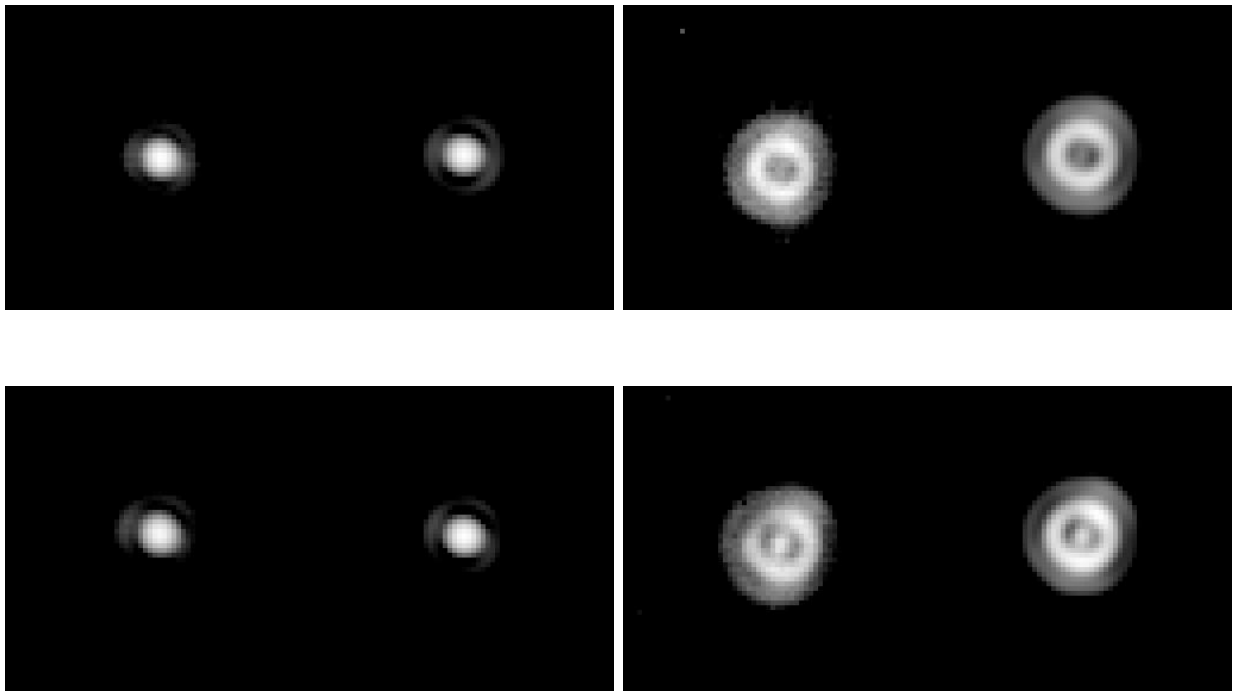


FIG. 3.14 – *Comparaison des images et des RI reconstruites pour deux filtres en K (en haut filtre 2,166 μm , en bas filtre 2,39 μm): à gauche dans le plan focal, à droite dans le plan défocalisé (la représentation est logarithmique). Dans chaque imagette, l'image est à gauche et la RI reconstruite à droite.*

3.4.3 Comparaison entre les RI estimées et mesurées

Une première validation qualitative des estimations effectuées par diversité de phase est proposée figure 3.14, elle repose sur la comparaison visuelle entre les images enregistrées dans le plan focal et dans le plan défocalisé et les RI reconstruites à partir des aberrations estimées par la diversité de phase. Pour s'affranchir au maximum de la présence de l'objet, cette comparaison visuelle est effectuée avec deux filtres de la bande K. Les images du haut correspondent au filtre centré sur 2,166 μm et celles du bas au filtre 2,39 μm . Les images plan focal et RI reconstruites correspondantes, sont placées à gauche, celles correspondantes au plan défocalisé à droite. C'est le logarithme des images qui est visualisé. Aussi bien dans le plan focal que pour celui défocalisé, visuellement l'adéquation entre mesure et estimée semble bonne.

3.4.4 Comparaison avec des calculs de Strehl sur l'image plan focal

On complète la comparaison qualitative présentée dans le paragraphe précédent par la comparaison de la valeur du rapport de Strehl calculé à partir de l'image plan focal avec celui calculé à partir des aberrations estimées.

3.4.4.1 Calcul du rapport de Strehl

A partir de l'image plan focal Pour calculer le rapport de Strehl à partir de l'image plan focal, on utilise la définition 1.21. Il convient, d'une part de calculer la FTO optique théorique (hors aberration) du système et d'autre part de déterminer la FTO optique réelle

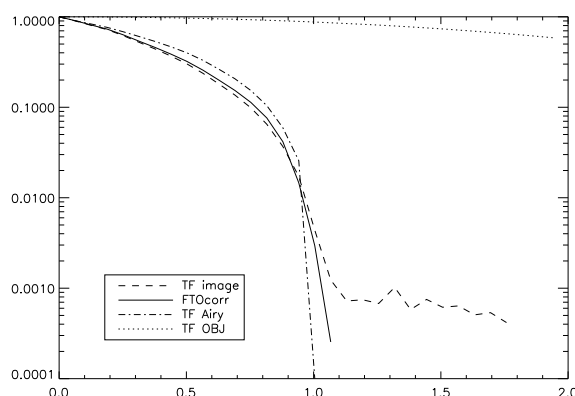


FIG. 3.15 – Calcul de Strehl à partir de l'image plan focal.

à partir de l'image plan focal enregistrée. Si le calcul de la FTO théorique est simple (par autocorrélation de la fonction pupille : disque de diamètre D), l'estimation de la FTO optique réelle est plus délicat. En effet, il faut tenir compte du fait que l'objet observé n'est pas ponctuel, qu'il y a une fonction de transfert associée au capteur et qu'il y a un fond résiduel sur l'image après soustraction du fond enregistré. L'image dont on dispose est le résultat de la convolution de l'objet observé par la RI totale du système : $i = o * (h_{optique} * h_{capteur})$ (voir paragraphe 1.1.3.1). L'influence de l'étendue de l'objet ainsi que celle de la réponse du capteur sont supprimées en divisant la TF image (corrigée préalablement du fond résiduel) d'une part par la transformée de Fourier théorique de l'objet – on suppose que c'est un disque de diamètre $10 \mu\text{m}$ (valeur théorique de perçage) dans le plan d'entrée – et d'autre part par la fonction de transfert théorique du capteur (pixels carrés). Ce qui revient à effectuer une déconvolution par filtre inverse. Un exemple de la moyenne circulaire de la transformée de Fourier de l'image corrigée du fond « TF image », de la FTO reconstruite à partir de l'image corrigée du fond et de l'étendue de l'objet « FTOcorr », de la FTO théorique « TF Airy » ainsi que de la TF théorique de l'objet « TF OBJ » sont présentées sur la figure 3.15. Le rapport de Strehl est calculé en divisant la somme de tous les points (jusqu'à la fréquence de coupure D/λ) de la FTO optique réelle (FTOcorr) par la somme de tous les points de la FTO théorique.

A partir des aberrations estimées Pour la diversité de phase, on utilise la relation qui lie le rapport de Strehl à la variance de la phase aberrante : pour de faibles aberrations, on peut écrire $SR \approx 1 - \sigma_\phi^2$ ou encore en utilisant la relation (1.31) $SR \approx 1 - \sum_{i=4}^{15} a_i^2$.

3.4.4.2 Comparaison

Cette comparaison a été faite avec la caméra C50S pour les différents filtres étroits de la bande JH ainsi que ceux de la bande K. La courbe 3.16 regroupe les résultats, à gauche pour la bande JH et à droite pour la bande K. Les barres d'erreur correspondent, pour le calcul de Strehl sur l'image plan focal, à une variation du diamètre de l'objet de 5 à 15 μm et pour le calcul à partir des aberrations estimées par diversité de phase, aux différents résultats obtenus avec les différentes paires de trous. La courbe correspondant aux filtres de la bande K montre une très bonne adéquation entre le calcul de Strehl fait sur l'image plan focal et celle donnée par les aberrations estimées par diversité de phase. Dans le cas

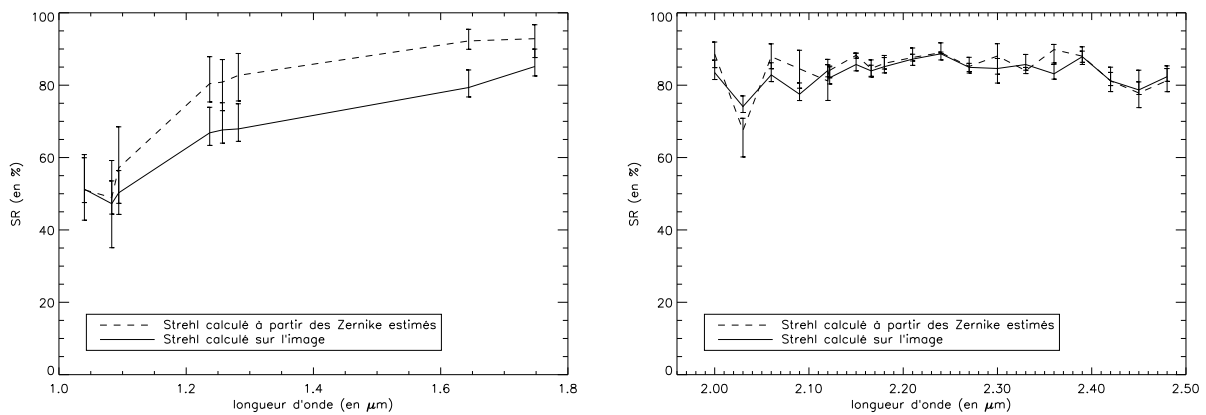


FIG. 3.16 – Comparaison du Rapport du Strehl (SR) calculé par diversité de phase et celui estimé à partir de l'image plan focal (à gauche les filtres de la bande JH, à droite ceux de la bande K). Les barres d'erreurs correspondent aux valeurs minimales et maximales obtenues en faisant varier, pour le calcul sur l'image, le diamètre objet et, pour la diversité de phase, la paire de trous utilisée.

de la bande JH, on observe un écart plus important entre les deux valeurs de Strehl ainsi estimées, excepté pour les deux premiers filtres. Le calcul exact du rapport de Strehl à partir de l'image mesurée est difficile notamment à cause de l'étendue de l'objet qui a d'autant plus d'importance de la longueur d'onde est faible. Vis-à-vis de la diversité de phase, nous avons montré que l'estimation de la défocalisation était assez imprécise notamment à cause de la connaissance imparfaite que nous avons de la distance entre les trous, l'erreur que l'on commet sur l'estimation des aberrations est d'autant plus importante là aussi que la longueur d'onde est faible (voir relation 3.1). Ces remarques fournissent quelques explications quant à l'écart observé entre les deux méthodes de mesure mais il reste à comprendre pourquoi cela n'affecte pas l'estimation des deux premiers filtres.

3.5 Conclusion

La diversité de phase a été appliquée à un problème concret : l'estimation des aberrations d'une caméra d'imagerie. Lors de la conception de cette caméra, rien n'avait été prévu pour calibrer ces défauts. Lorsqu'il a semblé important de connaître ces aberrations statiques afin d'augmenter les performances de NAOS+CONICA, les contraintes imposées par le système, entièrement enfermé dans une enceinte cryogénique, rendaient ces calibrations difficiles. La diversité de phase de part sa légèreté de mise en œuvre optique répondait aux exigences imposées par cette application. Il a cependant fallu adapter la méthode au cas particulier de CONICA ce qui a été à l'origine d'un certain nombre d'incertitudes ou non respect des contraintes imposées par la diversité de phase : la défocalisation est introduite en traduisant le plan d'entrée et non le détecteur, les distances inter-plans ne sont pas parfaitement connues, ce n'est pas le même objet ni strictement les mêmes aberrations qui sont observées par les deux plans de mesure. Des problèmes liés à CONICA ont été révélés par la diversité de phase au cours de l'étude : mauvais alignement de la pupille, présence d'un fond résiduel et incertitude sur la facteur d'échantillonnage. L'impact de ces différentes sources d'incertitudes sur l'estimation des aberrations ayant été étudié par simulations, des mesures

systematiques des différentes caméras et filtres ont été effectuées : l'estimation conjointe objet/aberrations a donnée de très bons résultats validés *in fine* par un gain mesuré en sortie du système NAOS+CONICA. Ces mesures ont été, par la suite, complétées par les mesures des aberrations introduites par les dichroïques de NAOS. Le système est à présent installé au foyer Nasmyth de l'UT4 du VLT et dispose d'une table d'aberrations correspondant aux défauts de chaque dichroïque, caméra et filtre (il est prévu de refaire périodiquement des mesures de ces aberrations pour tenir compte d'éventuelles variations). Ces défauts sont pris en compte et compensés en utilisant le miroir déformable.

Cette étude expérimentale a montré que la diversité de phase est une méthode dont la simplicité de mise en œuvre optique lui permet de s'adapter aux contraintes expérimentales d'un système préexistant et qui demeure fiable malgré différentes sources d'incertitudes.

Ce travail a fait l'objet d'un rapport technique pour l'ESO (Fusco et al., 2002) ainsi que d'un article pour A&A (soumis) qui est constitué de deux parties : « Calibration of NAOS and CONICA static aberrations: application of the phase diversity technique » et « Calibration of NAOS-CONICA aberrations: experimental results ». On pourra trouver dans cet article des compléments à l'étude présentée ici et notamment la calibration des lames dichroïques de NAOS. Cet article se trouve dans l'annexe G.

Chapitre 4

Estimation marginale des aberrations

4.1 Introduction

Au cours du chapitre 2, différents défauts de l'estimation conjointe de la phase aberrante et de l'objet ont été mis en évidence. L'erreur sur l'estimée des inconnues ne diminue pas lorsque le nombre de données traitées augmente à moins de dérégler l'estimateur. Cette méthode n'offre pas de possibilité

$\phi(l, m)$ représente les pixels de phase dans la pupille. En utilisant l'expression de $\partial h_t / \partial a_i$ établie dans (Paxman et al., 1992), on peut écrire :

$$\frac{\partial J_{\text{MAP}}}{\partial a_i} = \sum_t \left\{ \frac{2}{\mathcal{A}} \sum \right.$$

4.7 Propriétés asymptotiques

Après avoir validé les bons résultats obtenus par l'estimateur marginal J_{MV} à fort flux, on s'intéresse à l'évolution de l'erreur de reconstruction des aberrations avec le niveau de bruit et le nombre de données traitées. Encore une fois, les conditions de simulation sont identiques au paragraphe 2.3.2.1 et comme précédemment les hyperparamètres sont fixés à leur valeur vraie. Les coefficients de a_4 à a_{21} sont reconstruits à l'aide de l'estimateur marginal J_{MV} et les résultats, présentés figure 4.1 (concernant l'objet « vue aérienne »), 4.2 (obtenues avec l'objet « entrepôt ») et 4.3 (objet « maison ») proviennent de 50 estimations de ces coefficients obtenus par le traitement de 50 paires d'images bruitées. Tout comme dans le chapitre 2, pour la taille image 32×32 et 64×64 pixels, les courbes sont le résultat du moyennage des valeurs obtenues pour toutes les sous-images de 32×32 et 64×64 pixels contenues dans l'image de 128×128 pixels.

Pour les trois objets, les courbes montrent que le biais, l'écart-type et donc l'erreur (EQM) diminuent lorsque la taille de l'image traitée augmente. Ce comportement s'observe quel que soit le niveau de bruit. Cependant l'objet « entrepôt » présente des écarts faibles entre le biais associé aux images 32×32 pixels et celui des images 64×64 pixels pour des bruits de 10 et 14 %. De même la variance présente un écart faible, entre ces mêmes tailles images, pour 20 et 25 % de bruit. Pour la variance, ce comportement s'observe pour des niveaux de bruit très importants qui conduisent à des erreurs d'estimation des aberrations proche de 100%. Les données étant très peu informatives, le traitement d'une image 64×64 pixels ne suffit apparemment pas à améliorer l'estimation des aberrations (alors que c'est le cas pour l'image 128×128 pixels). En ce qui concerne le biais, le manque d'amélioration est plus surprenant. Pour les trois objets, l'évolution de l'écart-type en fonction du niveau de bruit est proche d'être linéaire. Pour l'objet « vue aérienne », l'évolution du biais avec le niveau de bruit est également linéaire, pour les deux autres objets, il est plus difficile de déduire un comportement global du biais certainement par manque de moyennage d'un plus grand nombre d'images qui permettrait de « lisser » les courbes.

On peut comprendre, intuitivement, le meilleur comportement asymptotique par rapport à la méthode jointe régularisée par le fait qu'ici, contrairement au cas de l'estimation jointe, le traitement d'une image plus grande ne modifie pas le nombre d'inconnues à estimer : le rapport données sur inconnues augmente avec le nombre de données traitées. Ces constatations sont en adéquation avec des propriétés connues du maximum de vraisemblance (Lehmann, 1983; Carvalho et Slock, 1997). En effet, sous des conditions mathématiques peu restrictives, cet estimateur est asymptotiquement non biaisé et à variance nulle.

Une deuxième différence par rapport aux courbes de l'estimateur joint est à noter : le problème des minima locaux, rencontré avec l'objet « vue aérienne » n'apparaît pas ici. On s'attend effectivement à ce que l'estimateur marginal, contrairement au joint, présente de moins en moins de minima locaux lorsqu'on augmente la taille des images traitées. En effet d'après un théorème de type centrale limite, la nappe de vraisemblance devient de plus en plus gaussienne lorsque le nombre de données traitées augmente.

On vient de montrer que l'estimateur marginal a de bien meilleures propriétés asymptotiques que l'estimateur joint : le biais, la variance ainsi que le nombre de minima locaux diminuent lorsque le nombre de données traitées augmente.

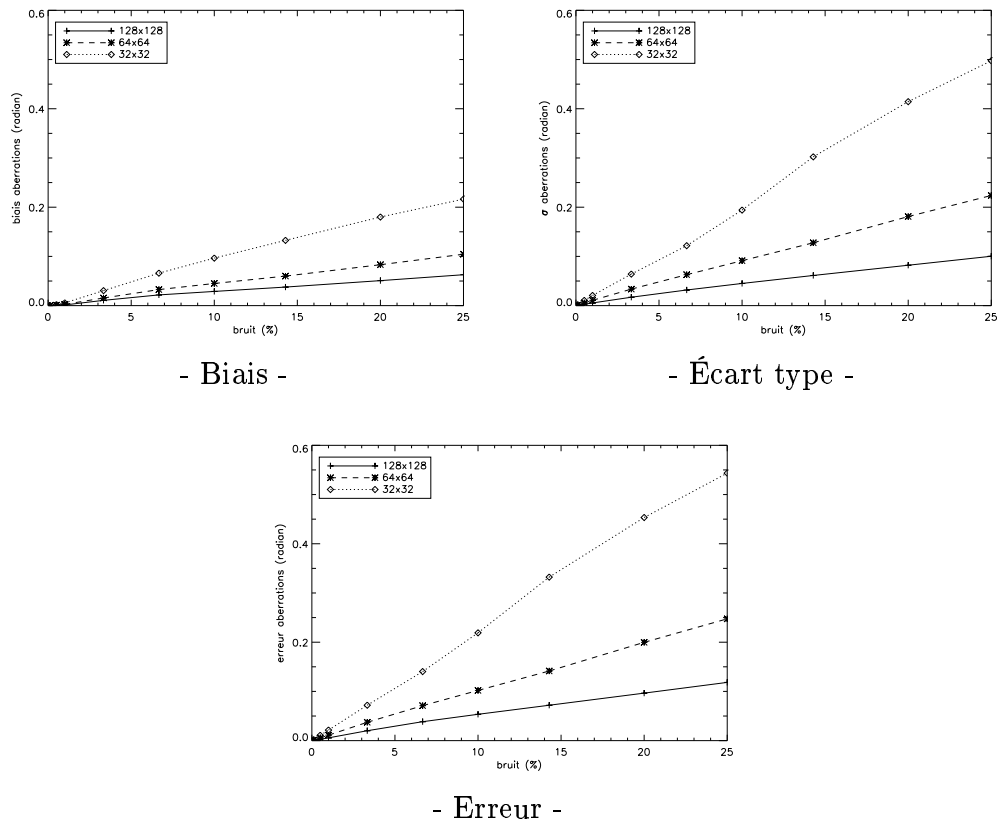


FIG. 4.1 – Propriétés asymptotiques - Scène vue aérienne haute.

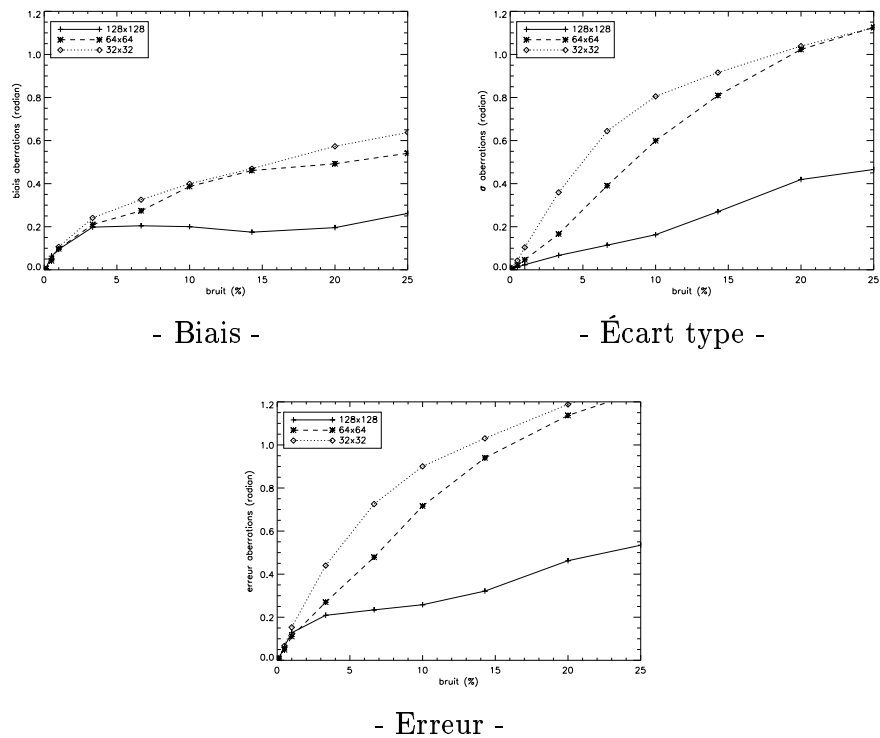


FIG. 4.2 – Propriétés asymptotiques - Scène entrepôt.

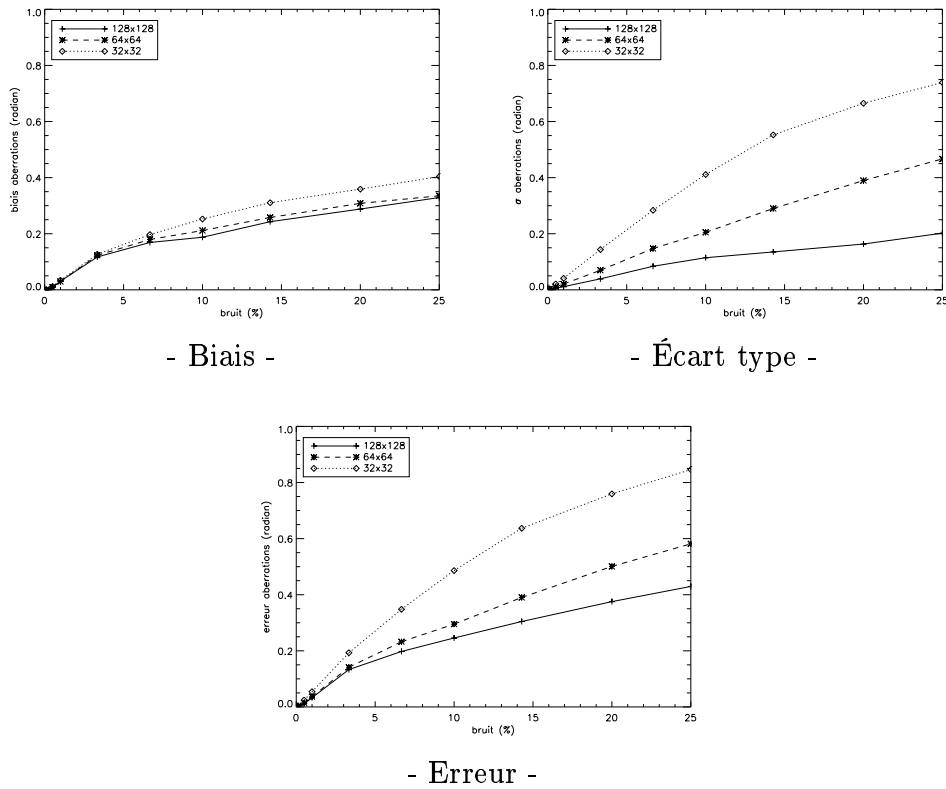


FIG. 4.3 – Propriétés asymptotiques - Scène Maison.

4.8 Estimation non supervisée des hyperparamètres

Tout comme dans la méthode jointe, l'estimation marginale nécessite le réglage de la valeur des hyperparamètres. Notamment, l'ajustement de ceux liés à l'objet (notés θ_o) est particulièrement difficile (voir paragraphe 2.2.3). Mais contrairement à l'estimation jointe, où il a été montré qu'il n'était pas possible de les estimer conjointement avec l'objet et les aberrations, la méthode marginale résout simplement ce problème : les hyperparamètres liés à l'objet θ_o ainsi que ceux liés au modèle de bruit θ_b peuvent être estimés conjointement avec les aberrations.

4.8.1 Principe

La densité de probabilité $f(i_f, i_d, \mathbf{a}; \boldsymbol{\theta})$ est maximisée par rapport aux aberrations \mathbf{a} ET par rapport aux hyperparamètres $\boldsymbol{\theta}_b$ et $\boldsymbol{\theta}_o$:

$$(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}}_o, \hat{\boldsymbol{\theta}}_b) = \arg \max_{\mathbf{a}, \boldsymbol{\theta}_o, \boldsymbol{\theta}_b} f(i_f, i_d, \mathbf{a}; \boldsymbol{\theta}) \quad (4.25)$$

L'expression du critère J_{MAP} reste inchangée mais il doit être, à présent, minimisé par rapport aux aberrations \mathbf{a} , et aux quatre hyperparamètres que sont l'écart type du bruit σ^2 et les trois paramètres du modèle de la densité spectrale de puissance de l'objet k , v_o et p .

4.8.2 Minimisation du critère

4.8.2.1 Changement de variable

En faisant le changement de variable $\mu = \sigma^2/k$ dans l'équation (4.18), on peut trouver une expression analytique de k qui minimise le critère J_{MAP} à \mathbf{a} , μ , v_o et p fixés. En remplaçant σ^2 par $k\mu$ dans l'expression (4.18), on obtient le critère $J_{\text{MAP}}(\mathbf{a}, \mu, k, v_o, p)$. L'annulation du gradient de J_{MAP} par rapport à k fournit une expression analytique de \hat{k} qui minimise le critère à \mathbf{a} , μ , v_o , et p fixés. En posant $S'_o = kS_o$, l'expression de \hat{k} s'écrit :

$$\hat{k} = \sum_v \frac{|\tilde{i}_d(v)\tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v)\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \left[|\tilde{h}_f(\mathbf{a}, v)\tilde{o}_m(v) - \tilde{i}_f(v)|^2 + [|\tilde{h}_d(\mathbf{a}, v)\tilde{o}_m(v) - \tilde{i}_d(v)|^2] \right]}{2N^2\mu \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \quad (4.26)$$

En remplaçant k par \hat{k} dans l'expression du critère, on obtient le critère $J'_{\text{MAP}}(\mathbf{a}, \mu, v_o, p)$ qui ne dépend alors plus que des aberrations et de trois hyperparamètres. Cependant afin d'alléger l'écriture du critère et surtout des gradients, on n'utilise pas l'expression réinjectée J'_{MAP} mais on conserve le critère J_{MAP} pour le calcul des gradients. Ceci est possible car $\frac{\partial J'_{\text{MAP}}(\mathbf{a}, \mu, v_o, p)}{\partial x} = \frac{\partial J_{\text{MAP}}(\mathbf{a}, \mu, k, v_o, p)}{\partial x} \Big|_{k=\hat{k}}$ avec $x = \{\mathbf{a}, \mu, v_o, p\}$ (la démonstration de cette relation est identique à celle rencontrée au paragraphe 2.2.4).

4.8.2.2 Calcul des gradients

L'annulation de la dérivée du critère par rapport aux autres hyperparamètres μ , v_o et p ne fournit pas d'expression analytique. La minimisation vis-à-vis de ces grandeurs devra être effectuée itérativement. Les expressions des trois gradients sont proposées dans l'annexe E.

4.8.2.3 Méthode de minimisation et reparamétrisation des hyperparamètres

Afin d'estimer les aberrations \mathbf{a} et les hyperparamètres μ , v_o et p , une minimisation itérative conjointe de J_{MAP} doit être effectuée par rapport à ces différentes grandeurs. Les aberrations et les hyperparamètres $\boldsymbol{\theta}' = (\mu, v_o, p)$ sont estimés par minimisation alternée (voir chapitre 2.2.4) de J_{MAP} vis-à-vis de \mathbf{a} d'une part et des hyperparamètres $\boldsymbol{\theta}'$ d'autre part. Deux méthodes de minimisation sont alors possibles pour les hyperparamètres, une minimisation semi-globale du critère vis-à-vis des trois ou une minimisation alternée. Ce qui conduit pour l'itération l de la minimisation aux deux schémas suivants :

$$\text{Méthode alternée} \begin{cases} \hat{\mathbf{a}}^{(l)} = \arg \min_{\mathbf{a}} J_{\text{MAP}}(\mathbf{a}, \mu^{(l-1)}, v_o^{(l-1)}, p^{(l-1)}) \\ \hat{\mu}^{(l)} = \arg \min_{\mu} J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \mu, v_o^{(l-1)}, p^{(l-1)}) \\ \hat{v}_o^{(l)} = \arg \min_{v_o} J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \hat{\mu}^{(l)}, v_o, p^{(l-1)}) \\ \hat{p}^{(l)} = \arg \min_p J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \hat{\mu}^{(l)}, \hat{v}_o^{(l)}, p) \end{cases} \quad (4.27)$$

$$\text{Méthode semi-globale} \begin{cases} \hat{\mathbf{a}}^{(l)} = \arg \min_{\mathbf{a}} J_{\text{MAP}}(\mathbf{a}, \mu^{(l-1)}, v_o^{(l-1)}, p^{(l-1)}) \\ \{\hat{\mu}^{(l)}, \hat{v}_o^{(l)}, \hat{p}^{(l)}\} = \arg \min_{\mu, v_o, p} J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \mu, v_o, p) \end{cases} \quad (4.28)$$

Des considérations physiques permettent de connaître approximativement l'ordre de grandeur des trois hyperparamètres (voir paragraphe 2.2.3) : p qui est la puissance de décroissance spectrale de l'objet est compris entre 1 et 4 et μ est proche de l'inverse du flux total dans l'image pour un bruit essentiellement photonique. On peut alors reparamétriser ces trois

grandeurs afin de tenir compte de ces connaissances et notamment imposer leur positivité. Cette contrainte, justifiée physiquement, est de plus indispensable afin que les paramètres ne prennent pas de valeurs négatives qui rendraient le critère incalculable (présence de terme logarithmique). La positivité de μ et v_o est imposée par projection, à chaque itération l de la minimisation, de l'estimée courante $\hat{\mu}^{(l)}$ et $\hat{v}_o^{(l)}$ sur \mathcal{R}^+ . Nous allons voir que suivant la méthode de minimisation employée (alternée ou globale), une reparamétrisation supplémentaire doit être effectuée.

Minimisation alternée A l'étape l , le critère est minimisé successivement par rapport à chacun des hyperparamètres. Ceci permet de ne pas avoir à gérer le problème de différence d'échelle entre les gradients. Dans ce cas, la contrainte de positivité imposée par reparamétrisation suffit. L'inconvénient de cette méthode est d'être coûteuse en temps de calcul. Le ralentissement de la convergence provient du fait que chaque hyperparamètre du modèle de DSP objet est modifié séparément à chaque étape. Pour accélérer la convergence on propose de modifier le schéma de minimisation en rajoutant une boucle supplémentaire : au cours de l'itération l , les minimisations successives du critère par rapport aux trois hyperparamètres sont répétées n fois ($n=3$ en pratique).

$$\left\{ \begin{array}{l} \hat{\mathbf{a}}^{(l)} = \arg \min_{\mathbf{a}} J_{\text{MAP}}(\mathbf{a}, \mu^{(l-1)}, v_o^{(l-1)}, p^{(l-1)}) \\ \text{Étape } l' \text{ (répétée } n \text{ fois)} \\ \left\{ \begin{array}{l} \hat{\mu}^{(l')} = \arg \min_{\mu} J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \mu, v_o^{(l'-1)}, p^{(l'-1)}) \\ \hat{v}_o^{(l')} = \arg \min_{v_o} J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \hat{\mu}^{(l')}, v_o, p^{(l'-1)}) \\ \hat{p}^{(l')} = \arg \min_p J_{\text{MAP}}(\hat{\mathbf{a}}^{(l)}, \hat{\mu}^{(l')}, \hat{v}_o^{(l')}, p) \end{array} \right. \end{array} \right.$$

Minimisation semi-globale A l'étape l , les trois hyperparamètres sont minimisés globalement. Au cours de la minimisation, les essais successifs effectués par l'algorithme de minimisation peuvent conduire à des jeux d'hyperparamètres aberrants, notamment des valeurs de la puissance trop importantes qui entraînent des divisions par zéro lors de l'évaluation du critère. On doit alors imposer une reparamétrisation plus forte que la seule contrainte de positivité. La puissance p est contrainte sur l'intervalle $[0,4]$ par reparamétrisation, on pose $p = 1,3 \left(\arctan(\alpha) + \frac{\pi}{2} \right)$. Moyennant cette reparamétrisation supplémentaire et en prenant garde que les gradients des trois hyperparamètres aient des poids équivalents (en reparamétrisant certains hyperparamètres – voir technique expliquée page 36), cette méthode permet d'atteindre plus rapidement la convergence que la méthode alternée.

Initialisation La rapidité de convergence dépend également du point d'initialisation. Afin de partir d'un jeu d'hyperparamètre proche des vrais, on effectue une première estimation des hyperparamètres par Maximum de Vraisemblance à partir de la valeur initiale donnée aux aberrations \mathbf{a}_{init} (généralement zéro):

$$(\mu, v_o, p)_{init} = \arg \max_{\mu, v_o, p} f(\mathbf{i}_f, \mathbf{i}_d, \mathbf{a}_{init}; \boldsymbol{\theta}_\phi) \quad (4.29)$$

Ce qui revient à minimiser le critère $J_{\text{MAP}}(\mathbf{a}_{init}, \boldsymbol{\theta})$ (équation 4.18) vis-à-vis uniquement des hyperparamètres. Cette première estimation est rapide et s'est avérée robuste.

4.8.3 Résultats de l'estimation non supervisée

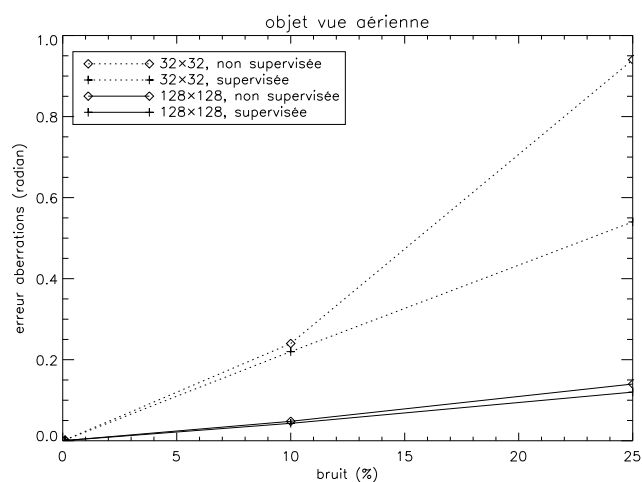
Pour valider la méthode marginale non supervisée, on présente les résultats d'estimations obtenus dans les mêmes conditions que celles du paragraphe 4.6 mais cette fois en estimant conjointement avec les aberrations, les hyperparamètres et en faisant varier le niveau de bruit (0,1%, 10% et 25%). Le tableau 4.2 regroupe les résultats d'estimations obtenus dans les mêmes conditions avec la méthode non supervisée et, pour comparaison, avec la méthode supervisée (c'est-à-dire en utilisant les hyperparamètres vrais). Ces résultats sont également présentés sous forme de courbes – figure 4.4. Les résultats correspondent, pour chaque objet, à l'erreur (rms) sur l'estimée des aberrations (voir paragraphe 2.3.2.2). A 0,1 % de bruit,

		32×32			128×128			
		bruit	0,1%	10%	25%	0,1%	10%	25%
« vue aérienne »	supervisée	0,002	0,22	0,54	0,0006	0,043	0,12	
	non supervisée	0,002	0,24	0,94	0,0006	0,048	0,14	
« entrepôt »	supervisée	0,011	0,9	1,3	0,005	0,26	0,53	
	non supervisée	0,011	1,5	2,5	0,005	0,74	2,0	
« maison »	supervisée	0,004	0,48	0,8	0,0011	0,24	0,43	
	non supervisée	0,004	0,77	1,3	0,0011	0,32	0,88	

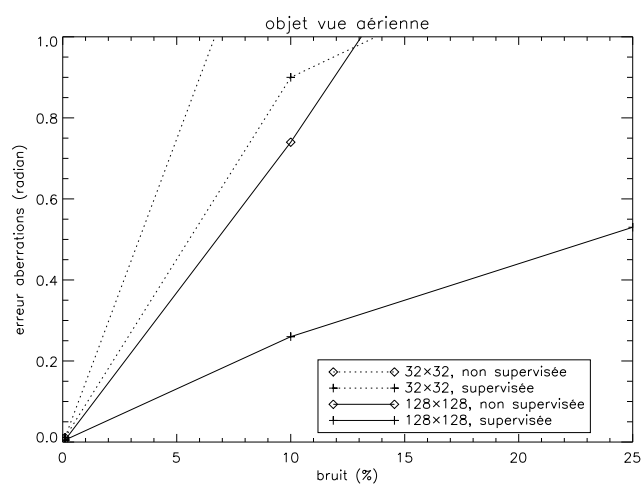
TAB. 4.2 – Comparaison de l'erreur (rms) en radian sur l'estimée des aberrations, obtenue par la méthode supervisée et par la méthode non supervisée (une moyenne sur 50 réalisations de bruit a été faite).

quel que soit l'objet et la taille image traitée, la qualité de l'estimée des aberrations obtenue avec la méthode non supervisée est aussi bonne qu'avec les hyperparamètres vrais. Pour l'image 32×32 pixels, à 10 % de bruit, l'estimation non supervisée reste proche de celle supervisée pour l'objet « vue aérienne », mais est de bien moins bonne qualité dans le cas des deux autres objets. Pour cette même taille image, à 25 % de bruit, pour les trois objets, l'erreur sur l'estimée des aberrations est proche d'être deux fois plus importante que celle obtenue avec les hyperparamètres vrais. Dans ce cas, les conditions d'estimations sont très défavorables : il y a beaucoup de bruit et peu de données. Il est alors difficile, pour la méthode non supervisée, d'extraire à la fois les aberrations et les hyperparamètres.

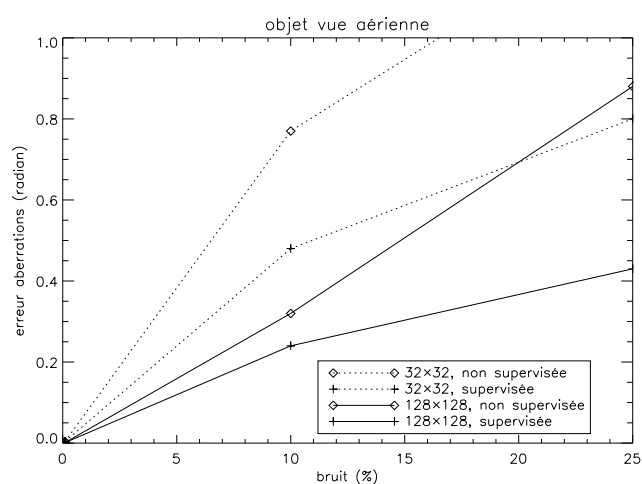
Pour l'image 128×128 pixels, quel que soit l'objet, l'écart entre l'estimée des aberrations obtenue avec la méthode supervisée et celle non supervisée s'est réduit par rapport au cas de l'image 32×32 pixels : l'ajout de données rend plus favorable la double estimation hyperparamètres-aberrations effectuée par la méthode non supervisée. Pour l'objet « vue aérienne », l'estimation non supervisée donne de très bons résultats même pour 25 % de bruit. Pour l'objet « maison », jusqu'à 10 % de bruit, les résultats non supervisés restent corrects par rapport à ceux supervisés, pour l'objet « entrepôt » à 10 % de bruit, la qualité de l'estimation s'est dégradée de plus de 50 %. L'estimation non supervisée reste décevante même à 128×128 pixels dans le cas de l'objet « entrepôt » qui comme on l'a déjà dit est nettement moins favorable à l'estimation des aberrations. Dans ce cas, il faudrait augmenter encore le nombre de données traitées pour espérer se rapprocher de la qualité d'estimation obtenue par la méthode supervisée.



- Vue aérienne -



- Entrepôt -



- Maison -

FIG. 4.4 – Comparaison de l'erreur (rms) en radian sur l'estimée des aberrations obtenue par la méthode supervisée (losange) et par la méthode non supervisée (croix) pour des images de taille 32×32 pixels (pointillé) et 128×128 pixels (ligne continue).

4.9 Validation sur données expérimentales

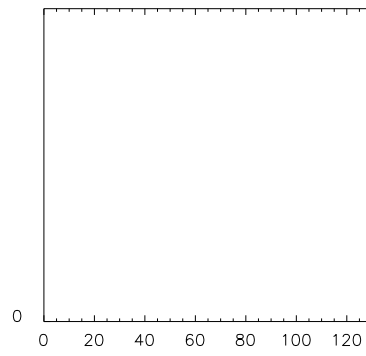
Avant de s'intéresser à une comparaison des performances obtenues avec chacun des estimateurs joint et marginal sur données simulées et expérimentales, ce paragraphe propose une première validation sur données expérimentales de l'estimateur marginal. Le problème du traitement des images très étendues est tout d'abord abordé, puis le montage expérimental est décrit et enfin on présente les résultats obtenus.

4.9.1 Objet étendu : Traitement des effets de bords

Afin d'accélérer le traitement des données, le critère J et les gradients sont calculés dans le domaine de Fourier. Cependant comme il a été dit dans le chapitre 2, lorsque l'objet observé est infiniment étendu (ce qui est le cas de l'observation de la Terre ou du soleil) le traitement par transformée de Fourier n'est plus possible : il y a, lorsque l'on effectue le passage dans le domaine de Fourier, une périodisation implicite des grandeurs concernées qui n'est alors plus valide. Il faut donc trouver une solution pour pouvoir continuer à effectuer proprement les calculs. Deux méthodes ont été proposées dans la littérature de la diversité de phase afin de résoudre ce problème dans le cas de l'estimation conjointe. L'expression du critère marginal ne différant de celle du critère joint que par l'ajout d'un terme, ces deux solutions sont donc également applicables au cas de l'estimateur marginal. Dans les paragraphes suivants, ces deux méthodes sont détaillées ainsi que leur extension au cas de l'estimateur marginal.

4.9.1.1 Apodisation

Principe Afin de pouvoir continuer à utiliser l'expression du critère J_{MAPJ} , équation (2.10) dans laquelle est réinjecté l'objet $\hat{\delta}$, relation (2.11), tout en minimisant les effets de bords, les images enregistrées peuvent être apodisées. L'apodisation entière de l'image par une fenêtre de Hanning est une première solution proposée par Paxman (Paxman et Crippen, 1990). Löfdahl (Löfdahl et Scharmer, 1994) en se basant sur la même idée, a utilisé une fenêtre de Hanning modifiée : seuls les bords de l'image sont apodisés (un exemple 1D de cette fenêtre est représenté figure 4.5). Afin de mettre en évidence l'influence des effets de bord sur le calcul du critère pour les différents traitements de l'image (image non apodisée, totalement apodisée ou uniquement apodisée sur les bords), la figure 4.6 présente, pour ces différents cas, la représentation 2D de l'erreur $|TF^{-1}(\tilde{\nu}_f - \tilde{h}_f \hat{\delta})|^2 + |TF^{-1}(\tilde{\nu}_d - \tilde{h}_d \hat{\delta})|^2$ dont la sommation sur tous les points image constitue une partie du critère J_{MAPJ} , équation (2.10); l'équivalence entre sommation dans le domaine réel et dans le domaine de Fourier est donnée par le théorème de Parseval. La représentation est logarithmique. Ces erreurs correspondent à une image 128×128 pixels de l'objet « vue aérienne » dégradée par les aberrations présentées au paragraphe 2.3.2.1 mais sans bruit, en se plaçant aux valeurs vraies des aberrations : l'erreur devrait alors être nulle partout. Pour l'image sans apodisation, on constate que l'erreur se répartit dans toute l'image avec un effet fort visible sur les bords. Pour l'image apodisée par une fenêtre de Hanning, l'erreur est beaucoup plus faible sur les bords mais est importante au centre. Pour l'image apodisée par une fenêtre de Hanning modifiée (apodisation de deux fois seize pixels dans chaque direction), l'erreur est importante sur les bords mais pratiquement nulle au centre du champ. On a calculé la valeur totale de l'erreur en sommant tous les points de la carte de l'erreur ou en ne sommant que les points du centre (entouré par un carré noir sur la figure 4.6) : le tableau 4.3 rassemble les résultats. Les effets de bords étant



Application au cas de l'estimateur marginal Le critère marginal et le critère joint sont liés par la relation 4.19, la différence porte sur la présence d'un terme supplémentaire dans le cas du critère marginal, ce terme ne dépend que des RIs associées à chaque plan. Le traitement par apodisation, proposé dans le cadre de l'estimation jointe, s'applique donc également dans le cas marginal.

4.9.1.2 Bande de Garde

Principe Une autre solution au traitement des objets très étendus a été proposée par Seldin (Seldin et Paxman, 1994). Elle permet de conserver le calcul de la convolution $h * o$ dans le domaine de Fourier mais sans pouvoir écrire de solution analytique pour l'objet. C'est donc la version alternée du critère joint qui est à présent utilisée : $J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta})$ (relation (2.10)). Afin de pouvoir continuer à calculer rapidement la convolution par transformée de Fourier, tout en résolvant le problème de la périodisation implicite ainsi effectuée de l'objet et de la RI, on ajoute une bande de zéros (appelée bande de garde) tout autour de ces grandeurs. Notons $N \times N$ la taille en pixels de l'image mesurée (domaine dénommé *FOV*) et B la largeur de la bande de garde ajoutée de part et d'autre, la taille des images, des RI et de l'objet qui interviennent dans l'équation (2.10) devient donc $(N + 2B) \times (N + 2B)$. Pour que la convolution circulante soit calculée correctement, la largeur de cette bande de garde doit être au moins égale à la largeur du support de la RI (elle dépend donc de la force des aberrations). Les pixels de la bande de garde ne correspondant pas à de vraies mesures, on prend soin de ne pas les inclure dans le calcul du critère, pour cela on effectue les sommations de l'équation (2.10) dans le domaine direct (ce qui est possible grâce au théorème de Parseval), en ne sommant que les points du centre $N \times N$.

L'avantage de cette solution de traitement des effets de bords est d'utiliser toutes les données enregistrées sans perte d'information mais l'inconvénient majeur est de considérablement augmenter le temps de traitement (la minimisation vis-à-vis de l'objet est effectuée itérativement).

Application au cas de l'estimateur marginal Afin d'appliquer cette bande de garde au cas du critère marginal, on va montrer que, tout comme dans le cas de l'estimateur joint, on peut construire une version « alternée » du critère marginal (c'est-à-dire minimiser également itérativement le critère vis-à-vis de l'objet). La relation qui lie le critère joint J'_{MAPJ} au marginal J_{MAP} peut être résumée ainsi : $J_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta}) = J'_{\text{MAPJ}}(\mathbf{a}, \boldsymbol{\theta}) + \varepsilon(\mathbf{a}, \boldsymbol{\theta})$. On appelle critère marginal alterné $J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta})$ le critère défini par $J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}) = J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}) + \varepsilon(\mathbf{a}, \boldsymbol{\theta})$. On a :

$$\begin{aligned}
 \arg \min_{\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}} J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}) &= \arg \min_{\mathbf{a}, \boldsymbol{\theta}} \left\{ \arg \min_{\mathbf{o}} J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}) \right\} \\
 &= \arg \min_{\mathbf{a}, \boldsymbol{\theta}} \left\{ \arg \min_{\mathbf{o}} [J_{\text{MAPJ}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta})] + \varepsilon(\mathbf{a}, \boldsymbol{\theta}) \right\} \\
 &= \arg \min_{\mathbf{a}, \boldsymbol{\theta}} [J'_{\text{MAPJ}}(\mathbf{a}, \boldsymbol{\theta}) + \varepsilon(\mathbf{a}, \boldsymbol{\theta})] \\
 &= \arg \min_{\mathbf{a}, \boldsymbol{\theta}} J_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta}) \tag{4.30}
 \end{aligned}$$

Minimiser $J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta})$ par rapport à \mathbf{o} , \mathbf{a} et $\boldsymbol{\theta}$ est donc équivalent à minimiser le critère $J_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta})$ par rapport à \mathbf{a} et $\boldsymbol{\theta}$ uniquement. En utilisant la relation (2.10) et (4.19), le

critère marginal alterné s'exprime ainsi :

$$\begin{aligned}
J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, \boldsymbol{\theta}) &= \frac{1}{2} \sum_v \left[\ln \left(|\tilde{h}_f(v)|^2 + |\tilde{h}_d(v)|^2 + \frac{\sigma^2}{S_o(v)} \right) + \frac{|\tilde{o}(v) - \tilde{o}_m(v)|^2}{S_o(v)} \right] \\
&+ \frac{N^2}{2} \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) + \frac{1}{2} \ln \det(R_a) + \frac{1}{2} \mathbf{a}^t R_a^{-1} \mathbf{a} + A' \\
&+ \frac{1}{2\sigma^2} \sum_{(l,m) \in FOV} \left[\left| i_f(l,m) - TF^{-1} \left(\tilde{h}_f(\mathbf{a}, v) \tilde{o}(v) \right) \Pi_{l,m} \right|^2 + \left| i_d(l,m) - TF^{-1} \left(\tilde{h}_d(\mathbf{a}, v) \tilde{o}(v) \right) \Pi_{l,m} \right|^2 \right]
\end{aligned} \tag{4.31}$$

$$\text{avec } \Pi_{l,m} = \begin{cases} 1 & \text{si } (l,m) \in FOV \\ 0 & \text{sinon} \end{cases} \tag{4.32}$$

L'ajout d'une bande de garde entraîne également une modification des expressions des gradients liés aux hyperparamètres. Ces relations sont rassemblées dans la deuxième partie de l'annexe E.

4.9.2 Montage expérimental

4.9.2.1 Dispositif

Afin de valider les performances de l'estimateur marginal sur données expérimentales, on utilise un dispositif permettant d'introduire sur le trajet optique des aberrations parfaitement connues. Ces aberrations sont générées par une lame à faces parallèles montée sur une platine rotative. Les aberrations introduites par la lame dépendent de manière connue de l'angle que celle-ci forme avec l'axe optique. Le montage expérimental est représenté sur la figure 4.7. Une diapositive (éclairée par un projecteur) positionnée dans le plan focal de la lentille L1 simule un objet étendu en lumière blanche situé à l'infini. Un diaphragme situé entre les lentilles L1 et L2 permet de modifier l'ouverture du système. En modifiant le diamètre de la pupille, on peut ainsi adapter le rapport f_c/f_p (voir page 16) des images. Une caméra CCD placée au foyer de la lentille L2 enregistre les images qui sont quasi-monochromatiques grâce à la présence d'un filtre interférentiel centré sur 633 nm ($\Delta\lambda = 10$ nm) situé devant le détecteur. La caméra utilisée dispose de pixels carrés de 9 μm de côté. La focale F du système est de 800 mm. Le diamètre D de la pupille est fixé à 28 mm afin que les images enregistrées soient échantillonnées à Shannon. Les deux images focalisée et défocalisée sont enregistrées successivement ce qui est possible puisque les aberrations recherchées sont lentement variables. On prend soin toutefois de capoter l'ensemble du dispositif expérimental afin de réduire l'influence de la turbulence du laboratoire.

4.9.2.2 Aberrations introduites

La lame introduit les aberrations suivantes : défocalisation a_4 , astigmatisme a_6 , coma a_7 et aberration sphérique a_{11} . L'ouverture du montage optique étant faible ($F/29$), seules les aberrations de défocalisation et d'astigmatisme sont significatives. Leur valeur dépend de l'angle α que forme la lame avec l'axe optique, de l'ouverture du système, de la longueur d'onde, de l'épaisseur de la lame ($e = 6,2$ mm) et de l'indice de la lame ($\epsilon = 1,515$ à 633 nm)

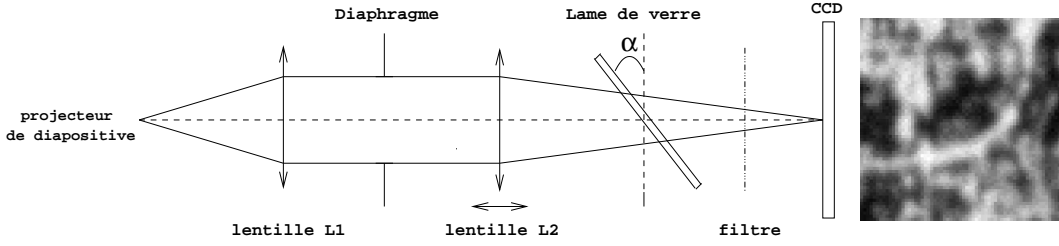


FIG. 4.7 – Schéma du montage expérimental

selon les expressions suivantes (Born et Wolf, 1983) :

$$a_4 = \frac{\pi e(\epsilon - 1)}{8\sqrt{3}\lambda (F/D)^2 \cos(\alpha)\epsilon} \quad (4.33)$$

$$a_6 = \frac{\pi e\alpha^2(\epsilon^2 - 1)}{8\sqrt{6}\lambda (F/D)^2 \epsilon^3} \quad (4.34)$$

4.9.3 Résultats

La validation expérimentale de l'estimateur marginal non supervisé est effectuée en utilisant le montage expérimental décrit ci-dessus². L'objet observé est une diapositive représentant une vue aérienne, le bruit est proche de 0,3% (le chapitre 5 étend cette validation à différents niveaux de bruits). On acquiert cinq paires d'images focalisées/défocalisées correspondant à cinq positions différentes de la lame : angle 0° , 15° , 30° , -30° et 45° . L'image défocalisée est enregistrée en translatant la lentille L2 de 4,2 mm soit une défocalisation de 2π crête-à-crête (voir relation 3.1). On recherche les coefficients de Zernike jusqu'au numéro 21. La translation de la lentille L2 qui permet d'introduire la défocalisation, entraîne un basculement de l'image défocalisée par rapport à l'image plan focal. On prend soin, tout comme dans le paragraphe 3.2, de recentrer préalablement les deux images et afin d'affiner le positionnement, on recherche, en plus des aberrations du système, les tilts relatifs de l'image défocalisée par rapport à celle focalisée. On extrait dans chacune des deux images enregistrées (plan focal et défocalisée), une zone contrastée de taille 64×64 pixels que l'on utilise pour l'inversion. Afin de traiter convenablement les effets de bords dus à l'étendu de l'objet, on utilise une bande de garde de largeur 64×64 pixels.

Le traitement de la paire d'images obtenue sans la lame permet de vérifier que le banc est quasiment exempt d'aberrations : l'amplitude de la phase aberrante est de $\lambda/90$ rms. Après cette vérification, on insère la lame dans le montage en prenant soin de translater la lentille L2 afin de compenser la valeur théorique de la défocalisation introduite par la lame à angle nul, soit $a_4 = 0,925$ radian, ce qui correspond à une translation de 2,10 mm (voir relation 3.1). Le traitement par diversité de phase de la paire d'images obtenue avec la lame à angle nul estime une défocalisation résiduelle de $a_4 = 0,07$ radian (soit 0,16 mm). Le tableau 4.4 présente la comparaison des résultats théoriques et expérimentaux des coefficients a_4 et a_6 pour différentes valeurs de l'angle que forme la lame avec l'axe optique. Pour la défocalisation a_4 , la valeur estimée pour la lame à angle nul a été soustraite des valeurs des a_4 estimées pour les autres angles.

L'erreur la plus importante est obtenue sur l'estimation du coefficients a_6 correspondant à une inclinaison de la lame de 45° , toutefois cet écart reste inférieur à $\lambda/100$. La dissymétrie

². Ce même montage a déjà servi à la validation de l'estimateur joint sans régularisation (Meynadier, 1997).

observée dans les résultats obtenus avec la lame à 30° et -30° peut peut-être s'expliquer par un positionnement incorrect de la lame par rapport à l'axe optique. L'estimation des aberrations par la méthode marginale donne donc, à fort flux, de très bons résultats.

α (degré)	a_4 théorique (radian)	a_4 estimée (radian)	a_6 théorique (radian)	a_6 estimée (radian)
15	0,030	0,055	-0,049	-0,028
30	0,143	0,163	-0,196	-0,188
-30	0,143	0,157	-0,196	-0,210
45	0,383	0,377	-0,442	-0,502

TAB. 4.4 – Comparaison des valeurs théoriques des aberrations introduites par la lame et celles estimées par la méthode marginale non supervisée (valeurs en radian).

4.10 Conclusion

Dans ce chapitre, une méthode d'estimation marginale des aberrations, originale dans le cadre de la diversité de phase, a été présentée et étudiée. Après avoir établi l'expression du critère marginal et montré la relation très simple qui lie ce critère au critère joint, nous avons montré les bonnes propriétés asymptotiques de cet estimateur : l'erreur sur l'estimée des aberrations diminue lorsque le nombre de données traitées augmente. Afin de résoudre le problème du réglage de la valeur des hyperparamètres, une méthode d'estimation non supervisée de ceux-ci a été proposée : les hyperparamètres sont estimés conjointement avec les aberrations. La validation de cette méthode non supervisée a été effectuée sur simulations. Enfin une première validation sur données réelles, à fort flux, de ce nouvel estimateur a été faite avec succès.

Cette nouvelle méthode d'estimation des aberrations ne présente donc pas les défauts de l'estimation jointe ; elle est, de plus, particulièrement simple à implémenter puisque l'expression du critère marginal diffère très peu de celui du joint. Enfin, les très bonnes performances de cette méthode ont été démontrées, à fort flux, sur données simulées et expérimentales.

Chapitre 5

Joint/Marginal : comparaison des performances sur l'estimation des aberrations

5.1 Introduction

Après nous être intéressés de manière séparée aux estimateurs joint et marginal, et notamment en avoir montré les propriétés asymptotiques, nous proposons dans les chapitres 5 et 6 des comparaisons de leurs performances. Ce chapitre s'intéresse à la qualité de restauration des aberrations obtenues par chacune des méthodes, le suivant porte sur la restauration de l'objet.

Lorsque l'on souhaite effectuer une correction temps réel des dégradations subies par un système d'imagerie, les paramètres d'intérêt à estimer à partir des images sont les aberrations. Les techniques de compensation en temps réel des aberrations sont connues sous le nom d'optique adaptative (dans le cas de la correction d'aberrations introduites par la turbulence atmosphérique – aberrations rapidement variables) et d'optique active (dans le cas des défauts propres d'un instrument – aberrations d'évolution plus lentes).

Ce chapitre porte sur la comparaison de la qualité de restauration des aberrations obtenues par estimation marginale et par la méthode jointe. Dans le souci d'une utilisation réaliste de chacune de ces méthodes, dans le cas de l'estimation marginale, les hyperparamètres sont réglés de façon non supervisée et la méthode jointe est utilisée sans terme de régularisation spécifique sur l'objet. Ce dernier choix est dicté par les constatations qui ont été faites dans le paragraphe 2.4.2 : le réglage des hyperparamètres est difficile puisqu'il doit être fait nécessairement empiriquement; de plus à fort flux, nous avons constaté que l'estimation des aberrations est de meilleure qualité dans le cas non régularisé (la présence de la régularisation objet dégradant les aberrations estimées). Ces comparaisons sont effectuées dans un premier temps sur données simulées. Cette étude est menée en fonction de l'objet observé, du niveau de flux et du nombre de données traitées. Le chapitre se termine par une comparaison des performances des deux méthodes sur des données expérimentales : un seul objet est étudié mais les résultats de différents niveaux de flux sont présentés.

5.2 Sur données simulées

5.2.1 Conditions de simulation et d'estimation

Les conditions de simulation sont identiques au paragraphe 2.3.2.1. Les coefficients de a_4 à a_{21} sont reconstruits à l'aide de chacun des estimateurs. Les résultats, présentés figure 5.1, 5.2 et 5.3 proviennent de 50 estimations de ces coefficients obtenus par le traitement de 50 paires d'images bruitées. Tout comme dans les chapitres 2 et 4, pour les tailles image 32×32 et 64×64 pixels, les courbes sont le résultat du moyennage des valeurs obtenues pour toutes les sous-images de 32×32 pixels et 64×64 pixels contenues dans l'image de 128×128 pixels.

5.2.2 Résultats de la comparaison

5.2.2.1 Objet vue aérienne

Les courbes présentées sur la figure 5.1 correspondent à l'objet vue aérienne (représenté page 38).

Erreur La comparaison de l'erreur d'estimation associée à chaque méthode montre qu'excepté pour les images 32×32 pixels dont le bruit est inférieur à 4%, l'estimateur marginal (MV) donne toujours de meilleurs résultats que l'estimateur joint (MVG). Lorsque la taille image augmente, on observe une diminution de l'écart entre les estimées des deux méthodes. Au contraire, cet écart augmente avec le niveau de bruit sauf pour la fin de la courbe des images 32×32 pixels où on observe un resserrement à partir de 15% de bruit.

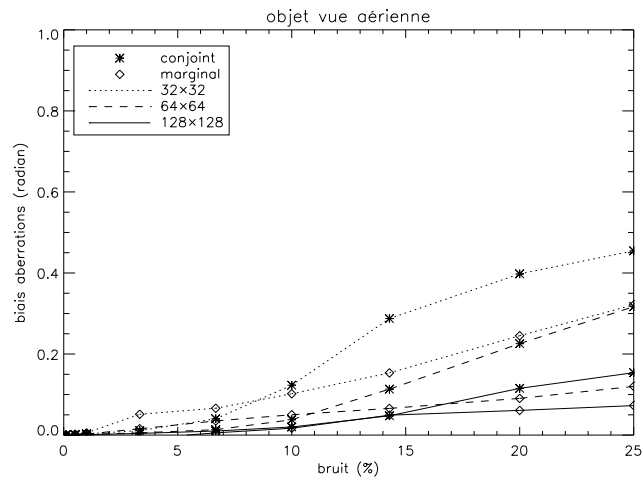
Biais On peut distinguer deux zones : à faible bruit, le biais de l'estimée par MVG est plus faible que celui du MV; à partir d'une certaine valeur du bruit (la limite dépend de la taille image), on observe une inversion : l'estimateur marginal devient moins biaisé que le joint. Dans cette deuxième zone, l'amplitude de l'écart augmente avec le niveau de bruit, alors que celui-ci diminue lorsque la taille image augmente.

Variance Les courbes liées à la variance des estimateurs montrent que la dispersion des estimées aberrations obtenues par MV est plus faible que celle des estimées MVG et ce quel que soit le niveau bruit et la taille image (on observe toutefois une légère inversion de cette comparaison dans le cas des images 32×32 pixels bruités avec un niveau de bruit inférieur à 4%).

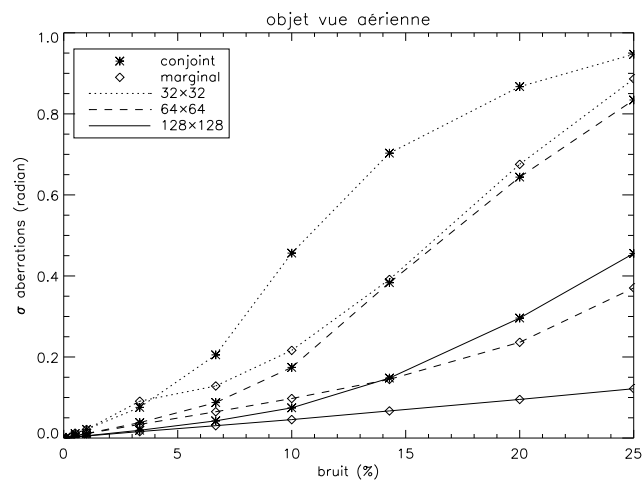
5.2.2.2 Objet maison

Les résultats correspondant à l'objet maison (présenté page 38) sont regroupés sur la figure 5.2.

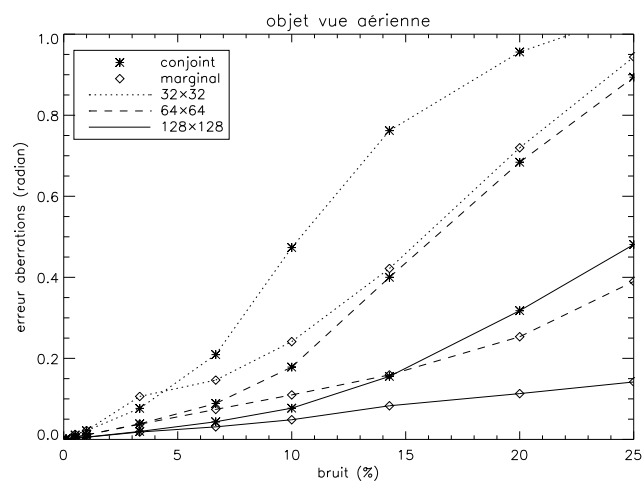
Erreur A faible niveau de bruit, l'erreur sur l'estimée des aberrations obtenue avec la méthode jointe est plus faible que celle obtenue avec la méthode marginale. Puis dans une zone intermédiaire de bruit (ou jusqu'à 25 % pour les images 64×64 pixels et 128×128 pixels), c'est l'estimateur marginal qui fournit les meilleurs résultats. Enfin, à fort bruit, dans le cas des images les plus petites, l'estimateur joint redevient meilleur. L'écart entre



- Biais -



- Écart type -



- Erreur -

FIG. 5.1 – Comparaison des résultats obtenus avec la méthode jointe et avec la méthode marginale - Scène vue aérienne.

les estimées de chaque méthode augmente avec la taille de l'image traitée mais reste assez stable lorsque le niveau de bruit évolue.

Biais Quels que soient le niveau de bruit et la taille image, l'estimateur marginal présente un biais plus fort que l'estimateur joint. L'écart entre les deux courbes augmente avec la taille de l'image traitée et reste à peu près stable lorsque le bruit augmente.

Variance Au contraire du biais, les courbes concernant la variance montre une supériorité de l'estimateur marginal (sauf tout comme l'erreur, pour les images 32×32 pixels à fort bruit) : celui-ci conduit à une dispersion des résultats plus faible que dans le cas de la méthode jointe. Cette supériorité est d'autant plus marquée que la taille image est grande et que le niveau de bruit augmente.

5.2.2.3 Objet entrepôt

La figure 5.3 rassemble les résultats concernant le dernier objet étudié : l'objet entrepôt (voir page 38).

Erreur On distingue trois zones : à faible bruit, l'erreur d'estimation associée à la méthode jointe est plus faible que celle associée à la méthode marginale, puis pour des bruits « moyens », l'estimateur marginal devient le plus performant et enfin lorsque le niveau de bruit est important, l'estimateur joint fournit à nouveau les meilleures estimées. On constate que pour cet objet, l'écart entre les courbes du joint et du marginal est assez faible.

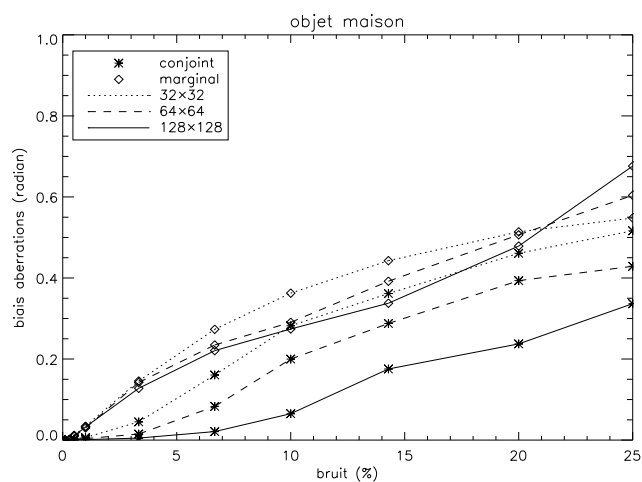
Biais La méthode marginale fournit les estimées les plus biaisées quelle que soit la taille image et le niveau de bruit. L'écart entre les courbes est important pour les images 64×64 pixels et 128×128 pixels et reste à peu près constant quel que soit le bruit ; il est plus faible pour les images 32×32 pixels surtout à fort bruit.

Variance La variance de l'estimateur marginal est plus faible que celle du joint à faible bruit puis devient plus importante. Le niveau de bruit où à lieu le croisement des deux courbes dépend de la taille image : plus celle-ci est faible plus il intervient tôt : 4,5% pour 32×32 pixels, 8,5% pour 64×64 pixels et au delà de 25% pour 128×128 pixels.

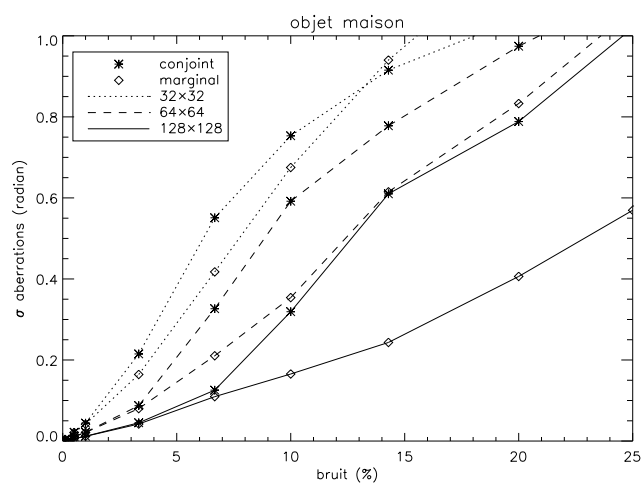
5.2.3 En résumé

L'apport de l'estimateur marginal sur la restauration des aberrations est très important dans le cas de l'objet vue aérienne puisqu'il permet une bien meilleure estimation quels que soient la taille image et le niveau de bruit. Pour l'objet maison, son intérêt apparaît lorsque le niveau de bruit est élevé et l'image supérieure à 32×32 pixels. L'objet entrepôt est un assez mauvais candidat pour la méthode marginale où ses performances sont très proches de celle de la méthode jointe.

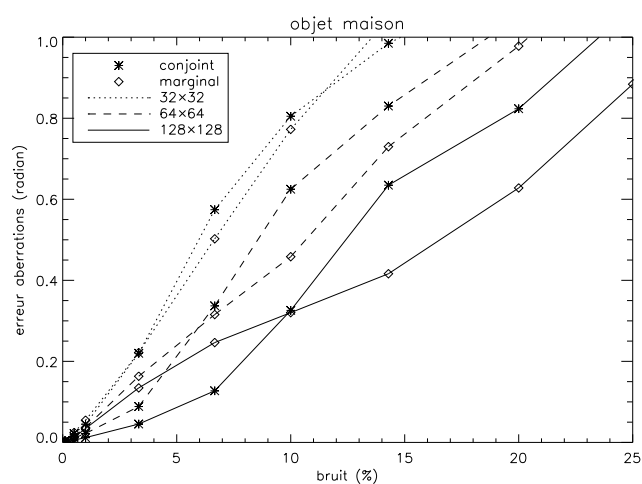
La synthèse de l'étude comparative menée sur les trois objets précédents conduit aux constatations suivantes : tant que l'erreur d'estimation est grossièrement inférieure à 80%, on distingue deux zones : une, correspondant aux faibles bruits, est plus favorable à une estimation par MVG (joint) et l'autre, liées aux bruits « moyens » est dominée par l'estimateur MV (marginal).



- Biais -

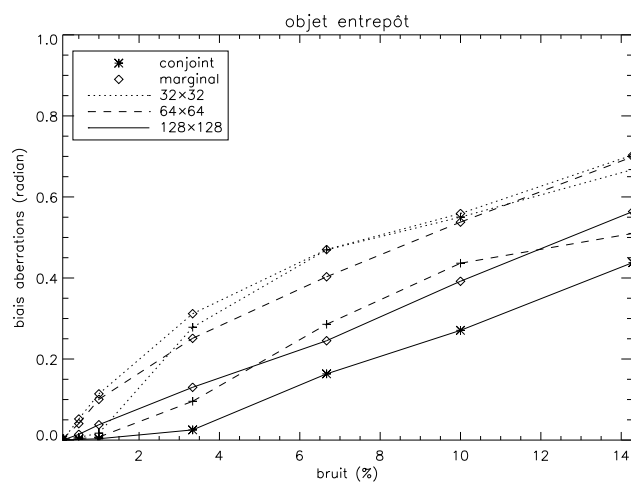


- Écart type -

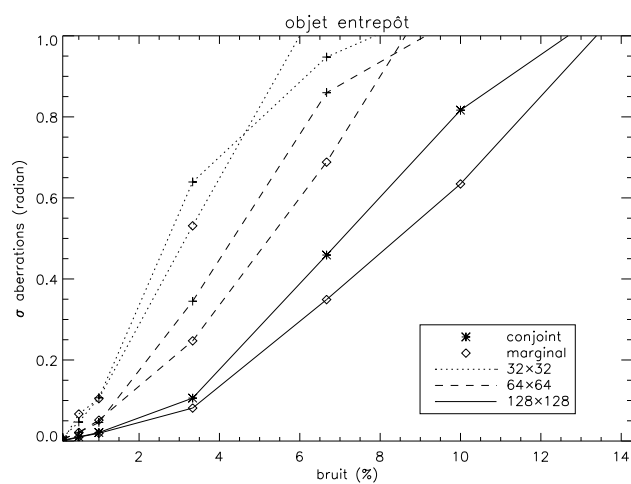


- Erreur -

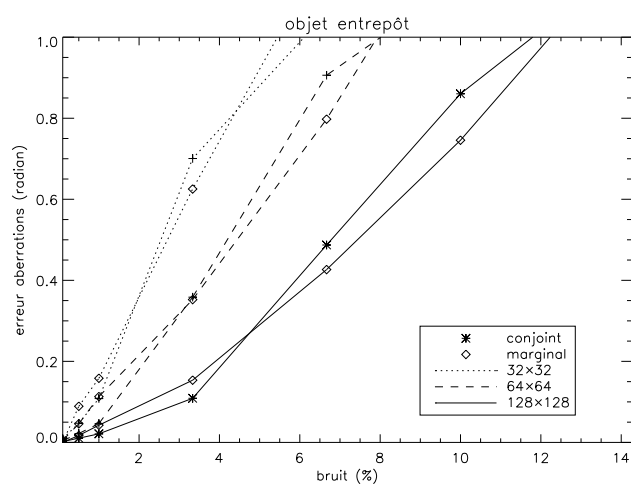
FIG. 5.2 – Comparaison des résultats obtenus avec la méthode jointe et avec la méthode marginale - Scène maison.



- Biais -



- Écart type -



- Erreur -

FIG. 5.3 – Comparaison des résultats obtenus avec la méthode jointe et avec la méthode marginale - Scène entrepôt.

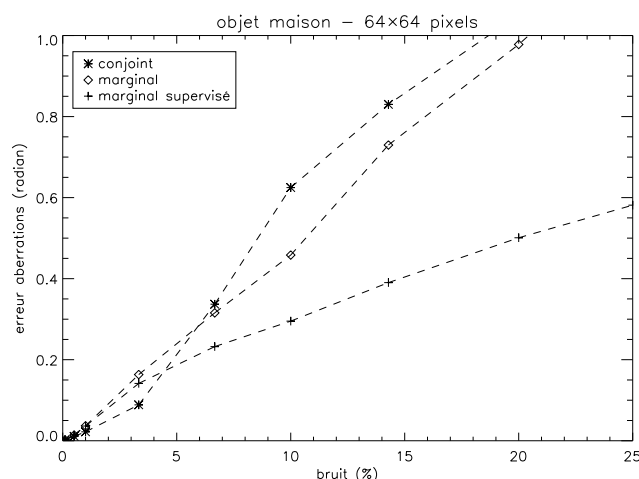


FIG. 5.4 – Comparaison de l'erreur (rms) en radian sur l'estimée des aberrations, obtenue par la méthode conjointe (étoile), marginale non supervisée (losange) et marginale supervisée (croix) sur une image de taille 64×64 pixels – objet maison.

5.2.4 Interprétations des résultats

Faibles bruits A faible bruit, la meilleure estimation est obtenue par l'estimateur joint pour les objets maison et entrepôt quelle que soit la taille image et pour l'objet vue aérienne pour les images 32×32 pixels (pour 64×64 pixels et 128×128 pixels, la qualité des estimées est identique). La séparation de l'erreur en ses deux composantes que sont le biais et la variance montrent, pour tous les objets, que l'estimateur marginal présente une variance plus faible mais un biais bien plus important que l'estimateur joint.

Afin de savoir si ce biais provient d'une erreur d'estimation des hyperparamètres du modèle de DSP associé à l'objet, nous avons ajouté, dans le cas de l'objet maison, sur la figure 5.4, en plus des courbes correspondant à l'estimateur joint et marginal non supervisé, celle du marginal supervisé. Nous avons vérifié que les hyperparamètres « vrais » utilisés dans la version supervisée, correspondent bien, à faible bruit, aux hyperparamètres meilleurs réglages (au sens de la minimisation de l'EQM entre aberrations vraies et estimées). On constate que même l'estimée fournie par la version supervisée du marginal présente une erreur plus importante que celle du conjoint. Cet écart ne provient donc pas d'une erreur liée au mauvais ajustement du modèle de DSP objet utilisé. Afin de construire l'estimateur marginal, nous avons intégré l'objet hors du problème et pour ce faire nous avons supposé qu'il provenait de la réalisation d'un processus gaussien. Ainsi l'estimateur MV (marginal) au contraire de l'estimateur MVG (joint) contient une hypothèse statistique sur l'objet. Dans le cas de conditions d'estimation favorables (bruit faible et simultanément nombre de données important), l'utilisation d'une hypothèse de gaussianité sur l'objet a tendance à dégrader l'estimation des aberrations en imposant un modèle à l'objet qui est forcément incorrect (les cas étudiés correspondant à des objets réels) et qui biaise la reconstruction. Ce biais sera d'autant plus important que l'objet observé est loin de correspondre au modèle gaussien qui lui est associé. L'erreur d'estimation associée à la méthode marginale est supérieure à celle de la méthode jointe car le gain apporté en variance par l'estimation MV, ne compense pas le biais important qui lui est associé.

bruits « moyens » Quel que soit l'objet, on observe un domaine de bruit (correspondant à des bruits « moyens » à « forts ») où l'estimation marginale est de meilleure qualité que l'estimation conjointe. La différence de comportement entre les trois objets étudiés portent sur l'importance de l'écart entre marginal et conjoint : il est grand pour l'objet vue aérienne, marqué pour les images 64×64 pixels, 128×128 pixels de l'objet maison et faible pour l'objet entrepôt.

A présent, le niveau de bruit ayant augmenté, l'apport de la marginalisation est significatif. Les conditions d'estimation sont moins favorables et la reconstruction conjointe d'un grand nombre de paramètres – pixels objet et aberrations – est difficile. Au contraire, la méthode marginale s'étant affranchie de la restauration des pixels objet, c'est-à-dire de l'estimation d'un grand nombre d'inconnues, permet une reconstruction des inconnues aberrations de meilleure qualité puisque non dégradées par l'estimation conjointe d'un objet de mauvaise qualité.

Nous avons complété la courbe 5.4 par les deux courbes équivalentes correspondant aux objets entrepôt (figure 5.5) et vue aérienne (figure 5.6) afin d'étudier la qualité d'estimation des hyperparamètres. L'écart entre les courbes du marginal non supervisé et supervisé est important pour les objets maison et entrepôt (excepté à très faible bruit) et augmente rapidement avec le niveau de bruit, tandis qu'il reste beaucoup plus faible pour l'objet vue aérienne. La comparaison des résultats obtenus par la méthode conjointe et par la estimation marginale supervisée montre que l'apport de la marginalisation est ici très significatif. Le faible écart que l'on constate pour l'objet maison à 32×32 pixels et pour l'objet entrepôt quel que soit la taille image n'est pas du à l'hypothèse de gaussianité utilisé dans le méthode marginale (comme à faibles bruits) mais à un mauvais ajustement des hyperparamètres. La qualité du réglage des hyperparamètres dépend, tout comme la reconstruction des aberrations, du niveau de bruit, du nombre de données traitées mais aussi de l'objet observé. Comme il a déjà été mentionné précédemment (paragraphes 2.3.2.2 et 4.6), la qualité de restauration dépend beaucoup de l'objet observé. Parmi les trois objets que nous avons étudiés, l'objet vue aérienne est de loin l'objet le plus favorable, au contraire de l'objet entrepôt qui conduit à des estimées de qualité médiocre. L'apport de la méthode marginale non supervisée est faible dans le cas de l'objet entrepôt car comme le montre la courbe 5.5 et comme il avait déjà été signalé dans le paragraphe 4.8.3, les hyperparamètres sont mal estimés.

Forts bruits A fort bruit, l'estimateur conjoint (MVG) redevient supérieur à l'estimateur marginal (MV). Il faut cependant noter que ce comportement se rencontre pour des conditions d'estimation telles que l'erreur de reconstruction sur les aberrations est supérieure à 80% et où donc aucun des estimateurs n'est, en pratique, utilisable. Le niveau de bruit où a lieu l'inversion dépend de l'objet observé et de la taille image : il est d'autant plus faible que l'objet est peu favorable à l'estimation et que la taille image est faible. Ainsi dans les conditions très défavorables d'estimation (bruit fort et faible nombre de données), l'estimateur MVG se révèle plus « performant » que l'estimateur MV. Des résultats concernant la supériorité de l'estimateur MVG par rapport au MV dans des conditions d'estimation difficiles (à faible nombre de données) ont été mise en évidence dans l'article (Champagnat et Idier, 1995).

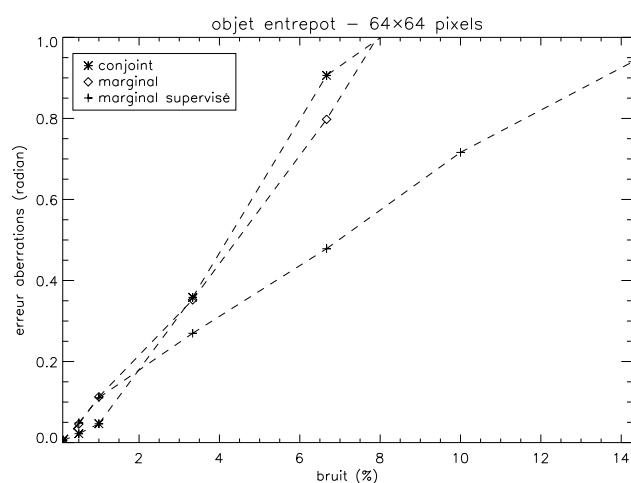


FIG. 5.5 – Comparaison de l'erreur (rms) en radian sur l'estimée des aberrations, obtenue par la méthode conjointe (étoile), marginale non supervisée (losange) et marginale supervisée (croix) sur une image de taille 64×64 pixels – objet entrepôt.

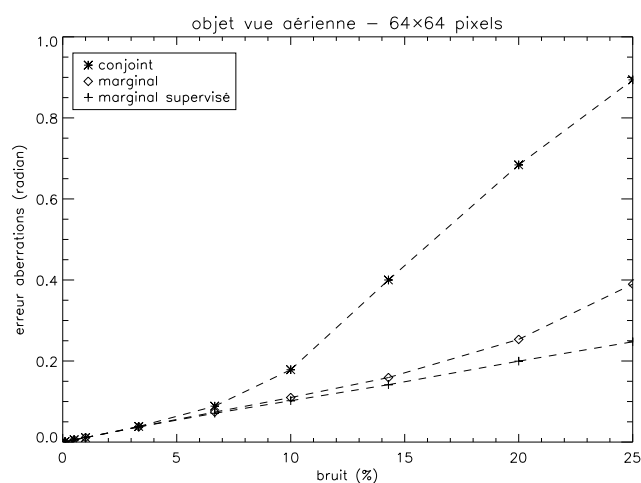


FIG. 5.6 – Comparaison de l'erreur (rms) en radian sur l'estimée des aberrations, obtenue par la méthode conjointe (étoile), marginale non supervisée (losange) et marginale supervisée (croix) sur une image de taille 64×64 pixels – objet vue aérienne.

5.3 Sur données expérimentales

5.3.1 Acquisition des données

Le montage expérimental présenté et utilisé dans le chapitre précédent (voir le schéma 4.7) est à nouveau employé pour acquérir les données expérimentales. Des aberrations parfaitement calibrées peuvent être introduites sur le trajet optique à l'aide de la lame à faces parallèles; les valeurs des aberrations introduites en fonction de l'angle que fait la lame avec l'axe optique sont données par les relations (4.33) et (4.34). Dans la suite, on s'intéressera aux aberrations introduites par la lame formant un angle de 45° avec l'axe optique : les aberrations introduites sont $a_4 = 0,383$ radian et $a_6 = -0,442$ radian. L'objet observé est une diapositive représentant la scène « vue aérienne »¹. Cependant l'objet « expérimental » vue aérienne est échantillonné plus finement que celui utilisé en simulation. Les images, comme dans le chapitre 4, sont échantillonnées en respectant le critère de Shannon.

On étudie le comportement des deux estimateurs en fonction du flux lumineux. Pour le faire varier, on modifie le temps de pose des images et afin de pouvoir atteindre des niveaux de flux faibles, on ajoute dans le montage expérimental une densité devant la caméra. Le RSB est calculé, en prenant le rapport du nombre de photons moyen dans l'image (le gain de la caméra étant connu) et de l'écart type du bruit. Ce dernier est égal à la racine carré de la somme de la variance du bruit photonique (qui est égal au nombre de photons par pixel) et de la variance du bruit électronique (13^2 électrons²).

5.3.2 Calibration des aberrations du banc à fort flux

On souhaite étudier la variance ainsi que le biais des coefficients de Zernike estimés. Afin d'évaluer l'erreur totale sur l'estimation des aberrations introduites par la lame à 45° , il faut connaître les aberrations associées à l'ensemble formé par le banc, la lame à angle nul et la densité introduite devant la caméra. Pour cela, nous avons effectué une première estimation par diversité de phase, à fort flux, des aberrations de cet ensemble. Par la suite, le terme de biais de l'erreur totale est estimé en soustrayant à la valeur des estimées des aberrations obtenues avec la lame à 45° , d'une part la valeur théorique des aberrations introduites par la rotation de la lame ($a_4 = 0,383$ radian et $a_6 = -0,442$ radian) et d'autre part la valeur des aberrations associées à l'ensemble banc+densité+lame à angle nul estimée à fort flux.

L'estimation des aberrations (aussi bien pour la calibration du banc que pour les mesures de la lame à 45°) est effectuée de la même façon que dans le paragraphe 4.9.3. Les données sont constituées d'une image plan focal et d'une image défocalisée (défocalisation de 2π crête-à-crête) acquises successivement. Le niveau de bruit associé aux images est de 0,5% (pour atteindre ce faible niveau de bruit, les images sont obtenues par accumulation de dix enregistrements). L'estimation est effectuée d'une part à l'aide de l'estimateur joint sans terme de régularisation sur l'objet et d'autre part par l'estimateur marginal avec réglage non-supervisé des hyperparamètres. Une zone de 64×64 pixels est extraite de chaque image enregistrée et utilisée pour l'estimation (une bande de garde de largeur identique à celle de l'image est employée pour traiter les effets de bords dus à l'étendu infinie de l'objet). Les estimations sont effectuées sur vingt-cinq paires d'images correspondant aux différentes combinaisons possibles des cinq paires d'images focalisées et défocalisées enregistrées. Les

1. Afin d'utiliser le même objet que celui des simulations, cette diapositive a été obtenue par photographie d'un écran de TX

valeurs moyennes des aberrations estimées sont rassemblées dans le tableau 5.1². L'écart

a_i (radian)	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
joint	0,045	0,145	-0,020	0,022	0,001	-0,020	0,046	0,004	-0,028
marginal	0,046	0,152	-0,020	0,011	0,006	-0,024	0,045	0,007	-0,028
a_i (radian)	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}
joint	0,014	0,005	0,015	-0,003	0,025	0,006	0,015	-0,006	0,006
marginal	0,014	0,003	0,017	0,005	0,007	0,004	0,009	-0,006	0,005

TAB. 5.1 – *Calibration à fort flux : valeur moyenne des coefficients de Zernike estimés par la méthode jointe et par la méthode marginale sur la lame à angle nul.*

entre les valeurs estimées par chacune des méthodes est très faibles : inférieur à $\lambda/250$. On utilisera dans la suite la moyenne des valeurs moyennes estimées par chacune des méthodes.

5.3.3 Résultats : évolution en fonction du niveau du flux

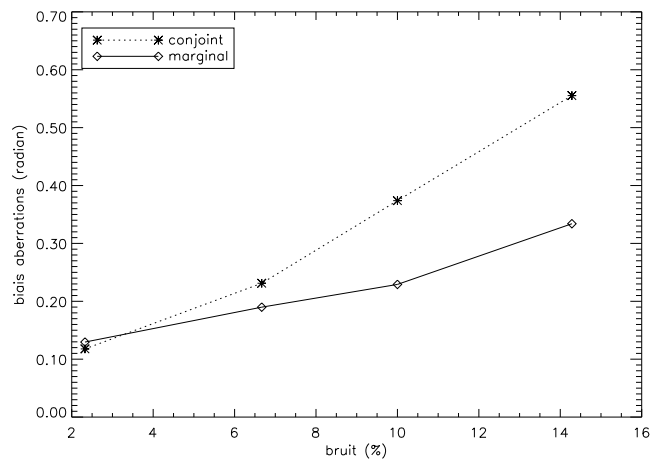
On fait tourner la lame de 45 degrés par rapport à la position « angle nul » (calibrée précédemment). Quatre séries d'images focalisées et défocalisées sont alors enregistrées avec des temps de pose différents tel que le niveau de bruit de chacune des séries soit respectivement de 2%, 7%, 10% et 14%.

Les résultats sont rassemblés sur la figure 5.7. Ils correspondent au moyennage des résultats obtenus à partir de vingt-cinq paires d'images. Les courbes montrent que l'estimation des aberrations obtenues avec la méthode marginale est de meilleure qualité que celle fournie par l'estimateur joint pour tous les niveaux de bruit excepté pour 2% où les deux résultats d'estimation sont quasiment identiques. La courbe de comparaison portant sur le biais montre un biais identique pour les deux estimateurs à faible niveau de bruit puis un écart qui augmente avec le niveau de bruit. La courbe présentant la variance des résultats indique une variabilité des résultats plus faibles dans le cas du marginal mais l'écart entre les deux méthodes est moins marqué que celui du biais. On retrouve sur la courbe expérimentale de l'évolution de l'erreur d'estimation des aberrations le comportement qui avait été observé sur les données simulées. On remarque toutefois que l'écart entre les variances associées à chaque méthode est moins favorable pour la méthode marginale que dans le cas simulé mais qu'au contraire l'écart entre les biais est plus marqué que dans le cas simulé. Des différences entre les conditions de simulation et celles expérimentales peuvent être mises en avant pour expliquer ces variations. Dans le cas des données expérimentales, une bande de garde a été utilisée lors de l'estimation pour gérer les problèmes liés à l'étendue infinie de l'objet. De plus, le niveau des aberrations à mesurer est plus faible que dans le cas des simulations (ici $\lambda/11$ rms), l'objet observé est échantillonné plus finement que la version utilisée en simulation et enfin nous n'avons pas pu utiliser la même zone de l'image qu'en simulation.

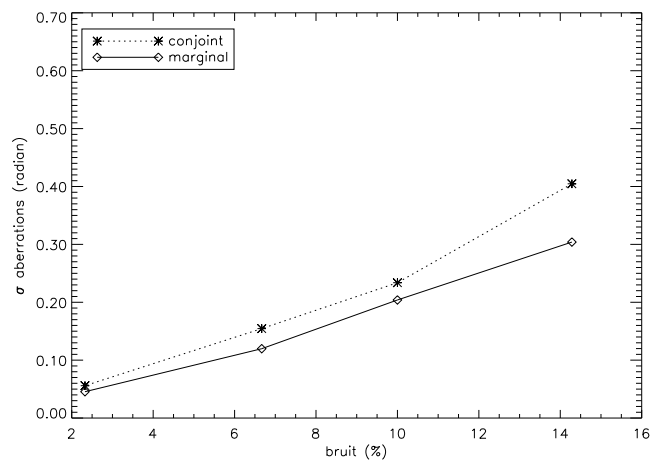
5.4 Conclusion

Après avoir étudié séparément les deux estimateurs joint et marginal, ce chapitre a présenté une comparaison des performances de chacune de ces méthodes sur l'estimation

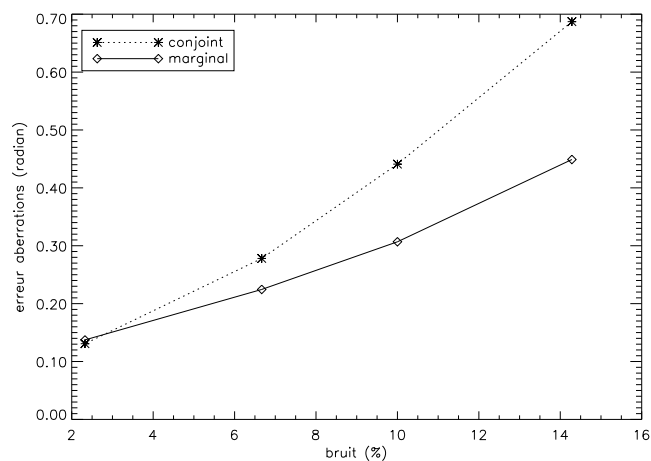
2. Remarque : les mesures utilisées ici n'ont pas été enregistrées en même temps que celles présentées à la fin du chapitre 4. Il s'est écoulé à peu près un an entre les deux séries d'acquisitions.



- Biais -



- Écart type -



- Erreur -

FIG. 5.7 – Comparaison des résultats obtenus avec la méthode jointe et avec la méthode marginale à partir de données expérimentales.

des aberrations. Cette étude basée sur trois objets peut être résumée ainsi. La variance de l'estimateur marginal est presque toujours plus faible que celle de l'estimateur joint. La comparaison du terme de biais dépend quant à lui fortement de l'objet observé et du niveau de bruit. A faible bruit, l'estimation marginale des aberrations donne des résultats identiques ou moins bons que l'estimateur joint sur la reconstruction des aberrations. L'importance du manque de performance de la méthode marginale dépend de l'objet observé et vraisemblablement de son adéquation au modèle gaussien qui lui a été associé lors de la construction du critère marginal. Passé les très faibles bruits, il apparaît un large domaine dans lequel l'apport de l'estimation marginale est significatif. L'amplitude de l'écart entre les estimées fournies par les deux méthodes dépend de l'objet observé. Une validation sur données expérimentales des bonnes performances de la méthode marginale en régime de forts bruits a été effectuée avec succès.

La présentation de la méthode marginale ainsi que sa comparaison avec la méthode jointe a été à la base de publications dans des congrès (Blanc et al., 1999; Blanc et al., 2000a; Blanc et al., 2000b). Les articles (Blanc et al., 1999) et (Blanc et al., 2000a) sont proposés dans les annexes H et I.

Chapitre 6

Méthodes possibles pour la restauration de l'objet

6.1 Introduction

Le chapitre 5 a proposé une comparaison des résultats de l'estimation des aberrations obtenues par l'estimateur joint et par l'estimateur marginal. A présent, cette comparaison va être menée vis-à-vis de l'estimation de l'objet. L'objet est la grandeur d'intérêt lorsque la diversité de phase est utilisée pour faire du traitement d'image *a posteriori*. L'application principalement développée en diversité de phase est l'imagerie solaire (Löfdahl et Scharmer, 1994; Seldin et Paxman, 1994).

L'estimateur joint, comme son nom l'indique, propose une estimation conjointe des aberrations et de l'objet. Le chapitre 2 a mis en évidence les difficultés soulevées par cette méthode et notamment sur la restauration de l'objet. Ces difficultés nous ont amenées à ne pas utiliser de terme de régularisation sur l'objet lorsque la diversité de phase est utilisée comme ASO. Si cette absence de régularisation objet ne détériore pas, à fort flux, l'estimation des aberrations, elle s'avère très pénalisante pour la restauration de l'objet, comme nous l'avions déjà évoqué dans ce même chapitre et comme nous allons le mettre en évidence ici. La méthode marginale, quant à elle, s'est affranchie de la restauration de l'objet. On montre qu'elle permet à partir des aberrations et des hyperparamètres estimés de restaurer également de façon simple et rapide l'objet. Enfin, une méthode de restauration de l'objet qui utilise à la fois une estimation marginale des hyperparamètres et une estimation des aberrations par la méthode jointe est proposée (méthode appelée estimation « mixte »).

6.2 Restauration de l'objet avec l'estimateur joint

L'estimateur joint, présenté dans le chapitre 2, est basé sur une estimation au sens du MAPJ des aberrations et de l'objet. Ce chapitre a mis en évidence le problème du réglage des hyperparamètres liés au terme de régularisation objet : ils doivent être ajustés nécessairement empiriquement et ce réglage est d'autant plus difficile que l'on s'intéresse à des objets étendus. On a montré qu'il est impossible de restaurer optimalement l'objet et les aberrations en conjoint (paragraphe 2.4). En particulier, si on choisit un paramètre de régularisation proche de zéro (voir page 45), les chapitres 2 et 5 ont montré que cette régularisation « quasi-nulle » sur l'objet permet, à fort flux, une bonne estimation des aberrations même si l'objet conjointement estimé est très dégradé. Au chapitre 2, un premier

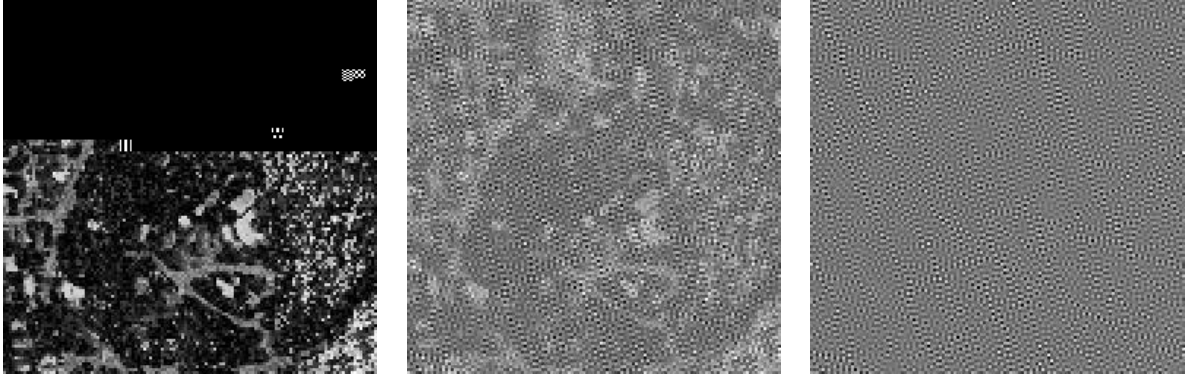


FIG. 6.1 – L'objet « vue aérienne » estimé par la méthode jointe. A gauche, l'objet vrai; au centre l'objet reconstruit pour un bruit de 1% , l'erreur (rms) entre l'objet vrai et l'objet estimé est de 115% ; à droite, l'objet reconstruit pour un bruit de 14% , l'erreur entre l'objet vrai et l'objet estimé est de 1500% .

exemple de restauration objet à régularisation « quasi-nulle » a été présenté. Ce résultat est rappelé ici, figure 6.1, (l'objet observé est la « vue aérienne », la taille image est de 64×64 pixels, les aberrations utilisées pour générer la phase aberrante sont celles données dans la table 3.1 et le bruit est de 1%) complétée par un exemple de restauration obtenue dans les même conditions mais pour un bruit de 14% . La qualité des objets ainsi restaurés est très mauvaise même à fort flux. Sans terme de régularisation explicite sur l'objet, il y a amplification du bruit (inévitablement présent dans les images) lors de la minimisation du critère. L'espoir de restaurer correctement l'objet passe forcément par l'utilisation d'un terme de régularisation spécifique mais pose alors les problèmes cités précédemment.

6.3 Restauration de l'objet avec l'estimateur marginal

6.3.1 Principe

L'estimateur marginal a été obtenu par intégration de l'objet hors du problème afin de construire un estimateur MAP des aberrations seules, qui comme il a été montré dans le chapitre 4 a de bonnes propriétés asymptotiques, donne de meilleurs résultats sur l'estimation des aberrations à faible flux que l'estimateur joint et permet une estimation des hyperparamètres objet (paramètres du modèle de DSP) et bruit (variance). Si l'on prend pour les aberrations \mathbf{a} et les hyperparamètres $\boldsymbol{\theta}$, les valeurs estimées par la méthode marginale $\hat{\mathbf{a}}_{\text{marg}}$ et $\hat{\boldsymbol{\theta}}_{\text{marg}}$, l'estimation de l'objet $\hat{\mathbf{o}}$ est alors très facile car la RI et les hyperparamètres sont connus. En particulier, cette estimation peut être faite par MAP :

$$\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}}) = \arg \max_{\mathbf{o}} f(\mathbf{o} | \mathbf{i}_f, \mathbf{i}_d, \hat{\mathbf{a}}_{\text{marg}}; \hat{\boldsymbol{\theta}}_{\text{marg}}) = \arg \max_{\mathbf{o}} f(\mathbf{i}_f, \mathbf{i}_d, \hat{\mathbf{a}}_{\text{marg}} | \mathbf{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) f(\mathbf{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) \quad (6.1)$$

En prenant l'opposé du logarithme du produit $f(\mathbf{i}_f, \mathbf{i}_d, \hat{\mathbf{a}}_{\text{marg}} | \mathbf{o}; \hat{\boldsymbol{\theta}}_{\text{marg}}) f(\mathbf{o}; \hat{\boldsymbol{\theta}}_{\text{marg}})$, on obtient le critère à minimiser :

$$J_{\text{MAP}}(\mathbf{o}) \propto \frac{\|\mathbf{i}_f - H_f(\hat{\mathbf{a}}_{\text{marg}})\mathbf{o}\|^2}{\sigma^2} + \frac{\|\mathbf{i}_d - H_d(\hat{\mathbf{a}}_{\text{marg}})\mathbf{o}\|^2}{\sigma^2} + (\mathbf{o} - \mathbf{o}_m)^t R_o^{-1} (\mathbf{o} - \mathbf{o}_m) \quad (6.2)$$

Lorsque l'on calcule le gradient de ce critère par rapport à l'objet, on trouve la solution analytique suivante :

$$\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}}) = (H_f^t H_f + H_d^t H_d + \sigma^2 R_o^{-1})^{-1} (H_f^t \mathbf{i}_f + H_d^t \mathbf{i}_d + \sigma^2 R_o^{-1} \mathbf{o}_m) \quad (6.3)$$

Dans le cas d'objets compacts, le critère J_{MAP} ainsi que la solution analytique $\hat{\mathbf{o}}_{\text{MAP}}$ peuvent s'écrire dans le domaine de Fourier. Dans le cas d'objets très étendus, le calcul du critère dans le domaine de Fourier est possible moyennant l'ajout d'une bande de garde (voir le paragraphe 4.9.1), la minimisation étant effectuée quant à elle itérativement.

On propose donc ici d'estimer l'objet en deux étapes : cette démarche nous permet d'effectuer une estimation des aberrations puis de l'objet par MAP et ainsi de bénéficier des bonnes propriétés de cet estimateur : lorsque le nombre de données tend vers l'infini, $\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}})$ tend vers $\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{vrai}}, \hat{\boldsymbol{\theta}}_{\text{vrai}})$.

De plus, l'estimation de l'objet ne nécessite en fait pas d'être effectuée dans une deuxième étape : aussi bien pour le critère marginal $J_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta})$ (voir équation 4.18) que pour sa version alternée $J_{\text{MAP}}^{\text{alt}}(\mathbf{a}, \boldsymbol{\theta})$ (voir équation 4.31), $\hat{\mathbf{o}}_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta})$ est calculé à chaque étape de la minimisation du critère. En effet, à l'étape l de la minimisation $\hat{\mathbf{o}}_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta})$ est calculé comme sous produit de l'estimation $J_{\text{MAP}}(\mathbf{a}, \boldsymbol{\theta})$: on a alors $\hat{\mathbf{o}}_{\text{MAP}}^l(\mathbf{a}^l, \boldsymbol{\theta}^l)$. En fin de minimisation, on a donc à la fois $(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}})_{\text{MAP}}$ et $\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}})$. L'estimation de $\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{marg}}, \hat{\boldsymbol{\theta}}_{\text{marg}})$ ne nécessite donc pas de calcul supplémentaire : il est fourni par la méthode marginale.

6.3.2 Exemples de restaurations

Les figures 6.2, 6.3 et 6.4 présentent des résultats de reconstruction de l'objet MAP (les conditions de simulations sont identiques au paragraphe précédent). Les trois objets étudiés précédemment en simulation sont présentés pour deux niveaux de flux différents. L'erreur entre l'objet vrai et l'objet estimé pour l'objet « vue aérienne » est de 22% pour un bruit de 1% et de 38% pour un bruit de 14%, pour l'objet « entrepôt », l'erreur est 6,7% pour un bruit de 1% et de 14% pour un bruit de 14% et pour l'objet « maison », l'erreur est 7,6% pour un bruit de 1% et de 17% pour un bruit de 14%. Les estimées ainsi obtenus sont de bonnes qualités. Cette méthode fournit donc très simplement, même si c'est indirectement, une reconstruction de l'objet au sens du MAP.

6.3.3 Adéquation entre $\hat{\mathbf{o}}_{\text{optimal}}$ et $\hat{\mathbf{a}}_{\text{optimal}}$

La section 2.4 a mis en évidence que l'estimation par MAPJ de l'objet et des aberrations ne permet pas de reconstruire de façon optimale à la fois l'objet et les aberrations. On peut se demander ce qu'il en est pour la méthode marginale. Pour cela, on trace sur une même courbe, le résultat de l'estimation MAP des aberrations et de l'objet en fonction de la valeur de l'hyperparamètre global k . Les courbes associées aux trois objets sont regroupées sur la figure 6.5. Elles ont été obtenues pour un bruit de 1% et sur une image de taille 32×32 pixels. Contrairement à l'estimateur joint (voir pour comparaison les courbes 2.8), ici c'est la même valeur de l'hyperparamètre qui fournit la meilleure estimée des aberrations *et* de l'objet. La méthode marginale fournit donc une estimation optimale à la fois des aberrations *et* de l'objet.



FIG. 6.2 – Objet « vue aérienne » estimé par la méthode marginale. A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 22%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 38%.



FIG. 6.3 – Objet « entrepôt » estimé par la méthode marginale. A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 6,7%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 14%.



FIG. 6.4 – Objet « maison » estimé par la méthode marginale. A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 7,6%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 17%.

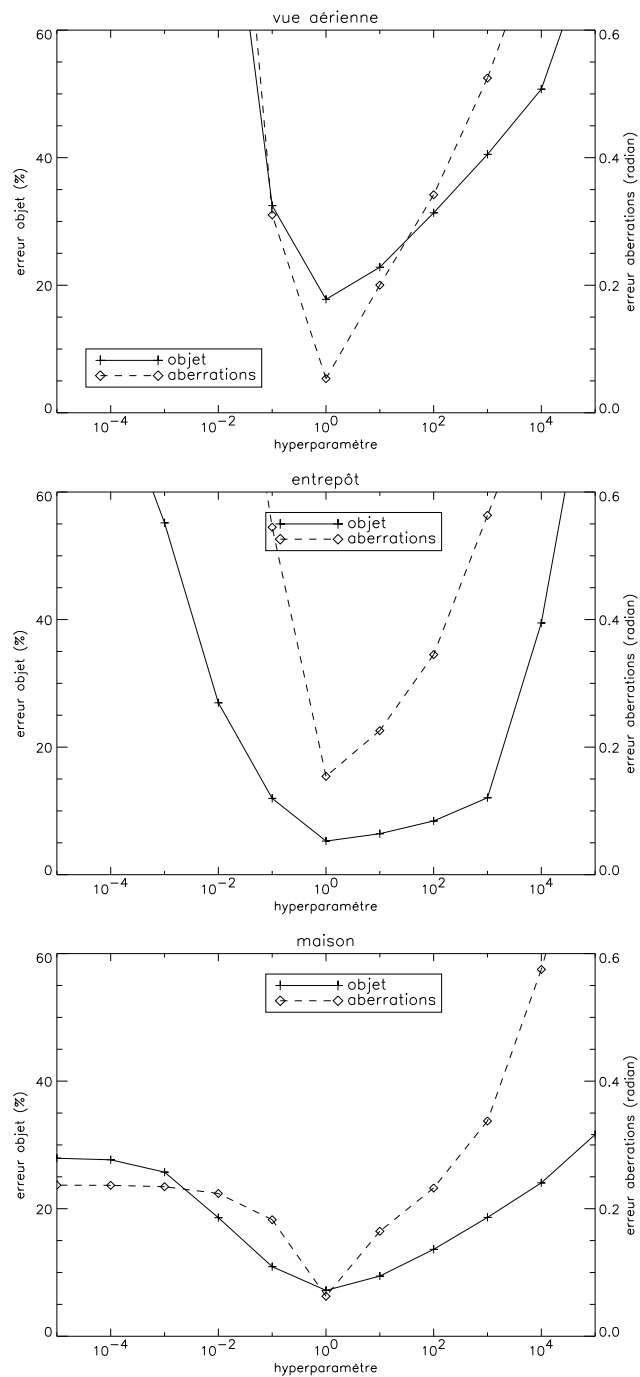


FIG. 6.5 – Influence de la valeur de l'hyperparamètre k sur l'estimation des aberrations et de l'objet obtenus à partir de l'estimateur marginal (image 32×32 pixels et bruit de 1%) : la même valeur de k fournit une estimation optimale des aberrations et de l'objet.

6.4 Restauration de l'objet par estimation « mixte »

6.4.1 Principe

La méthode conjointe, sans régularisation sur l'objet, fournit un objet restauré très dégradé mais conjointement des aberrations de très bonne qualité, à fort flux (et même pour certains objets meilleurs que le marginal). L'estimation des aberrations par la méthode conjointe nécessite un temps de calcul plus faible que dans le cas d'une restauration par la méthode marginale (ralentie par l'estimation des hyperparamètres). On peut donc penser à utiliser l'estimation jointe des aberrations (qui est rapide) puis d'effectuer une estimation marginale des hyperparamètres seuls à aberrations fixées; cette dernière estimation est rapide car très peu de paramètres sont estimés. On est alors ramené au paragraphe 6.3.1 à l'estimation de l'objet à RI et hyperparamètres connus, qu'on réalise par MAP. Cette méthode est appelée estimation « mixte ».

6.4.2 Les trois étapes

- Estimation des aberrations

Les aberrations sont estimées par MVG conjointement avec l'objet (voir chapitre 2) :

$$(\hat{\mathbf{o}}, \hat{\mathbf{a}})_{\text{MVG}} = \arg \max_{\mathbf{o}, \mathbf{a}} f(\mathbf{i}_f, \mathbf{i}_d | \mathbf{a}; \mathbf{o}, \theta) f(\mathbf{a}; \theta) \quad (6.4)$$

De cette estimation, seule $\hat{\mathbf{a}}_{\text{MVG}}$ est conservé.

- Estimation des hyperparamètres

L'estimation des hyperparamètres lié au bruit et à l'*a priori* objet est effectuée par Maximum de Vraisemblance. Pour ce faire la densité de probabilité jointe de l'objet \mathbf{o} , des images \mathbf{i}_f et \mathbf{i}_d et des aberrations \mathbf{a} est marginalisée par rapport à \mathbf{o} , tout comme dans le cas de l'estimateur marginal (la différence est que, dans ce cas, \mathbf{a} est fixé à sa valeur estimée par MVG) :

$$\hat{\theta} = \arg \max_{\theta} \int f(\mathbf{i}_f, \mathbf{i}_d, \mathbf{o}, \hat{\mathbf{a}}_{\text{MVG}}; \theta) d\mathbf{o} = \arg \max_{\theta} f(\mathbf{i}_f, \mathbf{i}_d, \hat{\mathbf{a}}_{\text{MVG}}; \theta) \quad (6.5)$$

L'expression du critère obtenu en prenant l'opposé du logarithme de la vraisemblance est identique à celui utilisé dans le paragraphe 4.9.1 pour l'estimation marginale des hyperparamètres et des aberrations. L'implémentation et la mise en œuvre pratique de cette estimation sont également identiques à ceux présentés dans le chapitre 4. Cette étape est très rapide puisque la minimisation est effectuée à aberrations fixées, seuls les hyperparamètres (c'est-à-dire trois inconnues) étant estimés.

- Estimation de l'objet MAP

A partir des aberrations $\hat{\mathbf{a}}_{\text{MVG}}$ et des hyperparamètres $\hat{\theta}$ estimés dans les deux étapes précédentes, l'objet est obtenu par l'estimateur MAP :

$$\hat{\mathbf{o}}_{\text{MAP}}(\hat{\mathbf{a}}_{\text{MVG}}, \hat{\theta}) = \arg \max_{\mathbf{o}} f(\mathbf{i}_f, \mathbf{i}_d, \hat{\mathbf{a}}_{\text{MVG}} | \mathbf{o}; \hat{\theta}) f(\mathbf{o}; \hat{\theta}) \quad (6.6)$$

Cette étape est identique à la reconstruction de l'objet à partir de l'estimateur marginal (voir paragraphe précédent) excepté qu'ici cette étape doit être explicitement effectuée.

6.4.3 Résultats

On présente le résultat des estimées objet obtenues par la méthode mixte sur les figures 6.6, 6.7 et 6.8. Les conditions de simulations sont identiques au paragraphe 6.2. L'erreur entre l'objet vrai et l'objet estimé pour l'objet « vue aérienne » est de 22% pour un bruit de 1% et de 41% pour un bruit de 14%, pour l'objet « entrepôt », l'erreur est 6,3% pour un bruit de 1% et de 24% pour un bruit de 14% et pour l'objet « entrepôt », l'erreur est 7,4% pour un bruit de 1% et de 20% pour un bruit de 14%. Les estimées objet obtenues par la méthode jointe sans régularisation objet et celles données par la méthode mixte sont sans comparaison : on retrouve ici de bons résultats proche de ceux obtenus à partir de la méthode marginale. A fort flux, les deux méthodes mixte et marginale fournissent des estimées identiques dans le cas de l'objet vue aérienne et légèrement meilleures pour la méthode mixte pour les deux autres objets. En effet pour ces objets, l'erreur sur l'estimée des aberrations à 1% de bruit est plus faible dans le cas de l'estimateur joint (se reporter au paragraphe 5.2). A faible flux l'estimée de l'objet donnée par la méthode mixte est de moins bonne qualité que celle obtenue par la méthode marginale, ce qui n'est pas surprenant puisqu'à faible flux, l'erreur sur l'estimée des aberrations fournie par la méthode marginale est plus faible que l'erreur obtenue avec la méthode jointe (voir paragraphe 5.2).

6.5 Conclusion

Ce chapitre consacré à la restauration de l'objet a permis de montrer que la méthode jointe sans régularisation objet n'est pas du tout adaptée à une bonne restauration de l'objet. Au contraire, la méthode marginale permet d'effectuer une estimation de bonne qualité des aberrations par MAP en bénéficiant des bonnes propriétés de cet estimateur puis à partir des estimées aberrations et hyperparamètres d'effectuer une restauration de l'objet également par MAP. Cette restauration objet est d'autant plus satisfaisante qu'elle ne nécessite pas de calcul supplémentaire et qu'elle permet une restauration optimale à la fois des aberrations et de l'objet. Enfin, on a développé une méthode appelée « mixte » qui à partir des bons résultats d'estimation des aberrations obtenues avec l'estimateur joint à fort flux, permet de restaurer l'objet. Cette estimation est effectuée en utilisant une méthodologie dérivée de la méthode marginale et présente essentiellement l'avantage d'être plus rapide.



FIG. 6.6 – Objet « vue aérienne » estimé par la méthode « mixte ». A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 22%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 41%.



FIG. 6.7 – Objet « entrepôt » estimé par la méthode « mixte ». A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 6,3%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 24%.



FIG. 6.8 – Objet « maison » estimé par la méthode « mixte ». A gauche l'objet vrai, au centre l'objet estimé pour un bruit de 1%, l'erreur entre l'objet vrai et l'objet estimé est de 7,4%. A droite pour un bruit de 14%, l'erreur entre l'objet vrai et l'objet estimé est de 20%.

Chapitre 7

Estimation des fortes aberrations

7.1 Introduction

Les données dont on dispose en diversité de phase sont constituées de l'image plan focal et d'une ou plusieurs images défocalisées. Ces images contiennent de l'information à la fois sur l'objet observé et sur la phase aberrante. Cependant les informations concernant cette dernière sont partielles. En effet dans la relation de formation d'image, la phase intervient dans l'argument d'une exponentielle complexe, l'image n'est donc porteuse d'information que sur la phase repliée sur l'intervalle $[0, 2\pi]$. Cette indétermination apparaît dès lors que l'amplitude de la phase dépasse 2π et affecte donc les systèmes fortement perturbés. C'est notamment le cas des télescopes d'observation de l'espace depuis le sol terrestre dont les images sont dégradées par la turbulence atmosphérique. A condition d'estimer correctement la phase modulo 2π , la restauration d'image par traitement *a posteriori* n'est pas gênée par cette indétermination. La correction temps réel des aberrations d'un système par optique adaptative par contre nécessite l'estimation de la phase aberrante dans la pupille. Cette application requiert donc la levée de l'ambiguïté.

Ce chapitre traite du problème spécifique des phases de grandes amplitudes. Le paragraphe suivant précise les spécificités du problème et s'intéresse aux différentes méthodes qui peuvent être utilisées pour estimer de telles phases. Le paragraphe 3 développe une méthode originale d'estimation MAP de la phase modulo 2π dans la pupille. Le paragraphe 4 présente un ensemble de résultats obtenus sur simulations.

7.2 Indétermination de 2π sur la phase

La réponse impulsionnelle monochromatique h est liée à la phase ϕ dans la pupille par une relation non linéaire qui est telle que la phase apparaît dans l'argument d'une exponentielle complexe :

$$h(x,y) = |TF^{-1} [P(x,y)e^{i\phi(x,y)}]|^2 \quad (7.1)$$

Toute variation de 2π de la phase laisse inchangée h et donc l'image. Il existe une infinité de phases qui correspondent donc à la même réponse impulsionnelle. La figure 7.1 propose deux exemples de phases équivalentes vis-à-vis de l'image. Les données seules (et ce quel que soit leur nombre) ne permettent pas de discriminer la vraie phase parmi toutes celles équivalentes puisqu'elles ne contiennent de l'information que sur la phase modulo 2π . Ainsi le terme d'attache aux données $J(\phi) = \|i - o * h(\phi)\|^2$ admet une infinité de minima

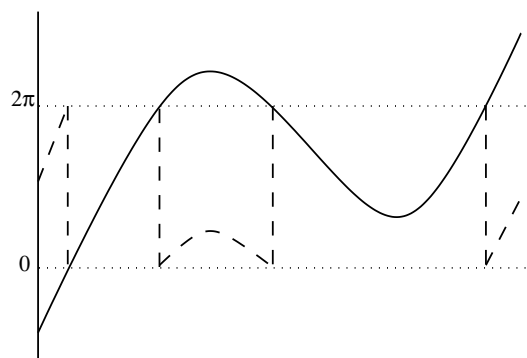


FIG. 7.1 – La phase en trait plein et la phase en tiret donnent la même réponse impulsionnelle.

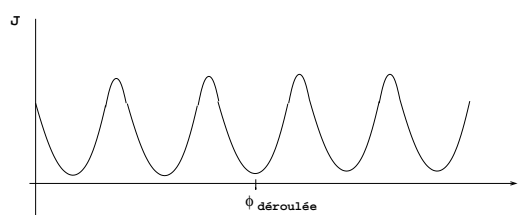


FIG. 7.2 – Allure du critère d'attache aux données

tous équivalents. Une représentation symbolique 1D d'un tel critère est tracé sur la figure 7.2: la phase vraie (déroulée) correspond à un des minimum de J mais elle n'est en rien discriminable, vis-à-vis du critère d'attache aux données, par rapport à l'infinité des autres solutions correspondant à des variations de 2π par rapport à la phase déroulée.

Remarque Si la diversité sur la phase introduite dans l'image défocalisée permet de lever l'indétermination de signe sur la partie paire de la phase (voir annexe A), elle n'a aucun effet sur l'incertitude de 2π .

7.3 Méthodes d'estimation possibles

Deux stratégies sont possibles pour estimer les phases de fortes amplitudes: soit rechercher directement la phase dépliée, soit reconstruire la phase modulo 2π puis dans un deuxième temps, lorsque cela est nécessaire, effectuer son déroulage.

7.3.1 Estimation de la phase dépliée

La première méthode possible consiste à estimer et dérouler simultanément la phase. Pour lever l'indétermination de 2π , il faut alors, lors de l'estimation, introduire un critère de régularité concernant la phase puisque les données ne contiennent pas d'information permettant son déroulage.

7.3.1.1 Choix d'une base

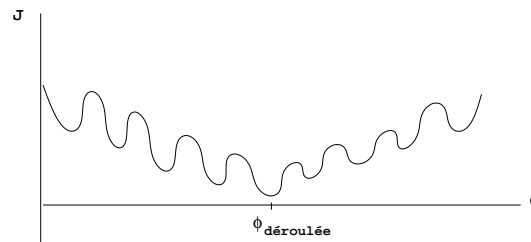


FIG. 7.3 – Allure du critère d'attache aux données lorsque la phase est décomposée sur la base des Zernike tronquée

Base des Zernike Dans les chapitres précédents, traitant du cas des aberrations faibles, nous avons choisi d'utiliser la base des Zernike pour représenter la phase (tout comme dans toute la littérature de la diversité de phase). Dans le cas d'une paramétrisation de la phase via sa décomposition infinie sur la base des Zernike, la forme du terme d'attache aux données reste inchangé : il présente toujours une infinité de minima tous strictement équivalents. En effet, toute phase (et donc $\phi[2\pi]$) peut être parfaitement modélisée par une infinité de coefficients de Zernike. Mais en pratique c'est une décomposition tronquée de la phase qui est effectuée. Dans ce cas, les minima du critère ne sont plus équivalents et le seul global correspond à la phase déroulée. Prenons l'exemple d'une phase correspondant à une défocalisation pure soit $\phi(x) = 3Z_4(x)$, la décomposition sur les 350 premiers coefficients de Zernike, de la version enroulée de cette phase sur l'intervalle $[0, 2\pi]$ est présenté sur le schéma 7.4, à droite. On observe la présence de coefficients d'ordre très élevés. Ces coefficients donnent naissance à la phase présentée sur la figure 7.4 (graphe de gauche) représentée en trait plein qui est à comparer avec la vraie phase enroulée sur l'intervalle $[0, 2\pi]$, tracée en pointillé. Malgré le grand nombre de modes ayant servi à représenter la phase, celle-ci est loin de correspondre à la phase vraie enroulée sur $[0, 2\pi]$. Le critère d'attache aux données prend alors, dans le cas d'une décomposition tronquée de la phase, une forme proche de celle schématisée sur la figure 7.3. Il présente un très grand nombre de minima quasi-équivalents qui rend impossible la découverte du minimum global. En pratique lorsque l'on recherche une phase de forte amplitude, le plus souvent ce n'est pas la phase dépliée qui est estimée mais une version « lissée » de la phase repliée sur l'intervalle $[0, 2\pi]$ correspondant à un minimum local du critère.

Base des indicatrices Une autre base de décomposition possible pour la phase est la base des indicatrices : on cherche à présent la phase point-à-point dans la pupille. Cette base est telle que la décomposition de la phase modulo 2π sur les indicatrices est identique à celle de la phase déroulée, à 2π près. Ainsi dans le cas de la décomposition de la phase sur cette base et contrairement à la base tronquée des Zernike, le terme d'attache aux données présente toujours une infinité de minima strictement équivalents (schéma 7.2).

7.3.1.2 Terme de régularisation

Bien que la base tronquée des Zernike ne conserve pas l'indétermination de 2π , en pratique, sa seule utilisation ne permet pas d'atteindre le minimum globale du critère (celui-ci étant « noyé » au milieu de très nombreux minima locaux). Quant à la base des indicatrices, elle laisse inchangée la forme du critère d'attache aux données. Il faut donc dans les deux cas, ajouter explicitement de l'information supplémentaire sur la phase lors de l'inversion du

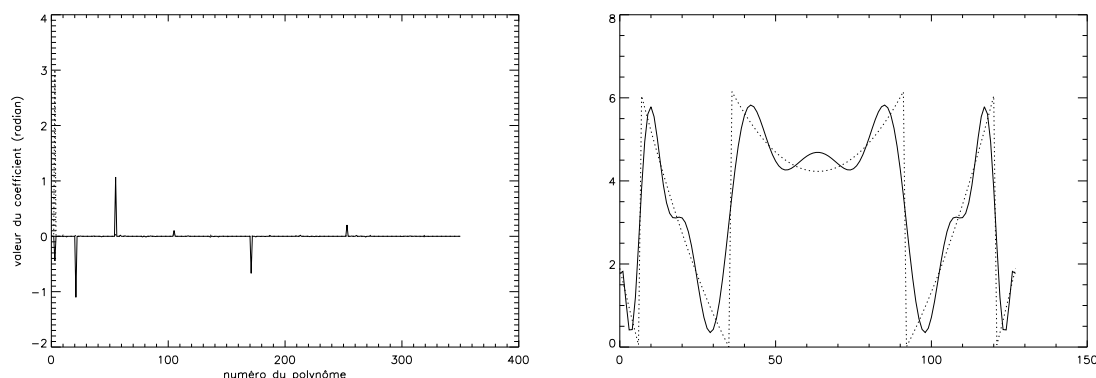


FIG. 7.4 – *A gauche, valeurs des coefficients de la décomposition de la phase modulo 2π sur les 350 premiers Zernike. A gauche (en trait plein) est représenté la phase reconstruite à partir des 350 coefficients (présentés sur la figure de gauche) et en pointillé la phase vraie repliée sur l'intervalle $[0, 2\pi]$.*

problème afin d'espérer pouvoir retrouver la phase dépliée. Pour cela on introduit un terme de régularité sur la phase afin de « déformer » le critère d'attache aux données pour qu'il ne présente plus qu'un seul minimum global le plus « marqué » possible qui, on l'espère, correspond à la phase dépliée. Pour la reconstruction de la phase sur la base des Zernike, on peut utiliser une régularisation du type de celle proposée au paragraphe 2.2.1.2.

Dans le cas des méthodes ayant pour objectif l'estimation de la phase dépliée, le critère à minimiser même régularisé présente un très grand nombre de minima qui ne sont pas équivalents, le minimum global correspondant à la phase dépliée. La minimisation d'un tel critère est très difficile. Dans le cas où elle est assurée par une méthode de minimisation locale (du type des méthodes à base de gradients), l'espoir de retrouver effectivement la phase dépliée est totalement conditionnée par le point d'initialisation de la minimisation. La phase initiale doit être assez proche de la phase vraie afin que le premier minimum rencontré lors de la minimisation corresponde effectivement à la phase dépliée. L'issue de l'estimation et du déroulage simultanée de la phase est donc très incertaine.

7.3.2 Estimation de $\phi[2\pi]$ (puis déroulage)

La résolution simultanée de l'estimation et du déroulage de la phase est ardue : chacune des deux étapes étant en soi difficile. La stratégie employée ici est différente, on ne cherche à estimer, dans un premier temps, que la phase modulo 2π puis, dans un deuxième temps, lorsque cela est nécessaire, elle est déroulée.

Estimation de $\phi[2\pi]$ On souhaite conserver l'équivalence entre tous les minima du critère correspondant à des variations de 2π de la phase, autrement dit on ne cherche pas à lever l'indétermination des données. Pour cela il faut choisir une base de décomposition de la phase telle que la décomposition de la phase vraie ou de sa version repliée sur $[0, 2\pi]$ soit la même modulo 2π . Comme nous l'avons dit précédemment, la base des indicatrices vérifie cette propriété et on peut montrer que c'est la seule. La reconstruction de la phase pixel-à-pixel dans la pupille introduit un grand nombre d'inconnues à reconstruire et nécessite donc l'ajout

d'un terme de régularisation explicite sur la phase. Dans le souci de conserver une équivalence stricte entre toutes les solutions équivalentes vis-à-vis des données, la régularisation doit elle aussi être insensible à toute variation de 2π en tout point de la pupille.

L'estimation de $\phi[2\pi]$ est faite en minimisant un critère qui présente un grand nombre de minima mais tous équivalents et correspondant chacun à une estimée de la phase pertinente par rapport à l'objectif fixé. La minimisation de ce critère est donc bien plus robuste que dans le cas précédent.

Déroulage Une fois la phase modulo 2π correctement estimée, il reste, lorsque cela est nécessaire, à la dérouler pour retrouver la phase dépliée. Le problème du déroulage à deux dimensions se retrouve dans de nombreuses disciplines et en particulier en imagerie SAR. La littérature comporte de nombreux articles traitant de ce problème de façon générale, on pourra notamment se reporter au livre de (Ghiglia et Pritt, 1998).

Nous avons développé, dans la suite de ce chapitre, une méthode basée sur l'estimation de la phase à 2π près fonctionnant sur objet étendu. Ce choix est basé d'une part sur la constatation que l'estimation et le déroulage simultanés de la phase est difficile et d'autre part sur le fait que l'une des applications à laquelle nous nous intéressons, la restauration d'images *a posteriori*, ne nécessite pas le déroulage de la phase. Le chapitre suivant présente une méthode originale d'estimation de la phase modulo 2π par Maximum *A Posteriori*. Nous n'avons pas abordé le problème du déroulage de phase éventuellement nécessaire. Mais avant cela, on présente, dans le paragraphe suivant, les différentes méthodes utilisées dans la littérature de la diversité de phase pour résoudre le problème des phases de fortes amplitudes.

7.3.3 Méthodes employées dans la littérature

7.3.3.1 Estimation et déroulage simultanés

Base des Zernike Dans le cas d'observation d'objet étendu depuis le sol terrestre, la décomposition sur la base des Zernike a été privilégiée par (Löfdahl et Scharmer, 1994; Seldin et Paxman, 1994; Van Kampen et Paxman, 1998), dans le cas de faibles D/r_o (2,5 à 5). Suivant la force de la turbulence le nombre de modes reconstruits varie de 15 à 55. La même paramétrisation a été employée par (Thelen et al., 1999) dans le cas d'un D/r_o de 5, en associant de plus, aux coefficients turbulents, un terme d'*a priori* lié aux connaissances statistiques que l'on possède sur la turbulence (il s'agit d'un terme de régularisation identique à celui présenté pages 31 et 33).

Base des indicatrices (Jefferies et al., 2002) propose de reconstruire les phases turbulentes point-à-point dans la pupille et de les paramétrer par une fonction convoluée par une gaussienne. Pour pallier le problème des nombreux minima locaux, Jefferies effectue une première estimation de la phase en la paramétrant sur trois coefficients de Zernike (a_4 terme de défocalisation et a_5 , a_6 correspondant aux astigmatismes), et se sert de cette première estimée comme point de départ pour l'estimation de la phase point-à-point.

7.3.3.2 Estimation puis déroulage

En diversité de phase, deux équipes ont privilégié une approche en deux étapes : estimation de $\phi[2\pi]$ puis déroulage (Löfdahl et al., 1998; Baba et Mutoh, 2001). Toutes deux recherchent les aberrations propres d'un télescope en observant un objet ponctuel à travers la turbulence atmosphérique. Les conditions de turbulence considérée dans ces deux articles sont sévères : D/r_o proche de 30. L'estimation de la phase repliée $\phi[2\pi]$ est effectuée par Maximum de Vraisemblance dans le cas de l'équipe de Löfdahl et par une méthode de type Gerchberg-Saxton pour l'équipe japonaise. Aucune précision n'est donnée quant à la méthode qui permet de dérouler la phase dans l'article de Löfdahl, pour l'article de Baba et Mutoh, le déroulage est basé sur une méthode proposée par (Ghiglia et Romero, 1994).

7.4 Construction du critère MAP

Ce paragraphe développe une méthode d'estimation de la phase modulo 2π par MAP.

7.4.1 Choix d'un terme de régularisation

7.4.1.1 Nécessité de régulariser

Deux types de régularisation ont été proposés dans le chapitre 2 (page 31) pour une phase reconstruite via sa décomposition sur la base des Zernike : une implicite consistant à limiter la décomposition aux tous premiers coefficients, l'autre basée sur l'ajout d'un terme d'*a priori* explicite sur les coefficients de Zernike. À présent, la phase est reconstruite point-à-point dans la pupille, le nombre d'inconnus « phase » est passé de quelques dizaines (coefficients) à quelques milliers (pixels de la pupille). Il faut construire un terme de régularisation adapté à la recherche de la phase modulo 2π respectant donc certaines contraintes.

7.4.1.2 Exigences

On souhaite pénaliser les écarts importants entre pixels (dus au bruit) mais tout en étant insensible aux différences correspondant à des variations de $2k\pi$ avec $k \in \mathbb{Z}$ entre pixels de phase. Ces considérations ainsi que d'autres exigences vis-à-vis du terme de régularisation ont guidé le choix de la forme de l'*a priori* qui a été développé. Ces contraintes sont basées sur un besoin commun, conserver les indéterminations des données. Le terme de régularisation doit être :

- Insensible aux sauts de 2π : l'objectif est de reconstruire la phase à 2π près. Le terme de régularisation doit donc être, tout comme les données, insensible à toute variation de $2k\pi$ d'un pixel de phase. Afin d'assurer cette insensibilité, on construit un terme de régularisation non pas sur la phase ϕ elle-même mais sur les phaseurs c'est-à-dire sur $e^{j\phi}$ (bien que le terme de régularisation porte explicitement sur les phaseurs, c'est malgré tout bien la grandeur ϕ qui est le paramètre recherché).
- Insensible au piston : le piston correspond à l'ajout d'un terme constant sur toute la phase. Il est donc sans effet lorsque l'on fait de l'imagerie. Les données étant insensibles à ce terme, on souhaite qu'il en soit de même pour le terme de régularisation.
- Insensible aux tilts : les tilts représentent de simples basculements globaux de la phase. Ils se traduisent pour la réponse impulsionnelle associée à la phase par un décalage. Lorsque l'objet observé est inconnu, ce décalage peut être indifféremment attribué à

l'objet ou à la RI (du fait de la relation de convolution qui lie objet et RI à l'image). Pour cette raison, le critère joint J'_{MAPJ} (équation 2.12) ainsi que le critère marginal J_{MAP} (équation 4.18) sont insensibles aux tilts. Le terme de régularisation doit donc également respecter cette indétermination.

7.4.1.3 Forme du terme de régularisation

Le terme de régularisation suivant vérifie toutes ces conditions :

$$J_{\text{regul}}(\phi) = \sum_{(l,m) \in S} \left[\left| e^{j(\phi_{l-1,m} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l+1,m})} \right|^2 + \left| e^{j(\phi_{l,m-1} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l,m+1})} \right|^2 \right] \quad (7.2)$$

La sommation est effectuée sur tous les pixels appartenant au support S de la pupille. Ce terme lie un pixel de phase $\phi_{l,m}$ à ces quatre plus proches voisins. Comme on impose une contrainte stricte de support à la phase, les pixels qui se trouvent en bord de pupille n'ont pas tous leurs voisins qui sont sur le support de la pupille. Lors de la sommation, tous les termes $|\dots|^2$ dont un pixel n'est pas sur la pupille sont supprimés.

La forme de la régularisation basée sur une contrainte imposée sur les phaseurs ne permet pas de déduire facilement l'effet de ce terme sur la phase. On peut à l'aide d'un développement limité au premier ordre montrer son effet sur ϕ :

$$J_{\text{regul}}(\phi) \simeq \sum_{(l,m) \in S} \left[|\phi_{l-1,m} - 2\phi_{l,m} + \phi_{l+1,m}|^2 + |\phi_{l,m-1} - 2\phi_{l,m} + \phi_{l,m+1}|^2 \right]$$

Ce qui correspond à une pénalisation quadratique sur la différence seconde de ϕ .

7.4.2 Critère MAP

Afin d'effectuer une estimation au sens du MAP de la phase ϕ , l'estimateur marginal, présenté au chapitre 4, est utilisé.

$$\hat{\phi}_{\text{MAP}} = \arg \min_{\phi} J_{\text{MAP}}(\phi) \quad (7.3)$$

L'expression de $J_{\text{MAP}}(\phi)$ est déduite de l'expression (4.18) en remplaçant le terme de régularisation portant sur les coefficients de Zernike \mathbf{a} par celui développé pour ϕ – relation (7.2).

$$\begin{aligned} J_{\text{MAP}}(\phi) &= N^2 \ln 2\pi + \frac{1}{2} \sum_v \ln \left(|\tilde{h}_f(\phi, v)|^2 + |\tilde{h}_d(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right) + \frac{N^2}{2} \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) \\ &+ \frac{1}{2} \sum_v \frac{|\tilde{i}_d(v) \tilde{h}_f(\phi, v) - \tilde{i}_f(v) \tilde{h}_d(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \left[|\tilde{h}_f(\phi, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\phi, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right]}{\sigma^2 \left(|\tilde{h}_f(\phi, v)|^2 + |\tilde{h}_d(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right)} \\ &+ \gamma \sum_{(l,m) \in S} \left[\left| e^{j(\phi_{l-1,m} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l+1,m})} \right|^2 + \left| e^{j(\phi_{l,m-1} - \phi_{l,m})} - e^{j(\phi_{l,m} - \phi_{l,m+1})} \right|^2 \right] \end{aligned} \quad (7.4)$$

avec γ l'hyperparamètre réglant le poids entre le terme de régularisation et celui d'attache aux données.

7.4.3 Réglage des hyperparamètres

7.4.3.1 Hyperparamètres liés au bruit et à l'objet

Tout comme au paragraphe 4.8, on propose d'estimer conjointement les hyperparamètres liés au bruit θ_b et à l'objet θ_o avec les pixels de phase. On rappelle :

$$(\hat{\phi}, \hat{\theta}_o, \hat{\theta}_b) = \arg \min_{\phi, \theta_o, \theta_b} J_{\text{MAP}}(\phi, \hat{\theta}_o, \hat{\theta}_b, \hat{\theta}_\phi) \quad (7.5)$$

7.4.3.2 Hyperparamètre lié à la phase

La valeur de l'hyperparamètre γ est ajustée au mieux empiriquement. Il ne peut pas, au contraire des hyperparamètres objet et bruit, être réglé de façon non supervisée (la solution $\gamma=0$, correspondant à une absence de régularisation, minimise le critère MAP quelle que soit la phase).

7.4.4 Minimisation du critère

Il n'y a pas de solution analytique donnant $\hat{\phi}_{\text{MAP}}$. La minimisation du critère $J_{\text{MAP}}(\phi)$ est effectuée itérativement à l'aide d'un algorithme de type gradient conjugué.

Expression des gradients L'expression du gradient du critère J_{MAP} par rapport aux pixels de phase est rappelé au paragraphe 4.5. Il convient cependant de lui ajouter un terme correspondant au gradient introduit par la régularisation. Ce gradient est donné par :

$$\begin{aligned} \frac{dJ_{\text{regul}}}{d\phi_{l,m}} &= \sin(\phi_{l,m} - 2\phi_{l+1,m} + \phi_{l+2,m}) - 2 \sin(\phi_{l-1,m} - 2\phi_{l,m} + \phi_{l+1,m}) \\ &\quad + \sin(\phi_{l-2,m} - 2\phi_{l-1,m} + \phi_{l,m}) + \sin(\phi_{l,m} - 2\phi_{l,m+1} + \phi_{l,m+2}) \\ &\quad - 2 \sin(\phi_{l,m-1} - 2\phi_{l,m} + \phi_{l,m+1}) + \sin(\phi_{l,m-2} - 2\phi_{l,m-1} + \phi_{l,m}) \end{aligned} \quad (7.6)$$

Le terme de régularisation fait intervenir les quatre plus proches voisins d'un pixel. Le gradient regarde, quant à lui, jusqu'à huit voisins. Dans le calcul du terme de régularisation $J_{\text{regul}}(\phi)$, les termes $|\dots|^2$ dont un des pixels n'était pas sur la pupille n'ont pas été pris en compte lors de la sommation. Pour être cohérent, on fait de même pour le gradient : tout terme $\sin(\dots)$ dont un des pixels de phase $\phi_{l,m}$ n'est pas sur le support est supprimé.

Forme du critère L'approche adoptée consiste à construire un critère dont tous les minima correspondant à des variations de 2π de la phase sont strictement équivalents. L'allure globale du critère ressemble donc à celle du critère présenté sur la figure 7.2. Cette infinité de minima n'est en rien gênante puisque que nous ne nous intéressons qu'à reconstruire de $\phi[2\pi]$, n'importe quelle solution correspondant à une des cuvette est un minimum global et est donc satisfaisante. Cependant du fait de la relation non linéaire qui lie la phase à l'image, il peut y avoir des minima locaux à l'intérieur d'une « cuvette » du critère (c'est-à-dire autour de chaque minimum global). On souhaite que le terme de régularisation introduit ait réduit au maximum ses minima mais dans le cas où il en resterait, on peut espérer « lisser » le critère en ajoutant aux données constituées par l'image plan focal et celle défocalisée, une deuxième image défocalisée. Une autre solution consiste à ne pas modifier la forme du critère mais à améliorer le point de départ de l'initialisation afin de se rapprocher le plus possible du minimum global.

7.4.5 Choix d'un critère de qualité

Le paragraphe qui suit propose des exemples de restaurations obtenues à l'aide du critère MAP. Comme l'estimation développée ici porte sur la reconstruction de phase modulo 2π et non sur la phase déroulée, il faut définir un critère permettant de quantifier la qualité de phase reconstruite sur l'intervalle $[0, 2\pi]$. Pour cela, on définit un critère d'erreur de reconstruction indiquant l'écart entre la phase estimée et la phase vraie et qui est insensible aux sauts de 2π au travers des phaseurs ($e^{j\phi}$):

$$Err = \sqrt{\frac{1}{\mathcal{A}} \sum_{(l,m) \in \mathcal{S}} c(l,m)^2} \quad (7.7)$$

$$\text{avec } c(l,m) = \sqrt{\left| e^{j(\phi_{l,m}^{\text{vraie}} - \phi_{l,m}^{\text{estimée}})} - \left\langle e^{j(\phi_{l,m}^{\text{vraie}} - \phi_{l,m}^{\text{estimée}})} \right\rangle \right|^2} \quad (7.8)$$

avec \mathcal{A} le nombre de pixels dans la pupille. La quantité c , que l'on appellera carte de l'erreur, représente l'écart entre les deux phases, en chaque point de la pupille. La présence d'un piston ou d'un tilt différentiel entre les deux phases n'ayant aucune incidence sur la qualité de reconstruction (voir paragraphe 7.4.1.2), on souhaiterait avoir un critère de qualité qui soit également insensible à une différence constante ou à un basculement entre ϕ^{vraie} et $\phi^{\text{estimée}}$. Si le critère précédent est bien insensible à une différence de piston entre la phase vraie et celle reconstruite, il n'est pas insensible à une différence de tilts. Nous n'avons, en effet, pas trouvé de critère ne prenant pas en compte les tilts différentiels mais tout en restant sensible à toutes les autres composantes de la phase. Toutefois, l'estimateur que nous avons construit étant insensible aux tilts, nous avons constaté qu'en partant d'une phase initiale ne comportant pas de composantes de tilts, il n'en apparaisse pas au cours de la minimisation¹.

En effectuant un développement limité au premier ordre de $e^{j(\phi_{l,m}^{\text{vraie}} - \phi_{l,m}^{\text{estimée}})}$, on remarque que l'erreur (Err) que nous avons définie est égale, lorsque $(\phi_{l,m}^{\text{vraie}} - \phi_{l,m}^{\text{estimée}})[2\pi]$ est petit, à l'écart type de la phase résiduelle.

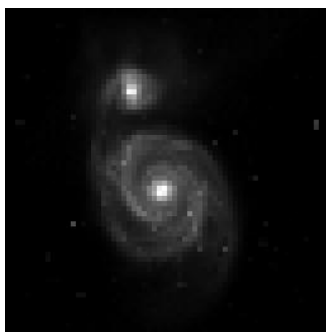
7.5 Résultats de simulations

Ce paragraphe regroupe des résultats d'estimation de phases fortement aberrantes estimées sur données simulées. Ces résultats ont été obtenus par deux méthodes différentes de restauration : d'une part, la méthode originale que nous venons de proposer; reconstruction des pixels de phase dans la pupille à l'aide du critère MAP – équation (7.4) – et d'autre part une méthode plus classique basée sur la décomposition tronquée de la phase sur la base des Zernike et utilisant le critère MAP présenté au chapitre 4 – équation (4.18).

7.5.1 Exemple montrant l'intérêt de la reconstruction point-à-point

On propose un premier exemple de restauration d'une phase de forte amplitude qui permet de mettre en évidence le problème évoqué au paragraphe 7.3.1.1 lié à la décomposition tronquée de la phase sur la base des Zernike.

1. Cette constatation est en accord avec une propriété connue des algorithmes de minimisation à base de gradient.

FIG. 7.5 – *Objet utilisé lors des simulations*

7.5.1.1 Conditions de simulation

La phase vraie est générée à partir des coefficients de Zernike $a_5 = -1,5$ radian, $a_6 = 2,5$ radian, $a_5 = 1$ radian. Elle est simulée sur une pupille circulaire, sans occultation centrale de 32×32 pixels. L'écart type spatial de la phase ainsi simulée est de $\lambda/2$ et son amplitude crête-à-crête de $2,8\pi$ radian. L'objet observé est la galaxie M51 représentée figure 7.5. Les images (de taille 64×64 pixels) sont obtenues par convolution de l'objet avec la RI associée à la phase vraie et bruitées par un bruit blanc gaussien tel que le niveau de bruit soit de 1%.

7.5.1.2 Résultats

Le résultat de l'estimation de la phase aberrante par la méthode point-à-point est présenté sur la figure 7.6. La valeur de l'hyperparamètre est fixée au mieux empiriquement (ici il est égal à 1000). On suppose que la forme de la pupille est parfaitement connue. Le point de départ de la minimisation correspond à une phase nulle dans la pupille. La figure 7.6 présente la phase vraie, la phase reconstruite et la carte de l'erreur c . L'erreur Err est de $\lambda/50$. On constate que la phase reconstruite est déroulée; cela provient du point d'initialisation dont le minimum le plus proche correspond à la phase continue mais on rappelle qu'aucune information dans le critère ne privilégie cette solution. La qualité de reconstruction est très bonne. Pour comparaison, on propose à présent de reconstruire la phase via sa décomposition modale tronquée sur la base des Zernike. La phase est reconstruite sur dix-huit modes (de a_4 à a_{21}). Aucun terme de régularisation sur les coefficients a_i n'est utilisé dans ce cas, le nombre de coefficients recherchés étant faible (voir chapitre 2) et la phase générée ne correspondant pas à une phase turbulente. La figure 7.7 regroupe le résultat de l'estimation. Même si la reconstruction de la phase par décomposition modale conduit à la reconstruction de la phase déroulée, c'est également le critère de qualité présenté au paragraphe 7.4.5 qui est utilisé dans le cas de la reconstruction modale de la phase, afin de pouvoir comparer les deux méthodes. La qualité de la phase estimée est très inférieure à celle obtenue en point-à-point : l'erreur est ici de $\lambda/7$. On observe ici le problème qui a été évoqué dans le paragraphe 7.3.1.1 : la phase « ajustée » par les dix-huit coefficients recherchés n'est pas celle continue mais une phase présentant des sauts de 2π , des coefficients d'ordres supérieurs à ceux ayant générés la phase vraie sont apparus pour essayer de suivre au mieux ces sauts et ont, de ce fait, dégradé l'estimation des coefficients de fréquences spatiales faibles.

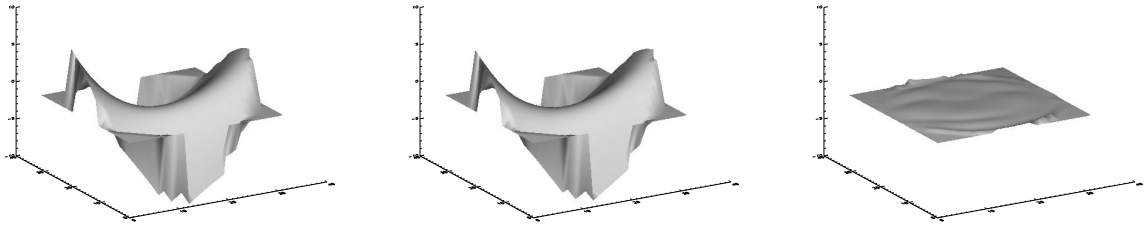


FIG. 7.6 – *Reconstruction de la phase sur la base des indicatrices : à gauche la phase vraie, au centre la phase estimée, et à droite la carte de l'erreur ($\lambda/50$).*

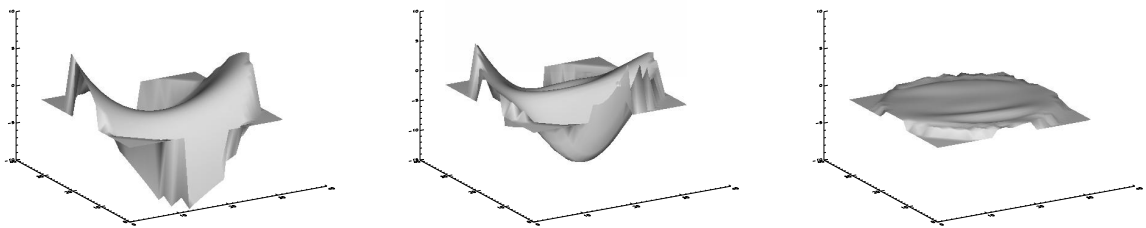


FIG. 7.7 – *Reconstruction de la phase sur la base des Zernike : à gauche la phase vraie, au centre la phase estimée, et à droite la carte de l'erreur ($\lambda/7$).*

7.5.2 Phases turbulentes

7.5.2.1 Conditions de simulation

Vingt phases turbulentes sont générées à partir de vingt tirages aléatoires décorrélés temporellement de coefficients de Zernike dont le spectre suit la loi de Kolmogorov pour $D/r_o = 1$ (voir page 17). On modifie la force de la turbulence en multipliant ces tirages par $(D/r_o)^{5/6}$. Les images simulées sont échantillonnées en respectant le critère de Shannon. Lorsque seulement deux images sont utilisées (et sauf mention contraire), l'amplitude crête-à-crête de défocalisation de la deuxième image est de 2π radian.

Objet ponctuel Pour les simulations à objet ponctuel, les phases turbulentes sont simulées à partir du tirage aléatoire des 253 premiers polynômes de Zernike (hors piston). La taille de la pupille est de 64×64 pixels et la valeur de D/r_o est fixée à 30. L'écart type spatial théorique de la phase turbulente est de $(2,7\lambda)$ (voir relation 1.24). Un bruit blanc gaussien est ajouté sur les images tel que le RSB soit de 100 sur l'image plan focal; dans le cas d'objet ponctuel, tout comme dans le chapitre 3, le RSB est défini comme le rapport entre la valeur maximale de l'image plan focal sur l'écart type du bruit.

Objet étendu L'objet observé est la galaxie M51. La phase est simulée sur une pupille de diamètre 32×32 pixels à partir de tirages aléatoires des 105 premiers Zernike (hors piston). Toutefois lors de la construction des phases turbulentes à partir des coefficients de Zernike, les composantes de tilts (coefficients a_2 et a_3) ont été supprimées (en effet, on a évoqué dans la paragraphe l'insensibilité du critère aux tilts, leur présence ou absence n'a donc aucune influence sur l'estimation). Pour les simulations qui suivent, on s'est intéressé à des valeurs

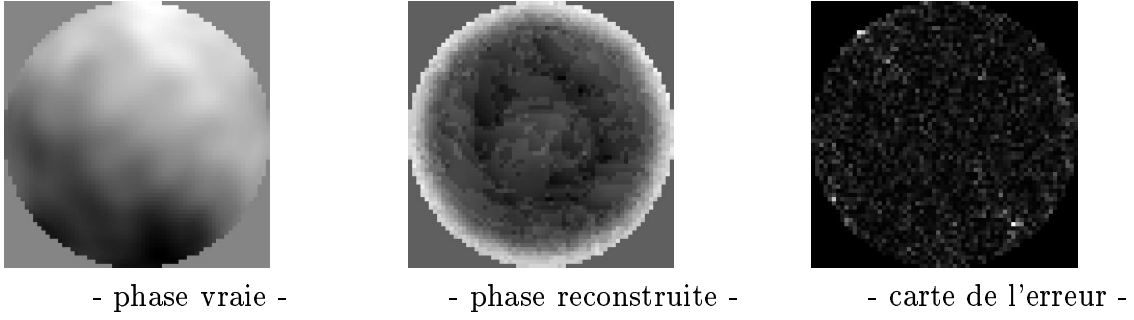


FIG. 7.8 – Exemple de restauration d'une phase turbulente ($D/r_o=30$) à partir d'un objet ponctuel (la carte de l'erreur a été multipliée par 10). L'écart type spatial de la phase vraie est de $(2,75\lambda)$ et son amplitude crête-à-crête de 22π radians. La phase est reconstruite point-à-point dans la pupille. L'erreur de reconstruction vaut $(\lambda/28)$.

de D/r_o allant de 4 à 10. Les variances théoriques des phases ainsi simulées (hors tilts) sont données par la relation (Noll, 1976) :

$$\sigma_\phi^2 = 0,134(D/r_o)^{5/3} \quad (7.9)$$

Les images sont obtenues par la convolution de l'objet observé avec les réponses impulsionnelles associées aux phases générées à partir des tirages aléatoires de coefficients turbulents puis bruitées par un bruit blanc gaussien stationnaire. Dans toute la suite, le niveau de bruit est fixé à 1% (dans le cas d'objet étendu, on rappelle que le niveau de bruit exprimé en pourcentage correspond au rapport entre l'écart type du bruit sur le flux moyen par pixel).

7.5.2.2 Résultats sur objet ponctuel

Un premier résultat sur objet ponctuel est présenté dans ce paragraphe. Les conditions de simulations sont celles indiquées dans le paragraphe 7.5.2.1. La force de la turbulence soit $D/r_o=30$ correspond approximativement aux cas traités dans (Löfdahl et al., 1998) et (Baba et Mutoh, 2001). Les estimations sont faites à partir d'une image plan focal et d'une image défocalisée. La phase est reconstruite pixel-à-pixel dans la pupille. L'estimation de la phase est faite par Maximum de Vraisemblance. En effet, lorsque l'objet est ponctuel, seule la phase est inconnue et on constate en pratique que pour des bruits modérés, il n'est pas nécessaire d'introduire un terme de régularisation pour effectuer une estimation de bonne qualité. L'erreur moyenne, sur vingt estimations obtenues à partir de vingt phases turbulentes tirées aléatoirement, est de $\lambda/30$. Un exemple de phase restaurée est proposé figure 7.8 : la phase vraie est présentée à gauche, la phase reconstruite au centre et la carte de l'erreur c à droite. La bonne adéquation entre la phase vraie et celle reconstruite n'est visible qu'à travers la carte de l'erreur ou sur la figure 7.9 qui propose les deux cartes de phases ramenée sur l'intervalle $[-\pi, \pi]$ car si la phase vraie est « continue », il n'en est rien de la phase reconstruite qui présente de très nombreux sauts de 2π .

En conclusion, dans le cas où l'objet est connu et où le niveau de bruit est faible, l'estimation de la phase sur l'intervalle $[0, 2\pi]$ se passe bien sans qu'il y ait besoin d'introduire de terme de régularisation.

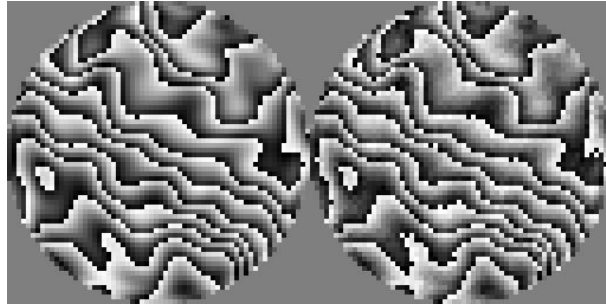


FIG. 7.9 – Exemple de restauration d'une phase turbulente ($D/r_o=30$) à objet ponctuel : à gauche, la phase vraie sur l'intervalle $[-\pi, \pi]$ et à droite, la phase reconstruite sur l'intervalle $[-\pi, \pi]$.

7.5.2.3 Résultats sur objet étendu

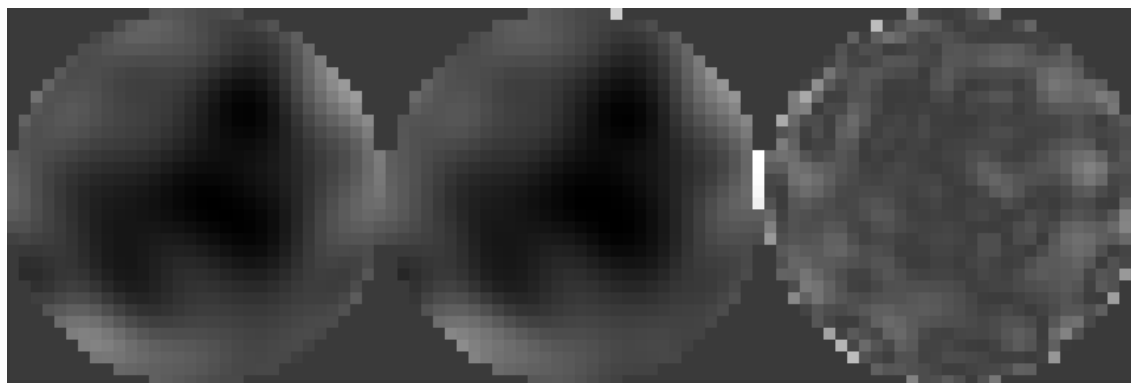
Dans la suite, nous nous sommes intéressés à l'estimation de phases turbulentes correspondant à des D/r_o allant de 4 à 10, à partir d'images d'objets étendus (la galaxie M51). Les conditions de simulations sont celles indiquées dans le paragraphe 7.5.2.1.

Le fait d'être à présent à objet inconnu, impose d'ajouter un terme de régularisation explicite sur la phase. Dans toute la suite, les phases sont estimées par MAP (c'est-à-dire en utilisant l'estimateur marginal). Pour la phase décomposée sur la base des Zernike, on associe aux coefficients turbulents une loi de probabilité gaussienne de moyenne nulle et de matrice de covariance R_a (voir paragraphe 2.2.1.2). La matrice de covariance est calculée en utilisant la relation 2.13, en fixant la valeur de D/r_o à sa valeur vraie. La phase vraie ayant été modélisée à partir des 105 premiers polynômes de Zernike, on prend soin de reconstruire la phase sur un nombre plus faible de modes (sauf mention contraire, 90) afin de tenir compte de la sous-modélisation toujours rencontrée dans un cas réaliste. Le terme de régularisation associé à la reconstruction de la phase point-à-point dans la pupille est celui proposé au paragraphe 7.4.1.3.

Le domaine de turbulence étudié (de 4 à 10) a été séparé en trois parties appelées : « faibles » aberrations, correspondant à $D/r_o=4$ et 5; aberrations « moyennes », D/r_o égaux à 6 et 7 et « fortes » aberrations regroupant les D/r_o de 8, 9 et 10.

- **Cas « faibles » aberrations** Les résultats présentés pour ce cas ont été obtenus à partir d'une image plan focal et d'une image défocalisée. La minimisation numérique du critère est faite à partir d'une phase initiale dans la pupille nulle. Pour $D/r_o=4$, sept phases parmi les vingt simulées ont une amplitude crête-à-crête supérieure à 2π , l'écart type théorique (hors tilt) étant de $\lambda/5$. Pour $D/r_o=5$, quinze des vingt phases étudiées présentent une amplitude crête-à-crête supérieure à 2π , l'écart type théorique (hors tilt) est de $\lambda/4$.

A partir de la phase point-à-point La valeur de l'hyperparamètre utilisée est fixée à 10, il correspond à l'hyperparamètre optimal (au sens « meilleur réglage » pour l'estimation des aberrations). Pour $D/r_o=4$, la moyenne sur les vingt réalisations de turbulence de l'erreur (Err) est de $\lambda/60$. La phase la moins bien estimée correspond à un erreur inférieure à

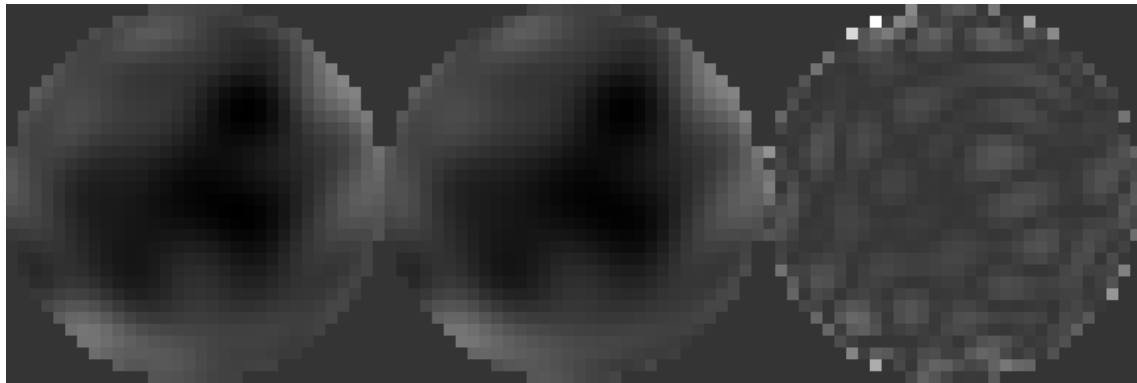


- phase vraie - - phase reconstruite - - carte d'erreur -

FIG. 7.10 – Exemple de restauration d'une phase turbulente ($D/r_o=4$) sur objet étendu (les cartes de la phase vraie et de celle reconstruite sont à la même échelle, la carte de l'erreur a été multipliée par 10). L'écart type spatial de la phase vraie est de $\lambda/4,4$ et son amplitude crête-à-crête de $1,15\pi$ radian. La phase est reconstruite point-à-point dans la pupille. L'erreur de reconstruction vaut $\lambda/40$

$\lambda/37$. Un exemple d'estimation est proposé sur la figure 7.10. La phase qui a servi à cette illustration fait partie des phases simulées ayant les variances spatiales les plus importantes. Son amplitude est supérieure à 2π et l'on observe que des pixels en bord de pupille (plusieurs à droite et un en haut) ont subi un saut de 2π par rapport aux mêmes pixels de la phase vraie. La carte de l'erreur c montre que les pixels les moins bien restaurés sont situés sur le bord de la pupille. Pour $D/r_o=5$, la valeur moyenne de l'erreur sur les vingt réalisations de turbulence est de $\lambda/25$. La qualité de reconstruction est bien inférieure au cas précédent à cause d'une phase qui a été mal restaurée, l'erreur d'estimation pour cette phase est de $\lambda/6$. Ce résultat correspond à une phase turbulente de fort écart type spatial $\lambda/3,4$ ainsi que de forte amplitude ($1,23\pi$ radians). La solution trouvée correspond à un minimum local du critère. Ainsi, comme nous l'avions évoqué au paragraphe 7.4.4, le critère peut présenter autour de ces minima globaux des minima locaux.

A partir des Zernike La même étude est menée avec la reconstruction de la phase sur la base des Zernike. Le qualité d'estimation obtenue pour $D/r_o=4$ est très proche de celle fournie par la reconstruction point-à-point. Mais il ne faut pas oublier que dans ce cas, la phase reconstruite correspond à la phase « physique » c'est-à-dire continue et non à une phase pouvant présenter des sauts de 2π , comme dans le cas précédent (et qui devra être suivie, dans le cas d'une correction temps réel, par une procédure de déroulage de phase). Le résultat d'estimation correspondant à la même phase turbulente que précédemment est présenté sur la figure 7.11. La phase reconstruite ne comporte pas de saut de 2π par rapport à la phase vraie, la continuité de la phase étant assurée par la paramétrisation de ϕ sur une base finie de Zernike. Le cas $D/r_o=5$ fournit une erreur moyenne de $\lambda/15$. Ce résultat est moins bon que dans le cas de la reconstruction point-à-point. La phase qui avait été mal reconstruite dans le cas point-à-point est ici aussi mal estimée ainsi que trois autres phases qui présentent toute la particularité d'avoir des écart types spatiaux importants.



- phase vraie - - phase reconstruite - - carte de l'erreur -

FIG. 7.11 – Exemple de restauration d'une phase turbulente ($D/r_o=4$) sur objet étendu (les cartes de la phase vraie et de celle reconstruite sont à la même échelle, la carte de l'erreur a été multipliée par 10). L'écart type spatial de la phase vraie est de $(\lambda/4,4)$ et son amplitude crête-à-crête de $1,15\pi$ radian. La phase est reconstruite sur 90 polynômes de Zernike. L'erreur de reconstruction vaut $(\lambda/40)$.

• **Cas aberrations « moyennes »** Les quelques mauvaises estimations rencontrées pour D/r_o égal à 5 concernant des phases de écart types importantes vont s'étendre à un plus grand nombre de cas ici, du fait de l'augmentation de la force de la turbulence. A partir de $D/r_o=6$, les vingt phases turbulentes que nous utilisons ont des amplitudes supérieures à 2π . Pour $D/r_o=6$, l'écart type spatial théorique de la phase (hors tilt) est de $\lambda/3,8$ et pour $D/r_o=7$ de $\lambda/3,3$. Une solution permettant de pallier le problème des phases mal estimées consiste à utiliser une image défocalisée supplémentaire lors de l'inversion. Ici aussi la phase initiale est prise nulle partout dans la pupille.

A partir de la phase point-à-point L'hyperparamètre optimal est toujours égal à 10. La restauration des phases correspondant à une force de turbulence de 6, à partir de deux images dont une est défocalisée de 2π radians conduit à une erreur moyenne sur les vingt réalisations de $\lambda/15$: la qualité d'estimation de deux phases est proche de $\lambda/6$, quatre autres présentent une erreur de $\lambda/15$ et les autres étant proches de $\lambda/30$.

Afin d'améliorer l'estimation, on propose, dans un premier temps, d'augmenter la défocalisation introduite entre les deux plans. En utilisant lors de l'estimation en plus de l'image plan focal, une image défocalisée de 4π radians par rapport au plan focal, l'erreur moyenne sur l'estimée de la phase diminue et devient $(\lambda/25)$. Deux phases parmi les vingt restent cependant mal estimées.

On utilise à présent, en plus de l'image focalisée et de celle défocalisée d'une distance de 2π radians, une deuxième image défocalisée de 4π radians. L'extension de l'expression du critère MAP au cas de trois images ainsi que celle de son gradient par rapport aux pixels de phase est présentée dans l'annexe F. Grâce à l'ajout de cette image, toutes les phases sont, à présent, estimées correctement. En moyenne, l'écart entre la phase vraie et celle estimée devient inférieur à $\lambda/40$. Une deuxième implémentation a été essayée, elle a consisté à utiliser une deuxième image défocalisée de -2π radians (les deux images défocalisées sont ainsi acquises de façon symétrique de part et d'autre du plan focal), elle a conduit à des résultats très proches de ceux obtenus pour les trois images de défocalisation 0, 2π , 4π radians.

Le cas $D/r_o=7$, conduit aux mêmes constatations que précédemment : l'utilisation de trois images permet d'obtenir une erreur moyenne proche de $\lambda/40$ contre $\lambda/10$ pour deux images. L'apport de l'image supplémentaire dans la minimisation est très vraisemblablement de « lisser » le critère qui, comme nous l'avons constaté dans la paragraphe précédent présente des minima locaux. Il faut noter que l'article de (Jefferies et al., 2002) portant également sur l'estimation des fortes aberrations (mais contrairement à notre cas il s'intéresse à la reconstruction de la phase ainsi que l'amplitude dans la pupille) utilise aussi plusieurs images défocalisées pour l'inversion du problème.

A partir des Zernike La reconstruction de la phase via sa décomposition tronquée sur la base des Zernike ne donne pas non plus de bons résultats lorsque l'inversion est effectuée à partir de deux images et d'une défocalisation de 2π . Pour $D/r_o=6$, l'erreur moyenne est de $\lambda/8$ (dans les mêmes conditions, l'erreur associé au point-à-point est de $\lambda/15$). En ajoutant une image défocalisée de 4π radians, l'estimation s'améliore. Pour $D/r_o=7$, l'utilisation des trois images de défocalisation 0, 2π , 4π radian fournit une erreur moyenne de $\lambda/15$. L'erreur moyenne de reconstruction est encore ici plus importante que dans le cas point-à-point, où dans une telle configuration l'erreur est de $\lambda/40$.

• **Cas « fortes » aberrations** Le traitement des aberrations qui ont été dénommées « moyennes » a montré la nécessité d'augmenter le nombre d'images défocalisées utilisées pour inverser le problème afin d'améliorer la qualité d'estimation de la phase. Les dernières forces de turbulence auxquelles nous nous sommes intéressés sont présentées ici et regroupées sous l'appellation « fortes »; D/r_o allant de 8 à 10. Les écart types théoriques (comme précédemment hors tilts) qui leur sont associés sont les suivantes : $\lambda/3$ pour $D/r_o=8$, $\lambda/2,7$ pour $D/r_o=9$ et $\lambda/2,5$ pour $D/r_o=10$.

A partir de la phase point-à-point A D/r_o égal à 8, l'estimation des vingt phases turbulentes à partir d'une image focalisée et de deux images défocalisées de 2π et 4π radians par rapport au plan focal conduit à une erreur moyenne sur l'estimée de la phase de $\lambda/17$. La qualité d'estimation des cinq phases ayant des écart types spatiaux les plus forts est inférieure à $\lambda/10$. Comme nous l'avons fait dans le paragraphe portant sur les aberrations moyennes, on propose, tout d'abord, pour réduire l'erreur d'estimation, d'augmenter la quantité de défocalisation introduite entre les images. L'amplitude de défocalisation de la deuxième image est à présent de 4π radians et celle de la troisième, 6π radians. L'estimation des phases à partir de ces trois images permet d'améliorer légèrement la qualité des phases reconstruites : $\lambda/20$ en moyenne.

Jusqu'à présent, le point de départ de la minimisation correspondait à une phase initiale nulle dans la pupille. Afin d'améliorer l'estimation, rendue difficile par la présence de minima locaux, on modifie le point d'initialisation de la minimisation pour se rapprocher du minimum global du critère. Pour cela, nous nous sommes inspirés de l'article de (Jefferies et al., 2002) qui propose d'utiliser comme point de départ de la minimisation de la phase point-à-point, le résultat d'une première estimation dans laquelle la phase est modélisée par trois polynômes de Zernike. Ces trois modes correspondent aux composantes de défocalisation et d'astigmatismes a_4 , a_5 et a_6 . Le point de départ de cette première estimation est donné par la phase nulle. L'utilisation de cette initialisation et des trois images 0, 4π , 6π radians permet d'atteindre, toujours pour $D/r_o=8$, une erreur moyenne sur les vingt réalisations de $\lambda/45$. Un exemple de restauration est présenté sur les figures 7.12 et 7.13. La figure 7.12

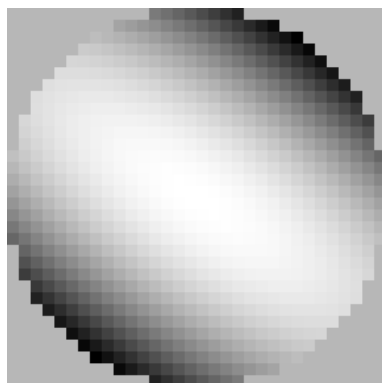


FIG. 7.12 – Phase reconstruite à partir des trois coefficients a_4 , a_5 et a_6 et utilisée comme point de départ de la minimisation de la phase point-à-point dans la pupille pour la simulation présentée figure 7.13.

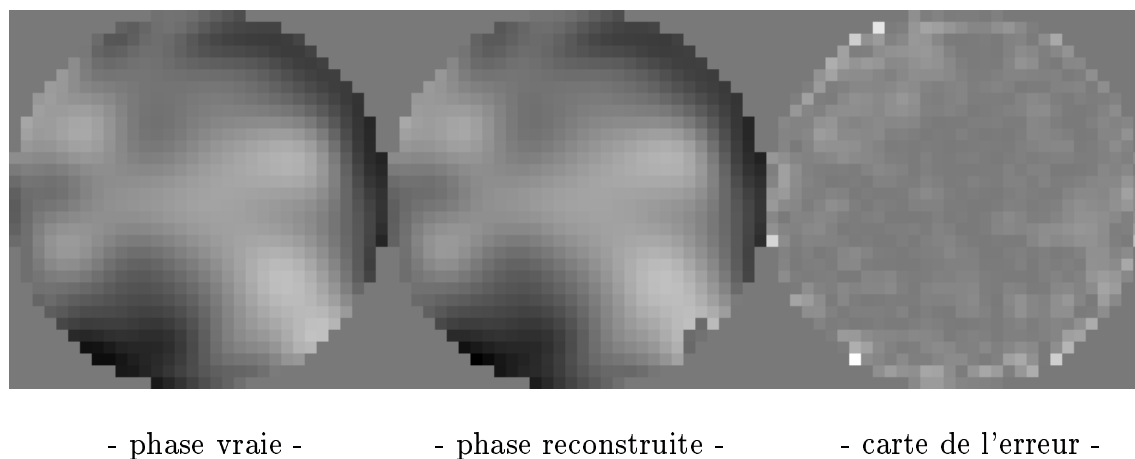


FIG. 7.13 – Exemple de restauration d'une phase turbulente ($D/r_o=8$) sur objet étendu (les cartes de la phase vraie et de celle reconstruite sont à la même échelle, la carte de l'erreur a été multipliée par 10). L'écart type spatial de la phase vraie est de $\lambda/2,2$ et son amplitude crête-à-crête de $2,2\pi$ radian. L'erreur de reconstruction vaut $\lambda/40$.

présente la phase reconstruite à partir des trois coefficients a_4 , a_5 et a_6 , c'est cette phase qui sert d'initialisation à la reconstruction de la phase point-à-point dans la pupille. Pour $D/r_o=9$, l'estimation des phases turbulentes à partir d'une première estimée par Zernike et l'utilisation des trois images 0 , 4π , 6π conduit à une erreur d'estimation moyenne de $\lambda/30$. On peut encore améliorer l'estimation en ajoutant un peu plus de défocalisation à la troisième image (en passant à une amplitude crête-à-crête de 8π radians), l'estimation conduit alors à une erreur moyenne de $\lambda/45$. Enfin pour $D/r_o=10$, en utilisant l'initialisation par Zernike et trois images de défocalisation 0 , 4π , 8π radians la qualité de restauration est bonne, en moyenne $\lambda/40$.

A partir des Zernike Les résultats d'estimation pour les phases fortement turbulentes à partir des Zernike conduisent aux mêmes constatations que pour la paramétrisation point-à-point. L'utilisation de trois images de défocalisation respectives 0 , 4π , 8π radians et d'une initialisation préalable de la phase par le résultat de l'estimation de ϕ sur les trois coefficients a_4 , a_5 et a_6 conduit à une estimation de bonne qualité de toutes les phases étudiées pour D/r_o égal à 8 et 9 . Pour une force de turbulence de 10 , la sous-modélisation intentionnelle que nous avons employées (90 modes sont recherchés alors que la phase vraie est modélisée sur 105) entraîne une dégradation de l'estimation de la phase qui devient importante pour $D/r_o=10$: sur 90 Zernike, l'erreur d'estimation est de $\lambda/16$, alors qu'elle n'est que de $\lambda/26$ pour une phase reconstruite sur le même nombre de polynômes que ceux ayant servis à la création de la phase vraie. Plus la force de la turbulence augmente, plus la valeur des coefficients supérieurs à 90 est importante et donc plus le fait de ne pas les prendre en compte dans la phase recherchée devient pénalisant. On remarque encore une fois que même dans le cas très favorable où les coefficients reconstruits correspondent exactement à ceux ayant servis à la simulation du problème direct, l'erreur est en moyenne plus forte dans le cas de la reconstruction à partir des Zernike qu'en point-à-point.

7.5.3 En résumé

Les principaux résultats concernant les valeurs moyennes (sur 20 réalisations de turbulence) de l'erreur d'estimation de la phase modulo 2π obtenue en fonction du niveau de turbulence sont regroupés dans le tableau 7.1. Ces estimées correspondent d'une part à une paramétrisation de la phase sur la base des Zernike et d'autre part à une recherche point-à-point de la phase. Les conditions d'estimation ont été indiquées dans la première colonne du tableau. Pour des conditions d'estimations identiques, c'est toujours la méthode point-à-point qui fournit de meilleures estimées de la phase modulo 2π que l'estimation sur la base tronquée des Zernike. Nous avons, en pratique, constaté que le critère MAP associé à la reconstruction de la phase par Zernike présente de très nombreux minima locaux qui rendent très incertain l'issue de la minimisation.

7.6 Conclusion

Le problème de la reconstruction de phases de forte amplitude par diversité de phase a été abordé très récemment dans la littérature. Cette estimation est rendue difficile par l'absence d'information dans les images autre que sur la phase modulo 2π . Nous avons proposé, dans ce chapitre, une première étude des méthodes de restauration possibles. Après avoir passé en revue les différentes façons d'aborder cette restauration, nous avons développé une méthode

		$D/r_o = 5$	$D/r_o = 7$	$D/r_o = 10$
deux images	point-à-point	$\lambda/25$	$\lambda/10$	\times
	Zernike	$\lambda/15$	$\lambda/8$	\times
trois images	point-à-point	\times	$\lambda/40$	$\lambda/16$
	Zernike	\times	$\lambda/15$	$\lambda/11$
trois images + init par Zernike	point-à-point	\times	\times	$\lambda/40$
	Zernike	\times	\times	$\lambda/16$

TAB. 7.1 – Valeurs moyennes (sur 20 réalisations de turbulence) de l'erreur d'estimation de la phase modulo 2π obtenue à partir d'une paramétrisation de la phase sur la base des Zernike ainsi que par une recherche point-à-point de la phase dans la pupille pour des forces de turbulence (D/r_o) de 5, 7 et 10. Les conditions d'estimations (nombre d'images utilisées et initialisation employée) sont indiquées dans la première colonne du tableau.

originale de restauration de la phase modulo 2π point-à-point dans la pupille. L'utilisation de la base des indicatrices est justifiée par le souhait de ne pas créer de minima locaux. Le choix de reconstruire la phase à 2π près est motivée par le fait que d'une part une partie des applications visées (la restauration d'image *a posteriori*) ne requiert pas le déroulage de la phase et d'autre part que les problèmes de déroulage de phase sont difficiles à résoudre. Un terme de régularisation adapté à cette estimation (c'est-à-dire ne créant pas de minima locaux supplémentaires) a été proposé. Une étude des performances de cette méthode sur la restauration de différentes phases fortement aberrantes a été faite. Nous avons, en parallèle, étudié les performances d'une méthode d'estimation plus classique basée sur une recherche de la phase via sa décomposition modale tronquée sur la base des Zernike. Nous avons mis en évidence, pour les deux méthodes, la nécessité d'utiliser plus d'une image défocalisée lorsque l'amplitude des aberrations devient importante afin d'assurer une bonne reconstruction de la phase : dans le cas de nos simulations (pour une phase turbulente) lorsque D/r_o est supérieur à 5. Enfin pour des phases encore plus turbulentes ($D/r_o > 7$), l'utilisation comme point de départ de la minimisation de la phase point-à-point ou décomposée sur un grand nombre de modes, du résultat d'une première estimation de la phase décomposée sur un très petit nombre de coefficients de Zernike rend plus robuste l'estimation en évitant certains minima locaux présentés par les critères. La comparaison des deux méthodes de restauration montre que l'estimation point-à-point fournit de meilleures estimées de la phase modulo 2π que la reconstruction sur la base des Zernike. Ceci est dû à la présence d'un moins grand nombre de minima locaux dans le critère MAP associé à la reconstruction point-à-point que dans celui associé à l'estimation par Zernike. On peut d'ailleurs, dans le cas du point-à-point, espérer faire disparaître ces minima locaux situés autour des minima globaux du critère en améliorant le terme de régularisation employé (en utilisant par exemple les connaissances statistiques que l'on a sur les phases turbulentes). On peut sans doute également améliorer la qualité des estimées en introduisant une régularisation non quadratique pénalisant moins fortement les fortes variations inter-pixels (afin d'améliorer la reconstruction des pixels en bord de pupille notamment).

Conclusion

Bilan

La diversité de phase est une technique qui a connu un essor important ces dix dernières années où elle a été mis en œuvre dans différentes applications : observation du soleil, co-phasage de télescopes segmentés, analyse de surface d'onde pour l'optique active, etc. Si la diversité de phase est très simple optiquement à mettre en œuvre, elle requiert la résolution d'un problème inverse difficile consistant à extraire d'images focalisée et défocalisée(s) la grandeur d'intérêt qu'est l'objet observé ou l'ensemble des aberrations affectant le système. Peu de travaux spécifiques se sont intéressés à l'étude et à l'amélioration des méthodes d'estimation permettant d'extraire les inconnues, des images. Une partie importante du travail de cette thèse a porté sur l'étude de méthodes d'estimation et notamment le développement d'une technique novatrice dans le cadre de la diversité de phase.

Une première étude originale de la méthode classique fondée sur une estimation conjointe des aberrations et de l'objet a montré ses mauvais comportements asymptotiques et le problème du réglage des hyperparamètres introduits par les termes d'*a priori* : impossibilité de régler leurs valeurs de façon non supervisée, inadéquation entre les hyperparamètres optimaux vis-à-vis de l'objet et ceux optimaux vis-à-vis des aberrations. Cela a permis également de montrer qu'en l'absence de terme de régularisation spécifique sur l'objet, bien que celui-ci soit très mal reconstruit, la qualité des aberrations estimées conjointement est très bonne à faible bruit, expliquant ainsi l'absence de régularisation sur l'objet fréquemment rencontrée dans la littérature s'intéressant à la reconstruction de front d'onde. Un estimateur marginal des seules aberrations a été proposé afin de s'affranchir des problèmes présentés par l'estimateur joint. Après avoir mené à bien les calculs analytiques permettant l'implémentation de cette méthode, nous avons montré ses bonnes propriétés asymptotiques et développé une méthode d'estimation non supervisée des hyperparamètres. La comparaison des performances sur l'estimation des aberrations obtenues avec chacune de ses méthodes a été faite sur données simulées et validée expérimentalement. Elle a montré que l'estimateur joint fournit des résultats identiques ou légèrement meilleurs que ceux obtenus par l'estimateur marginal dans le domaine des faibles bruits mais que la méthode marginale s'avère plus performante pour des niveaux de bruits plus importants. De plus, la méthode marginale est asymptotiquement exempte de minima locaux dont la présence nuit parfois à la méthode jointe. L'importance des écarts entre chacune de ces méthodes varie selon l'objet observé. Ces constatations ont été confirmées par les résultats expérimentaux. Enfin nous avons comparé plusieurs méthodes de restauration de l'objet à partir des deux méthodes jointe et marginale. Pour chaque estimateur, une méthode de reconstruction de l'objet a été proposée et validée sur simulations.

Nous nous sommes également intéressés au problème spécifique de l'estimation des fortes aberrations. Nous avons présenté les différentes stratégies possibles permettant d'effectuer cette estimation : estimation et déroulage simultanés ou estimation de la phase modulo 2π

puis éventuellement déroulage. Le déroulage de la phase étant difficile et surtout pas toujours nécessaire (notamment pour la restauration d'images *a posteriori*) nous avons choisi de développer une méthode originale permettant, sur objet étendu, d'estimer la phase modulo 2π . Nous avons montré, sur simulations, les bonnes performances de cette méthode et mis en évidence la nécessité d'utiliser au moins deux images défocalisées lors de l'estimation de phases turbulentes de forte amplitude. Le critère associé à une recherche de la phase à 2π près que nous avons proposé, présente moins de minima locaux qu'un critère lié à une recherche de la phase dépliée. Nous avons montré, à partir de simulations, que notre méthode était plus robuste quant à l'estimation de la phase modulo 2π qu'une méthode plus classique basée sur une décomposition de la phase sur la base des Zernike.

L'ensemble de ces travaux a été complété par une mise en œuvre expérimentale de la méthode jointe pour l'estimation des aberrations d'un système astronomique NAOS-CONICA. Cette étude a permis de démontrer la très grande adaptabilité de la méthode à un problème concret et les bons résultats qu'elle permet d'obtenir. Cette étude nous a amenés à étudier l'influence de différentes sources d'incertitudes sur la qualité d'estimation des aberrations. Elle a mis en évidence la robustesse de la méthode à un grand nombre d'écart à son implémentation idéale.

Perspectives

La caractérisation des défauts statiques des futurs systèmes d'optique adaptative à très haut Strehl sera indispensable afin d'atteindre les très hautes performances attendues de ces systèmes (et nécessaire pour la recherche d'exo-planètes par exemple). La démonstration sur l'instrument NAOS-CONICA de la facilité de mise en œuvre de la diversité de phase et des bonnes performances obtenues sur l'estimation des défauts des dichroïques et de la caméra, ouvre des perspectives intéressantes quant à l'utilisation de cette technique pour les futurs instruments.

Les résultats présentés dans cette thèse ouvrent de nombreuses perspectives à plus long terme. Des études sont encore à mener sur la méthode marginale que nous avons développée afin d'améliorer ses performances. Notamment, il pourrait être étudié la pertinence d'introduire un autre modèle d'*a priori* objet lors de la marginalisation prenant en compte, par exemple, la présence de bords francs dans l'objet. Il serait intéressant également, afin de compléter sa caractérisation, d'étudier la propagation du bruit, tout comme il a été fait pour la méthode conjointe.

Jusqu'à présent la diversité de phase a surtout été utilisée dans des applications où la phase était faiblement aberrante. Il conviendrait, de façon générale, d'étudier précisément les possibilités offertes par cette technique plan focal dans la reconstruction de fortes aberrations. Notre approche basée sur la recherche point-à-point de la phase à 2π près s'est avérée prometteuse. Elle reste cependant encore à améliorer, notamment en modifiant le terme de régularisation pour pénaliser moins fortement les fortes variations inter-pixels (hors sauts de 2π) ou pour intégrer les connaissances statistiques que l'on possède sur les phases turbulentes.

Les travaux réalisés ouvrent la voie à de nouvelles applications de la diversité de phase : l'apport de la méthode marginale dans d'autres applications que celles présentées dans cette thèse reste à poursuivre, notamment pour le cophasage de télescopes segmentés (Baron et al., 2002) ou pour la restauration de front d'onde et d'objet en imagerie turbulente anisoplanétique.

Annexe A

Domaine de validité de l'expression de la réponse impulsionnelle d'un système imageur

A.1 Introduction

On propose, dans cette annexe, d'étudier le domaine le long de l'axe optique d'un système imageur où la réponse impulsionnelle (RI) du système peut être approchée par le module carré de la transformée de Fourier inverse du champ électromagnétique dans la pupille. Cette étude est motivée par le fait que la technique étudiée dans cette thèse, la diversité de phase requiert l'acquisition d'images défocalisées. On souhaite donc s'assurer que les plans de mesures de ces images défocalisées restent dans le domaine de validité de l'expression de la RI établie au chapitre 1.

A.2 Expression de la réponse impulsionnelle d'un système imageur et domaine de validité

Le rôle d'un système optique de focale f est de transformer une onde plane dans le plan de la pupille en une onde sphérique de centre F (on a supposé pour simplifier que la pupille est placée dans le plan principal). La transmission pupillaire d'un tel système s'écrit donc :

$$T(X,Y) = \exp(jk(f - \sqrt{f^2 - (X^2 + Y^2)})) = \exp(jk\delta(X,Y)) \quad (\text{A.1})$$

(Afin de mener à bien cette étude, il a été nécessaire d'abandonner l'expression approchée 1.9 donnant la transmission d'une pupille qui a été utilisée dans le chapitre 1; cette expression est une approximation parabolique valable dans le cas de systèmes faiblement ouverts.)

En utilisant l'équation (1.3), l'approximation paraxiale et en introduisant la transmission pupillaire, l'amplitude complexe au point z s'écrit :

$$U(x,y) = \frac{1}{j\lambda z} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U'_s(X,Y) \exp[jk(\delta(X,Y) - r)] dXdY \quad (\text{A.2})$$

A.2.1 Au foyer

On se place au foyer c'est à dire $z = f$. On peut alors faire un développement limité de r et de $\delta(X,Y)$ autour de f et approximer les termes d'amplitude (pour l'expression de r ,

voir la relation 1.4):

$$\begin{aligned}
\delta(X,Y) - r &\approx f - f\left(1 - \frac{X^2 + Y^2}{2f^2} + \frac{(X^2 + Y^2)^2}{8f^4} + \dots\right) \\
&\quad - f\left(1 + \frac{(x - X)^2 + (y - Y)^2}{2f^2} - \frac{[(x - X)^2 + (y - Y)^2]^2}{8f^4} + \dots\right) \\
&\approx -f - \frac{x^2 + y^2}{2f} + \frac{xX + yY}{f} \\
&\quad + \frac{[(x - X)^2 + (y - Y)^2]^2 - (X^2 + Y^2)^2}{8f^3} + \dots
\end{aligned} \tag{A.3}$$

Ce développement appelle quelques remarques :

- le terme de Fresnel $jk\frac{X^2+Y^2}{2f}$ a disparu, compensé par la transmission pupillaire du système optique,
- le terme $jkf + jk\frac{x^2+y^2}{2f}$ représente le déphasage entre plan pupillaire et le point d'intérêt : il ne joue pas de rôle dans le calcul de la réponse impulsionnelle en éclairage incohérent,
- le terme d'erreur d'ordre 4 contient les termes d'aberrations du troisième ordre (coma, astigmatisme et distorsion). Ce terme peut être négligé dans l'intégrale si il reste très petit devant 2π quelques soient les valeurs prises par x , y , X et Y . Comme l'ordre de grandeur de la valeur maximum prise par X et Y est donné par le rayon de la pupille R alors que celui de x et y est donné par la dimension du domaine où il est nécessaire d'estimer la RI (pour un système à haute résolution, cette dimension est de l'ordre de $\frac{3\lambda f}{2R}$), l'ordre de grandeur de l'expression $[(x - X)^2 + (y - Y)^2]^2 - (X^2 + Y^2)^2$ est donné par les termes de la forme xX^3 :

$$k \frac{[(x - X)^2 + (y - Y)^2]^2 - (X^2 + Y^2)^2}{8f^3} \approx \frac{2\pi}{\lambda} \frac{-4\frac{3\lambda f}{2R}R^3}{8f^3} = \frac{-3\pi R^2}{2f^2} = \frac{3\pi}{2}\alpha^2 \tag{A.4}$$

où α représente la demi-ouverture numérique de sortie.

Ainsi on retrouve bien l'expression (1.8) établie au chapitre 1 : l'amplitude complexe du champ électromagnétique au foyer d'un système imageur peut être approché par la transformée de Fourier du champ dans la pupille **si l'ouverture de sortie** $\alpha^2 \ll 1$.

A.2.2 Validité le long de l'axe optique

Intéressons nous à présent à la validité de cette expression le long de l'axe optique.

Soit $M(x,y,z_f)$ le point d'observation où z_f représente la coordonnée axiale par rapport au plan focal. On peut récrire le développement de r en se limitant au cas sur l'axe ($x = y = 0$) :

$$\begin{aligned}
r &= \sqrt{(f + z_f)^2 + X^2 + Y^2} = f\sqrt{1 + 2\frac{z_f}{f} + \frac{z_f^2 + X^2 + Y^2}{f^2}} \\
&\approx f\left(1 + \frac{z_f}{f} + \frac{z_f^2 + X^2 + Y^2}{2f^2} - \frac{(2fz_f + z_f^2 + X^2 + Y^2)^2}{8f^4} + \dots\right) \\
&\approx f\left(1 + \frac{z_f}{f} + \frac{z_f^2 + X^2 + Y^2}{2f^2} \right. \\
&\quad \left. - \frac{1}{4f^4}(2fz_f^3 + 2fz_f(X^2 + Y^2) + z_f^2(X^2 + Y^2)) - \frac{z_f^4 + (X^2 + Y^2)^2}{8f^4} + \dots\right)
\end{aligned} \tag{A.5}$$

Ainsi (cf. A.1) :

$$\begin{aligned}
\delta(X,Y) - r &\approx f - f\left(1 - \frac{X^2 + Y^2}{2f^2} + \frac{(X^2 + Y^2)^2}{8f^4} + \dots\right) \\
&\quad - f\left(1 + \frac{z_f}{f} + \frac{X^2 + Y^2}{2f^2}\right) \\
&\quad - \frac{1}{4f^4}(2fz_f^3 + 2fz_f(X^2 + Y^2) + z_f^2(X^2 + Y^2)) \\
&\quad - \frac{z_f^4 + (X^2 + Y^2)^2}{8f^4} + \dots \\
&\approx -f\left(1 + \underbrace{\frac{z_f}{f} - \frac{z_f^3}{2f^3} - \frac{z_f^4}{8f^4}}_{\text{déphasage constant}} - \underbrace{\frac{z_f(X^2 + Y^2)}{2f^3}}_{\text{défocalisation}}\right) \\
&\quad - \underbrace{\frac{z_f^2(X^2 + Y^2)}{4f^4} + \dots}_{\text{termes d'erreur}}
\end{aligned} \tag{A.6}$$

Remarques

- le premier groupe de termes fonctions de z_f induit une phase constante dans l'intégrale : il n'a pas d'influence sur la réponse impulsionnelle en imagerie incohérente,
- le second groupe peut être interprété comme la défocalisation créée par déplacement suivant z_f (en continuant le DL, on fait apparaître d'autres termes de forme $\frac{z_f(X^2+Y^2)^{2n}}{f^{2n}}$ qui peuvent être interprétés de la même façon),
- le terme suivant en $\frac{z_f^2(X^2+Y^2)}{f^2}$ est le premier terme qui ne peut pas être relié à la défocalisation créée. C'est le terme d'ordre le plus faible à la fois vis-à-vis de $\frac{z_f}{f}$ et de $\frac{X^2+Y^2}{f^2}$. On le choisit donc pour évaluer l'ordre de grandeur de l'erreur commise.

Pour que l'erreur sur la phase reste négligeable dans l'intégrale, il faut que $k\frac{z_f^2(X^2+Y^2)}{f^2} \ll 2\pi$, soit :

$$z_f \ll \frac{2\sqrt{\lambda f}}{\alpha} \tag{A.7}$$

Un déplacement axial d'amplitude z_f correspond à une défocalisation crête-à-crête au bord de la pupille de $\delta_d = \frac{z_f R^2}{2f}$. La condition A.7 peut donc se traduire par :

$$\delta_d \ll \frac{R}{\sqrt{\lambda f}} \times \lambda \tag{A.8}$$

Application aux systèmes étudiés dans la thèse :

- CONICA ($\lambda=2 \mu\text{m}$, $f/2R=15$, $2R=8 \text{ m}$)
On obtient $z_f \ll 0,2 \text{ m}$ et $\delta_d \ll 260\lambda$
- banc « optique active » ($\lambda=0,633 \mu\text{m}$, $f = 0,8 \text{ m}$, $2R=0,028 \text{ m}$)
On obtient $z_f \ll 0,08 \text{ m}$ et $\delta_d \ll 20\lambda$

Ces distances de défocalisations limites sont très supérieures aux défocalisations que nous introduirons pour créer la diversité de phase. L'expression approchée de la RI établie au chapitre 1 reste donc valable pour tous les plans de mesures que nous considérerons.

Annexe B

Indétermination sur les aberrations estimées à partir d'une seule image

Cette annexe montre que le signe de la partie paire de la phase ne peut pas être déduit de la seule image plan focale. Dans le cas d'une décomposition sur la base des Zernike, c'est le signe global des coefficients d'ordre radial pair qui est indéterminé. Enfin on montre que l'utilisation d'une deuxième image défocalisée permet de lever l'incertitude.

B.1 Décomposition de la phase

Toute fonction (et donc la phase ϕ) peut s'écrire, de manière unique, comme la somme d'une fonction paire ϕ_p et d'une fonction impaire ϕ_i :

$$\phi(x,y) = \phi_p(x,y) + \phi_i(x,y) \quad (\text{B.1})$$

$$\text{avec} \quad \begin{cases} \phi_p(-x, -y) = \phi_p(x,y) \\ \phi_p(x,y) = \frac{\phi(x,y) + \phi(-x,-y)}{2} \end{cases} \quad \text{et} \quad \begin{cases} \phi_i(-x, -y) = -\phi_i(x,y) \\ \phi_i(x,y) = \frac{\phi(x,y) - \phi(-x,-y)}{2} \end{cases}$$

B.2 Propriété de l'autocorrélation

Par définition, l'autocorrélation de la fonction f est donnée par :

$$\Gamma_f(x) \triangleq \int f^*(t)f(t+x)dt \quad (\text{B.2})$$

On pose $t' = -t - x$, on a donc

$$\Gamma_f(x) = \int f^*(-t' - x)f(-t')dt \quad (\text{B.3})$$

Soit $f'(t) = f^*(-t)$, on voit que

$$\Gamma_f = \Gamma_{f'} \quad (\text{B.4})$$

B.3 Indétermination sur la phase

Soit A l'amplitude dans la pupille,

$$A(t) = P(t)e^{i\phi(t)} = P(t)e^{i[\phi_p(t)+\phi_i(t)]} \quad (\text{B.5})$$

On définit

$$A'(t) = A^*(-t) = P^*(-t)e^{-i\phi_p(t)+i\phi_i(t)} \quad (\text{B.6})$$

Or $P(t) \in \mathfrak{R}^+$ d'où $P^*(t) = P(t)$. Si de plus P est à symétrie centrale $P(-t) = P(t)$, on a

$$A'(t) = P(t)e^{-i\phi_p(t)+i\phi_i(t)} \quad (\text{B.7})$$

En utilisant la propriété de l'autocorrelation $\Gamma_A = \Gamma_{A'}$, on en déduit que $\phi = -\phi_p + \phi_i$ et $\phi = \phi_p + \phi_i$ donne la même FTO.

On ne peut donc pas déduire, à partir d'une seule image, le signe de la partie paire de la phase.

B.4 Indétermination sur la décomposition de la phase sur les Zernike

Lorsque la phase est décomposée sur la base des Zernike, cette indétermination conduit à ne pas pouvoir connaître le signe global des coefficients d'ordre radial (ou azimutal) pair. En effet :

$$\phi(x,y) = \sum_{i=1}^{+\infty} a_i Z_i(x,y) \quad \text{avec} \quad Z_i(x,y) = R_n^m(\sqrt{(x^2+y^2)})_{[\sin]}^{[\cos]}(m\theta(x,y)) \quad (\text{B.8})$$

D'après sa définition, la fonction R_n^m est une fonction paire, donc la parité de la fonction Z_i dépend de celle de $_{[\sin]}^{[\cos]}(m\theta(x,y))$. θ est l'angle polaire associé à (x,y) d'où $m\theta(-x, -y) = m\theta(x,y) + m\pi$ et donc Z_i est paire lorsque (n,m) sont paires et impaires lorsque (n,m) sont impaires (rappel : n et m ont forcément la même parité). La phase se décompose en la somme d'une fonction paire ϕ_p et impaire ϕ_i telle que :

$$\phi_p = \sum_{i=1; m \text{ pair}}^{+\infty} a_{i; m \text{ pair}} Z_{i; m \text{ pair}}(x,y) \quad (\text{B.9})$$

$$\phi_i = \sum_{i=1; m \text{ impair}}^{+\infty} a_{i; m \text{ impair}} Z_{i; m \text{ impair}}(x,y) \quad (\text{B.10})$$

Ainsi la paire $[\{a_{i; m \text{ pair}}\}, \{a_{i; m \text{ impair}}\}]$ et $[-\{a_{i; m \text{ pair}}\}, \{a_{i; m \text{ impair}}\}]$ donnent la même FTO.

B.5 Utilisation d'une deuxième image défocalisée

Avec une seule image, on ne peut reconstruire les coefficients d'ordre radial pair qu'à un facteur multiplicatif près valant $+1$ ou -1 . A présent, on considère deux images : une plan focal, l'autre défocalisée. Soient a_{if} les coefficients associés au plan focal et a_{id} ceux

associés à l'image défocalisée. Les deux ensembles de a_i diffèrent uniquement pour a_4 qui correspond à la défocalisation : $|a_{4d}| = |a_{4f} + \delta a_4|$ avec δa_4 la défocalisation connue ajoutée à la deuxième image. Grâce à l'ajout de ce coefficient qui est d'ordre radial pair, on peut donc déterminer le signe du coefficient multiplicatif et lever l'indétermination. Remarquons que n'importe quelle « diversité » ne permet pas de lever l'ambiguïté, l'aberration introduite doit être d'ordre radial (ou azimutal) pair.

Annexe C

Tableau récapitulatif de quelques filtres de CONICA - Caméra C50S

Pour chaque filtre, ont été reporté: le nom (colonne 1), la longueur d'onde centrale (colonne 2), la valeur de l'échantillonnage théorique associée (colonne 3) et l'amplitude de défocalisation crête-à-crête introduite par le trou positionné à 1mm, 2mm et 4 mm du plan focal d'entrée de la caméra.

Filtre (μm)	λ_c (μm)	f_c/f_p	défoc c-à-c (radian) pour		
			1mm	2mm	4mm
NB1.04	1.040	2.02	1.07π	2.14π	4.27π
HeI	1.083	2.11	1.03π	2.05π	4.10π
P_γ	1.094	2.13	1.02π	2.03π	4.06π
OII	1.237	2.41	0.90π	1.80π	3.59π
FeII	1.257	2.45	0.88π	1.77π	3.54π
P_β	1.282	2.50	0.87π	1.73π	3.47π
FeII	1.644	3.20	0.68π	1.35π	2.70π
H2(1-0)S(7)	1.748	3.40	0.64π	1.27π	2.54π
NB2.00	2.000	3.89	0.56π	1.11π	2.22π
NB2.06	2.060	4.00	0.54π	1.08π	2.16π
NB2.09	2.090	4.07	0.53π	1.06π	2.13π
H2(1-0)S(1)	2.122	4.13	0.52π	1.05π	2.09π
NB2.15	2.150	4.18	0.52π	1.03π	2.07π
Br_γ	2.166	4.21	0.51π	1.03π	2.05π
NB2.18	2.180	4.24	0.51π	1.02π	2.04π
NB2.21	2.210	4.30	0.50π	1.01π	2.01π
NB2.24	2.240	4.36	0.50π	0.99π	1.98π
NB2.27	2.270	4.42	0.49π	0.98π	1.96π
NB2.33	2.330	4.53	0.48π	0.95π	1.91π
NB2.39	2.390	4.65	0.46π	0.93π	1.86π
NB2.42	2.420	4.71	0.46π	0.92π	1.84π
NB2.45	2.450	4.77	0.45π	0.91π	1.81π

Annexe D

Propriétés matricielles

D.1 Lemme d'inversion des matrices partitionnées

Soit A une matrice partitionnée de forme:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (\text{D.1})$$

Si les matrices A , A_{11} et A_{22} sont non singulières, l'inverse de A s'écrit :

$$A^{-1} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -A_{11}^{-1}A_{12}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \\ -A_{22}^{-1}A_{21}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \quad (\text{D.2})$$

D.2 Le déterminant d'une matrice bloc

Soit Δ une matrice bloc,

$$\text{Soit } \Delta = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad (\text{D.3})$$

son déterminant est donné par :

$$\det(\Delta) = \det(A) \det(D - CA^{-1}B) \quad (\text{D.4})$$

Annexe E

Expression des gradients du critère marginal par rapport aux hyperparamètres

E.1 A partir du critère marginal J_{MAP}

En utilisant le relation 4.18 et en posant $S_o = kS'_o$, le critère marginal s'écrit :

$$\begin{aligned}
 J_{\text{MAP}}(\mathbf{a}, k, \mu, v_o, p) &= \sum_v \ln \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right) \\
 &+ N^2 \ln \mu + \sum_v \ln S'_o(v) + 2N^2 \ln k + \ln \det(R_a) + \mathbf{a}^t R_a^{-1} \mathbf{a} \\
 &+ \frac{1}{k\mu} \sum_v \frac{\left(|\tilde{i}_d(v) \tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v) \tilde{h}_d(\mathbf{a}, v)|^2 \right) + \frac{\mu}{S'_o(v)} \left(|\tilde{h}_f(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right)}{\left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)}
 \end{aligned} \tag{E.1}$$

E.1.1 Expression du gradient par rapport à μ

$$\begin{aligned}
 \frac{\partial J_{\text{MAP}}}{\partial \mu} &= \frac{N^2}{\mu} + \sum_v \left[\frac{1}{S'_o(v) \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\
 &\quad \left. - \frac{\frac{1}{\mu^2} \left(|\tilde{i}_d(v) \tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v) \tilde{h}_d(\mathbf{a}, v)|^2 \right) + \frac{\mu}{S'_o(v)} \left(|\tilde{h}_f(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right)}{\left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right]
 \end{aligned} \tag{E.2}$$

E.1.2 Expression du gradient par rapport à v_o

$$\begin{aligned}
\frac{\partial J_{\text{MAP}}}{\partial v_o} = \sum_v p v_o^{p-1} & \left[-S'_o(v) + \frac{\mu}{\left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\
& + \frac{\left(|\tilde{h}_f(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right) \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 \right)}{k \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \\
& \left. - \frac{\left(\tilde{i}_d(v) \tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v) \tilde{h}_d(\mathbf{a}, v) \right)^2}{k \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right] \quad (\text{E.3})
\end{aligned}$$

E.1.3 Expression du gradient par rapport à p

$$\begin{aligned}
\frac{\partial J_{\text{MAP}}}{\partial p} = \sum_v - \left(v_o^p \ln v_o + v^p \ln v \right) & \left[-S'_o(v) + \frac{\mu}{\left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\
& + \frac{\left(|\tilde{h}_f(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_f(v)|^2 + |\tilde{h}_d(\mathbf{a}, v) \tilde{o}_m(v) - \tilde{i}_d(v)|^2 \right) \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 \right)}{k \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \\
& \left. - \frac{\left(\tilde{i}_d(v) \tilde{h}_f(\mathbf{a}, v) - \tilde{i}_f(v) \tilde{h}_d(\mathbf{a}, v) \right)^2}{k \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right] \quad (\text{E.4})
\end{aligned}$$

E.2 A partir du critère marginal alterné $J_{\text{MAP}}^{\text{alt}}$

En utilisant la relation 4.31 et en posant $S_o = kS'_o$, le critère marginal alterné s'écrit :

$$\begin{aligned}
J_{\text{MAP}}^{\text{alt}}(\mathbf{o}, \mathbf{a}, k, \mu, v_o, p) = \sum_v & \left[\ln \left(|\tilde{h}_f(\mathbf{a}, v)|^2 + |\tilde{h}_d(\mathbf{a}, v)|^2 + \frac{\mu}{S'_o(v)} \right) + \frac{|\tilde{o}(v) - \tilde{o}_m(v)|^2}{k S'_o(v)} \right] \\
& + N^2 \ln k \mu + \sum_v \ln k S'_o(v) + \ln \det(R_a) + \mathbf{a}^t R_a^{-1} \mathbf{a} \\
& + \frac{1}{k \mu} \sum_{(l, m) \in \text{FOV}} \left[\left| i_f(l, m) - TF^{-1} \left(\tilde{h}_f(\mathbf{a}, v) \tilde{o}(v) \right) \Pi_{l, m} \right|^2 + \left| i_d(l, m) - TF^{-1} \left(\tilde{h}_d(\mathbf{a}, v) \tilde{o}(v) \right) \Pi_{l, m} \right|^2 \right] \quad (\text{E.5})
\end{aligned}$$

E.2.1 Expression de \hat{k}

L'annulation du gradient de $J_{\text{MAP}}^{\text{alt}}$ par rapport à k fournit une expression analytique de \hat{k} :

$$\begin{aligned} \hat{k} &= \sum_v \frac{N^2 |\tilde{\delta}(v) - \tilde{\delta}_m(v)|^2}{S'_o(v)} \\ &+ \frac{N^2}{\mu} \sum_{(l,m) \in \text{FOV}} \left[\left| \iota_f(l,m) - TF^{-1} \left(\tilde{h}_f(\mathbf{a},v) \tilde{\delta}(v) \right) \Pi_{l,m} \right|^2 + \left| \iota_d(l,m) - TF^{-1} \left(\tilde{h}_d(\mathbf{a},v) \tilde{\delta}(v) \right) \Pi_{l,m} \right|^2 \right] \end{aligned} \quad (\text{E.6})$$

E.2.2 Expression du gradient par rapport à μ

$$\begin{aligned} \frac{\partial J_{\text{MAP}}^{\text{alt}}}{\partial \mu} &= \frac{N^2}{\mu} + \sum_v \left[\frac{1}{S'_o(v) \left(|\tilde{h}_f(\mathbf{a},v)|^2 + |\tilde{h}_d(\mathbf{a},v)|^2 \right) + \mu} \right] \\ &- \frac{1}{k\mu^2} \sum_{(l,m) \in \text{FOV}} \left[\left| \iota_f(l,m) - TF^{-1} \left(\tilde{h}_f(\mathbf{a},v) \tilde{\delta}(v) \right) \Pi_{l,m} \right|^2 + \left| \iota_d(l,m) - TF^{-1} \left(\tilde{h}_d(\mathbf{a},v) \tilde{\delta}(v) \right) \Pi_{l,m} \right|^2 \right] \end{aligned} \quad (\text{E.7})$$

E.2.3 Expression du gradient par rapport à v_o

$$\frac{\partial J_{\text{MAP}}^{\text{alt}}}{\partial v_o} = \sum_v p v_o^{p-1} \left[\frac{\mu}{\left(|\tilde{h}_f(\mathbf{a},v)|^2 + |\tilde{h}_d(\mathbf{a},v)|^2 + \frac{\mu}{S'_o(v)} \right)} + \frac{|\tilde{\delta}(v) - \tilde{\delta}_m(v)|^2}{k} - S'_o(v) \right] \quad (\text{E.8})$$

E.2.4 Expression du gradient par rapport à p

$$\frac{\partial J_{\text{MAP}}^{\text{alt}}}{\partial p} = \sum_v \left(v_o^p \ln v_o + v^p \ln v \right) \left[\frac{\mu}{\left(|\tilde{h}_f(\mathbf{a},v)|^2 + |\tilde{h}_d(\mathbf{a},v)|^2 + \frac{\mu}{S'_o(v)} \right)} + \frac{|\tilde{\delta}(v) - \tilde{\delta}_m(v)|^2}{k} - S'_o(v) \right] \quad (\text{E.9})$$

Annexe F

Extension du critère marginal à trois images

F.1 Le critère

On rappelle la relation qui lie le critère marginal au critère joint (voir équation (4.19) et (4.17))

$$J_{\text{MAP}}(\phi, \boldsymbol{\theta}) = J'_{\text{MAPJ}}(\phi, \boldsymbol{\theta}) + \frac{1}{2} \ln \det(R_I) - N^2 \ln \sigma^2 - \frac{1}{2} \sum_v \ln S_o(v) + A'' \quad (\text{F.1})$$

où A'' est une constante. Le critère J'_{MAP} peut être aisément généralisé à plus de deux images en utilisant les résultats établis dans (Paxman et al., 1992) :

$$J'_{\text{MAPJ}}(\phi) = N^2 \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) + J_{\text{regul}}(\phi) + \frac{1}{2} \sum_v \frac{\sum_p |\tilde{i}_p(v)|^2 \sum_q |\tilde{h}_q(\phi, v)|^2 - |\sum_p \tilde{h}_p^*(\phi, v) \tilde{i}_p(v)|^2 + \frac{\sigma^2}{S_o(v)} \sum_p |\tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v)|^2}{\sigma^2 \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right)} \quad (\text{F.2})$$

avec p (et q) = $\{f, d1, d2\}$ où f est l'indice lié au plan focal, $d1$ au premier plan défocalisé et $d2$ au deuxième. L'expression du critère marginal nécessite encore le calcul le $\ln \det(R_I)$ dans le cas de trois images. En utilisant la définition $R_I \triangleq E[\mathbf{I}\mathbf{I}^t] - E[\mathbf{I}]E[\mathbf{I}]^t$, on obtient l'expression :

$$R_I = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} = \begin{pmatrix} H_f R_o H_f^t + \sigma^2 I_d & H_f R_o H_{d1}^t & H_f R_o H_{d2}^t \\ H_{d1} R_o H_f^t & H_{d1} R_o H_{d1}^t + \sigma^2 I_d & H_{d1} R_o H_{d2}^t \\ H_{d2} R_o H_f^t & H_{d2} R_o H_{d1}^t & H_{d2} R_o H_{d2}^t + \sigma^2 I_d \end{pmatrix} \quad (\text{F.3})$$

Il faut calculer son déterminant. Pour cela on utilise l'algorithme de Gauss généralisé (Gantmacher, 1966) qui permet d'écrire (la matrice R_{11} étant carré et régulière) :

$$\det(R_I) = |R_{11}| \begin{vmatrix} R_{22} - R_{21} R_{11}^{-1} R_{12} & R_{23} - R_{21} R_{11}^{-1} R_{13} \\ R_{32} - R_{31} R_{11}^{-1} R_{12} & R_{33} - R_{31} R_{11}^{-1} R_{13} \end{vmatrix} \quad (\text{F.4})$$

En utilisant la formule donnant le déterminant d'une matrice bloc à quatre éléments rappelé dans l'annexe D, on peut calculer le déterminant de la matrice bloc de l'expression F.4. Le calcul de $\ln \det(R_I)$ est tout à fait similaire à celui mené dans le cas deux images (paragraphe 4.2.4), nous nous contenterons donc ici de donner le résultat final :

$$\ln \det(R_I) = 2N^2 \ln \sigma^2 + \sum_v \ln S_o(v) + \sum_v \ln \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right) \quad (\text{F.5})$$

Le critère marginal généralisé à trois images s'écrit alors :

$$\begin{aligned} J_{\text{MAP}}(\phi, \theta) &= N^2 \ln \sigma^2 + \frac{1}{2} \sum_v \ln S_o(v) + J_{\text{regul}}(\phi) \\ &+ \frac{1}{2} \sum_v \ln \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right) \\ &+ \frac{1}{2} \sum_v \frac{\sum_p |\tilde{i}_p(v)|^2 \sum_q |\tilde{h}_q(\phi, v)|^2 - \left| \sum_p \tilde{h}_p(\phi, v) \tilde{i}_p(v) \right|^2 + \frac{\sigma^2}{S_o(v)} \sum_p |\tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v)|^2}{\sigma^2 \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\sigma^2}{S_o(v)} \right)} \end{aligned} \quad (\text{F.6})$$

F.2 Le gradient par rapport à ϕ

En utilisant la relation donnant le gradient du critère joint vis-à-vis des phases qui a été établie dans (Paxman et al., 1992), on en déduit l'expression étendue de $\frac{\partial J'_{\text{MAP}}}{\partial \phi(l, m)}$ à trois images. Il reste à calculer le gradient du terme différentiel entre le critère joint et le critère marginal, celui-ci se calcule simplement et donne :

$$\frac{\partial \ln \det(R_I)}{\partial h_p} = TF^{-1} \left(\frac{\tilde{h}_p(\phi, v)}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S_o(v)} \right)} \right) \quad (\text{F.7})$$

F.3 Les hyperparamètres

Il ne reste plus qu'à calculer l'expression analytique de k dans le cas de trois images et d'étendre les gradients du critère par rapport aux hyperparamètres μ , v_o et p au cas trois images. Afin d'alléger les expressions, on pose

$$\text{num}(\phi, v) = \sum_p |\tilde{i}_p(v)|^2 \sum_q |\tilde{h}_q(\phi, v)|^2 - \left| \sum_p \tilde{h}_p^*(\phi, v) \tilde{i}_p(v) \right|^2 + \frac{\sigma^2}{S_o(v)} \sum_p \left| \tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v) \right|^2 \quad (\text{F.8})$$

L'annulation du gradient du critère par rapport à k , donne l'expression suivante pour \hat{k} (en posant $\mu = \sigma^2/k$) :

$$\hat{k} = \sum_v \frac{\text{num}(\phi, v)}{3\mu \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S_o(v)} \right)} \quad (\text{F.9})$$

Les gradients par rapport aux autres hyperparamètres se calculent à partir de l'expression F.6 et en posant ici aussi $\mu = \sigma^2/k$.

Expression du gradient par rapport à μ

$$\begin{aligned} \frac{\partial J_{\text{MAP}}}{\partial \mu} &= \frac{2N^2}{\mu} + \sum_v \left[\frac{1}{2S'_o(v) \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\ &\left. + \frac{1}{2k\mu} \left(\frac{\sum_p |\tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v)|^2}{S'_o \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} - \frac{\text{num}(\phi, v) \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{2\mu}{S'_o(v)} \right)}{\mu \left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right) \right] \end{aligned} \quad (\text{F.10})$$

Expression du gradient par rapport à v_o

$$\begin{aligned} \frac{\partial J_{\text{MAP}}}{\partial v_o} &= \sum_v p v_o^{p-1} \left[S'_o(v) - \frac{\mu}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\ &\left. - \frac{1}{2k} \left(\frac{\sum_p |\tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v)|^2}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} - \frac{\text{num}(\phi, v)}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right) \right] \end{aligned} \quad (\text{F.11})$$

Expression du gradient par rapport à p

$$\begin{aligned} \frac{\partial J_{\text{MAP}}}{\partial p} &= \sum_v (v_o^p \ln v_o + v^p \ln v) \left[S'_o(v) - \frac{\mu}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} \right. \\ &\left. - \frac{1}{2k} \left(\frac{\sum_p |\tilde{h}_p(\phi, v) \tilde{o}_m(v) - \tilde{i}_p(v)|^2}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)} - \frac{\text{num}(\phi, v)}{\left(\sum_p |\tilde{h}_p(\phi, v)|^2 + \frac{\mu}{S'_o(v)} \right)^2} \right) \right] \end{aligned} \quad (\text{F.12})$$

Annexe G

Article A&A

**Calibration of NAOS and CONICA static aberrations.
Application of the phase diversity technique.**

Cet article a été accepté pour publication dans le journal *Astronomy&Astrophysics*.

A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
missing; you have not inserted them

ASTRONOMY
AND
ASTROPHYSICS

Calibration of NAOS and CONICA static aberrations

Application of the phase diversity technique

A. Blanc^{1,2}, T. Fusco¹, M. Hartung³, L. M. Mugnier¹, and G. Rousset¹

¹ Office National d'Etudes et de Recherches Aéronautiques, Département d'Optique Théorique et Appliquée,
BP 72, F-92322 Châtillon cedex, France
E-mail: laurent.mugnier@onera.fr

² Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité, Plateau de Moulon, F-91192 Gif-sur-Yvette, France

³ Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

the date of receipt and acceptance should be inserted later

Abstract. We describe and evaluate the performance of a phase diversity wavefront sensor used to measure the static aberrations of the VLT instrument NAOS-CONICA. The main limitations of this phase diversity technique are compiled. We investigate the systematic errors due to the experimental implementation and the design restrictions. Further error sources stem from the imperfect knowledge of the system, and from limitations of the algorithm. The influence of these errors on the wavefront estimation is evaluated on numerical and experimental data. This study highlights the essential verifications and calibrations needed to obtain accurate results and gives a practical guideline for the application of a phase diversity wavefront sensor. The comprehensive calibration results and the final gain in optical performance are presented and discussed in a complementary paper (Hartung et al., this issue).

Key words: Wavefront sensing, phase diversity, active optics, adaptive optics

1. Introduction

The VLT instrument NAOS-CONICA comprised of the adaptive optics (AO) system NAOS¹ (Rousset et al., 1998; Rousset et al., 2000) and the high resolution camera CONICA² (Lenzen et al., 1998; Hartung et al., 2000), aims at providing very high quality images on one (UT4-Yepun) of the 8-m telescopes of the Cerro Paranal observatory. To achieve its specification, NAOS has been designed to provide high Strehl ratios ($\geq 70\%$) under good observing conditions. But even if the AO system ideally corrects for atmospheric turbulence, the optical train of CONICA as well as a part of the optical train of NAOS which is not sensed by the wavefront sensor also contribute to the over-

all degradation of the instrument performance. Indeed, the optical train is not free of aberrations and consequently reduces the Strehl Ratio (SR) on the images. In order to reach the ultimate performance with the instrument, it is necessary to pre-compensate these down-stream static aberrations by the AO system. This is a feature provided by NAOS which is able to introduce known static aberrations on the deformable mirror (DM) in closed loop.

In the NAOS-CONICA instrument it is not possible to introduce an additional wavefront sensor in the optical train of the imaging device CONICA to estimate the wavefront. Hence, the CONICA images themselves must be used to derive the unknown aberrations.

In this paper, we present a phase diversity (PD) (Gonçalves, 1982) approach to calibrate these unseen aberrations using a couple of focused and defocused images obtained on the CONICA detector. Because of the huge number of observation modes of NAOS-CONICA, the calibration of the unseen aberrations must be split in several parts: NAOS dichroics, i.e. the beam splitters between the wavefront sensor and imaging paths, and CONICA aberrations, i.e. the different filters and objectives. Detailed explanations of the global calibration procedures as well as a comprehensive presentation of the results are presented in (Hartung et al., this issue) (hereafter called Paper II).

In the present paper we focus on the validation of the phase diversity approach used to estimate these aberrations. In Section 2, a description of the phase diversity concept is given along with the specific algorithm used. In Section 3, a brief presentation of the experimental setup is proposed. A more complete description can be found in Paper II. In Section 4, a validation of our PD algorithm is proposed on simulated data. It allows us to assess qualitatively and quantitatively the accuracy of the aberration estimation.

Finally, we exploit the simulations and the experimental results to present the limitations and the overall per-

¹ NAOS stands for Nasmyth Adaptive Optics System

² CONICA stands for COudé Near Infrared CAMERA

formance of the aberration calibration procedure in Section 5. We list, analyze and quantify all the possible error sources which may induce a loss of accuracy in the PD results. This Section allows us to define the optimal parameters to calibrate the whole system.

In Section 6, we present a practical example and give a detailed description of the global procedure to estimate NAOS-CONICA static aberrations. This procedure is quite general but the particularities of CONICA and NAOS-dichroic aberration measurements are underlined.

2. Estimation of static aberrations by Phase Diversity

2.1. Phase diversity principle

The estimation of the aberrations from the sole focused image does not ensure the uniqueness of the solution. This indetermination is due to the relationship between the point spread function (PSF) h and the aberrated phase: a couple (ϕ, ϕ') exists such that $h(\phi) = h(\phi')$. Phase diversity (Gonsalves, 1982; Paxman et al., 1992) was proposed to add information and removes this indetermination. The idea is to collect at least one additional image, which differs from the focused one by a known phase variation. Figure 1 illustrates the phase diversity principle. One can note that there are several ways to introduce the known aberration. A possibly non comprehensive list of the possibilities is the following :

- simultaneously using a beam splitter and two detectors as presented in Figure 1,
- introducing a beam-splitter and a prism and recording the focused and defocused images on the same detector (Gates et al., 1994),
- sequentially using a translation of the detector or introducing the known aberration in the optical path (see Subsection 2.3.2). In this case, it is assumed that the aberrations and the object do not evolve between the two acquisitions (which is the case of the NAOS-CONICA static aberrations for instance).

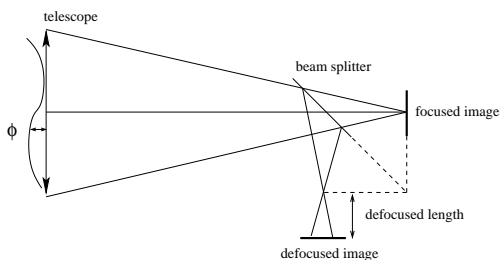


Fig. 1. Phase diversity principle

This technique has been successfully used by some authors to determine aberrations (Carreras et al., 1994;

Kendrick et al., 1994; Lee et al., 1997; Thelen et al., 1999; Löfdahl et al., 2000) and also to restore images, as in solar imaging through turbulence (Restaino, 1992; Löfdahl and Scharmer, 1994; Seldin and Paxman, 1994). It uses a low-cost, optically simple wavefront sensor which consists in the imaging camera itself, but it requires a complex numerical and iterative processing to restore the unknowns from the images.

2.2. Imaging model

In the isoplanatic patch of the imaging system, the image is the noisy convolution of the PSF h in the observation plane and the object o :

$$i(\mathbf{r}) = (h * o)(\mathbf{r}) + n(\mathbf{r}) \quad (1)$$

where $*$ denotes the convolution product, \mathbf{r} is a two-dimensional vector in the image plane and n is an additive noise. The PSF associated with the focused image is given by:

$$h_1(\mathbf{r}) = |FT^{-1}\{P(\mathbf{u}) \cdot \exp[i\phi(\mathbf{u})]\}|^2 \quad (2)$$

where \mathbf{u} is a two-dimensional vector in the pupil plane, ϕ is the unknown aberrated phase function, P is the aperture function and FT^{-1} denotes the inverse Fourier transform. The phase function is expanded on a set of Zernike polynomials. Indeed, aberrations in an optical system can be mathematically represented by Zernike polynomials (Noll, 1976).

$$\phi(\mathbf{u}) = \sum_{i=2}^k a_i Z_i(\mathbf{u}) \quad (3)$$

Theoretically, k should tend to infinity to describe any wave form, but in the particular case of static aberration estimation, the first polynomials (typically the first twenty) are enough to describe the aberrations arising from misalignment of systems made mostly of spherical surfaces or of surfaces with a modest degree of asphericity. In the following, we will note $\mathbf{a} = (a_1, \dots, a_k)^T$ the unknown aberration coefficients, where T denotes transposition.

In the defocused plane, the PSF is given by :

$$h_2(\mathbf{r}) = |FT^{-1}(P(\mathbf{u}) \cdot \exp\{i[\phi(\mathbf{u}) + \phi_d(\mathbf{u})]\})|^2 \quad (4)$$

where ϕ_d is the known diversity phase function. In our case, $\phi_d(\mathbf{u}) = a_4^d Z_4(\mathbf{u})$ where Z_4 is the defocus Zernike polynomial.

In practice, data are discrete arrays because of the spatial sampling of the images. Equation 1 takes the form:

$$i = \mathbf{H}o + n \quad (5)$$

where \mathbf{H} is the Toeplitz matrix corresponding to the convolution by h (Ekstrom and Rhoads, 1974), and where i , o and n are the discrete forms of the previous variables.

The problem is to estimate the aberration parameters \mathbf{a} (the set of a_i) from the data (focused i_1 and defocused i_2 images) and the defocused distance, without knowing the object \mathbf{o} .

2.3. Aberration estimation principle

Using a Joint Maximum A Posteriori (JMAP) approach (Little and Rubin, 1983), an estimator can be defined as

$$\begin{aligned} (\hat{\mathbf{o}}, \hat{\mathbf{a}})_{\text{JMAP}} &= \arg \max_{\mathbf{o}, \mathbf{a}} p(i_1, i_2, \mathbf{o}, \mathbf{a}) \\ &= \arg \max_{\mathbf{o}, \mathbf{a}} p(i_1 | \mathbf{o}, \mathbf{a}) p(i_2 | \mathbf{o}, \mathbf{a}) p(\mathbf{o}) p(\mathbf{a}) \end{aligned} \quad (6)$$

where $p(i_1, i_2, \mathbf{o}, \mathbf{a})$ is the joint probability density function of the data (i_1, i_2) , of the object \mathbf{o} and of the aberrations \mathbf{a} . $p(i_1 | \mathbf{o}, \mathbf{a})$ and $p(i_2 | \mathbf{o}, \mathbf{a})$ denote the likelihood of the data i_1 and i_2 , $p(\mathbf{o})$ and $p(\mathbf{a})$ are the *a priori* probability density function of \mathbf{o} and \mathbf{a} . The noise is modeled as a stationary white Gaussian noise with a variance σ^2 (the same for the two images). In our case, we do not use an explicit probability density function for \mathbf{a} ($p(\mathbf{a}) = 1$) since the regularization is brought by the small number of estimated Zernike coefficients. Consequently, this joint estimation will be called a Generalized Maximum Likelihood (GML) approach. We choose a Gaussian prior probability distribution for the object with a covariance matrix R_o . The maximization of the a posteriori probability law defined in Equation 6 is equivalent to the minimization of its neg-logarithm (J_{GML}) defined as

$$\begin{aligned} J_{\text{GML}}(\mathbf{o}, \mathbf{a}) &= \frac{1}{2\sigma^2} \|\mathbf{i}_1 - \mathbf{H}_1 \mathbf{o}\|^2 + \frac{1}{2\sigma^2} \|\mathbf{i}_2 - \mathbf{H}_2 \mathbf{o}\|^2 \\ &\quad + \frac{1}{2} \mathbf{o}^t \mathbf{R}_o^{-1} \mathbf{o} \end{aligned} \quad (7)$$

It is important to note that the derivative of $J_{\text{GML}}(\mathbf{o}, \mathbf{a})$ with respect to the object gives a closed-form expression for the object $\hat{\mathbf{o}}(\mathbf{a})$ that minimizes the criterion for a given \mathbf{a} (Paxman et al., 1992). This expression is that of a bi-frame Wiener filter:

$$\hat{\mathbf{o}}(\mathbf{a}) = (\mathbf{H}_1^t \mathbf{H}_1 + \mathbf{H}_2^t \mathbf{H}_2 + \sigma^2 \mathbf{R}_o^{-1})^{-1} (\mathbf{H}_1^t i_1 + \mathbf{H}_2^t i_2) \quad (8)$$

Substituting $\hat{\mathbf{o}}(\mathbf{a})$ into the criterion yields a new criterion that does not depend explicitly on the object:

$$J'_{\text{GML}}(\mathbf{a}) = J_{\text{GML}}(\hat{\mathbf{o}}(\mathbf{a}), \mathbf{a}) \quad (9)$$

The interest of this new criterion is its drastic reduction of the solution space.

Furthermore, using a periodic approximation (which corresponds to approximate a Toeplitz by a circulant matrix) (Demoment, 1989), the criterion can be expressed in the discrete Fourier domain:

$$\begin{aligned} J_{\text{GML}}(\mathbf{o}, \mathbf{a}) &\propto \sum_f \left[\frac{|\tilde{h}_1(f) \tilde{o}(f) - \tilde{i}_1(f)|^2}{\sigma^2} \right. \\ &\quad \left. + \frac{|\tilde{h}_2(f) \tilde{o}(f) - \tilde{i}_2(f)|^2}{\sigma^2} + \frac{|\tilde{o}(f)|^2}{S_o(f)} \right] \end{aligned} \quad (10)$$

where $\tilde{\cdot}$ stands for a Discrete Fourier Transform (DFT), f the spatial frequency and S_o the power spectral density (PSD) of the object.

Using the same periodicity approximation, $\hat{\mathbf{o}}(\mathbf{a})$ can be written as :

$$\tilde{\hat{\mathbf{o}}}(\mathbf{a}) = \frac{\tilde{h}_1^*(f) \tilde{i}_1(f) + \tilde{h}_2^*(f) \tilde{i}_2(f)}{|\tilde{h}_1(f)|^2 + |\tilde{h}_2(f)|^2 + \sigma^2 / S_o(f)}. \quad (11)$$

$\tilde{\hat{\mathbf{o}}}(\mathbf{a})$ can be introduced in Equation (10) so that the new criterion $J'_{\text{GML}}(\mathbf{a})$ only depends on the unknown phase aberrations (\mathbf{a}) and can be minimized using an iterative algorithm (for instance a conjugate gradient).

2.3.1. Object regularization

The *a priori* information required on the object consists in the choice of the object power spectral density model. In our case a focused and a defocused image are acquired with high Signal to Noise Ratio (SNR) and the object is close to be a Dirac function. Consequently the phase estimation can be obtained without any knowledge or estimation of the prior object spectral density. The dimensionless ratio $\sigma^2 / S_o(f)$ is simply set to a small arbitrary constant (10^{-6} in our case) in order to avoid numerical problems due to computer precision. Notice that, in the case of lower SNR, the use of a marginal estimator which estimates the sole aberration parameters, gives better results than this ad-hoc regularization (Blanc et al., 2000).

2.3.2. Choice of the defocus distance

The choice of the known defocus distance is essential to obtain accurate results. The RMS defocus coefficient a_4^d depends on the defocus distance d of the second image, the telescope diameter D and the focal length F through:

$$a_4^d = \frac{\pi d}{8\sqrt{3}\lambda(F/D)^2} \text{ (in radian)} \quad (12)$$

The corresponding peak-to-valley optical path Δ is equal to

$$\Delta = \frac{\sqrt{3}\lambda a_4^d}{\pi} = \frac{d}{8(F/D)^2} \quad (13)$$

In the following sections, a_4^d will be given in nanometers. For NAOS+CONICA, $F/D = 15$. It has been shown (Lee et al., 1999; Meynadier et al., 1999) that a defocus Δ equal to λ provides accurate results. That is a defocus distance equal to 4 mm for $\lambda = 2.2 \mu\text{m}$. In fact the "optimal" defocus distance depends on the object structure, the phase amplitude ϕ and the SNR on the images. In practice a large domain around this value (typically $\lambda \pm \lambda/2$) still provides accurate results (Meynadier et al., 1999).

In the case of NAOS-CONICA static aberration estimation, two different procedures are applied to introduce the defocus. They are explained briefly in Section 3 and in detail in Paper II.

3. Practical implementation

3.1. CONICA calibration

The estimation of CONICA stand-alone aberrations (objectives and filters) is obtained through the use of pinholes (diameter $10\mu\text{m}$) located at different defocused positions in the camera entrance focal plane (0, 1, 2 and 4 mm). The telescope pupil is simulated by a cold pupil placed inside CONICA. A detailed explanation of the experimental setup is given in Paper II. The defocus distances induced through the use of different pinhole pairs are summarized in Table 1.

pinhole pairs	0-1	0-2	0-4	1-4	2-4
Defocus distance (mm)	1.0	2.0	4.0	3.0	2.0

Table 1. Defocus distances induced by the use of various pinhole pairs. 0-2 and 0-4 pairs are recommended for J-H and K filters respectively.

The necessity of introducing enough diversity between the two images and the higher SNR of images obtained with the pinhole 0 (which is in the focal plane) lead us to choose the pair 0-2 for J and H filters and the pair 0-4 for K filters. Note that the use of pinholes in the entrance focal plane is not optimal for the phase diversity algorithm since

- the known aberration is not a pure defocus (a longitudinal translation in the entrance focal plane of the camera is not completely equivalent to a detector translation in the imaging focal plane) – see Subsection 5.1.5,
- the use of different pinholes may induce errors in the aberration estimation (the PD algorithm assumes that the same object is used to obtain focused and defocused images, see Subsection 5.1.4). Shape differences between two pinholes can induce phase estimation errors,
- the focused and defocused images are not at the same position on the detector and need to be re-centered – see Subsection 6.2 – since PD can not estimate relative tip-tilt greater than λ between the two images. Indeed, the phase is estimated modulo 2π (see Equations 2 and 4),
- lastly, there may be field aberrations due to the different pinhole position in the beam – see Subsection 5.1.6.

The procedure described in Subsection 2.3 allows us to estimate a set of aberrations for each CONICA configuration, that is a filter plus a camera objective.

3.2. Dichroic calibration

The estimation of the NAOS dichroic aberrations is obtained through the use of the AO system. 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 given defocus is introduced on the DM with the AO loop still closed to record the defocused image (See Paper II for a complete description of this procedure.). This approach gives the input data for the estimation of the NAOS dichroic aberrations together with the CONICA aberrations. The value of the separated dichroic aberrations is obtained by subtracting the value of the previously estimated CONICA aberrations.

The introduction of a defocus by the DM avoids the difficulties of object defocussing highlighted above for the CONICA calibration. Now, the same object is considered and thus the two images are located at the same position on the detector.

4. Simulation results

In order to validate the algorithm and to quantify its precision, we first consider simulated images. The conditions for this simulation are given by a point-like object, an imaging wavelength of $2.166\mu\text{m}$ and a pure defocus equal to λ (peak to valley) between the two images, corresponding to a defocus coefficient a_4^2 of 641.5nm RMS . We degrade the PSF by a wavefront deformation described by its Zernike coefficients. The coefficients are arbitrary but chosen to have comparable values as observed in the calibration procedure (see Paper II). The phase is generated with the first 15 Zernike polynomials, note that the estimated phase will be expanded on the same polynomials. We add white noise to each image in order to obtain a SNR of 200 which corresponds roughly to the SNR of the CONICA data (this is the typical average SNR corresponding to the time exposure used, which has been estimated on several measurements). The SNR is defined as the ratio of the maximum flux in the focus image over the RMS noise. The same noise statistics is applied to the defocused image. That results in a lower SNR on this image since the defocus spreads the PSF and reduces its maximum value. The focused and defocused images are presented in Figure 2. In this example only two possible limiting parameters (see Section 5) have been taken into account: noise and image re-centering. The system is assumed to be perfectly adjusted and the images to be perfectly pre-processed without having any residual background features.

Figure 3 shows a comparison between the input images of the simulation and the PSFs being reconstructed by the estimated aberrations and visualizes the quality of the aberration estimation. This visualization is helpful to judge real calibration data when only in-focus and out-of-focus images are available. Table 2 quantifies the performance of wavefront estimation by comparing the true and the estimated Zernike coefficients. On each coefficient a good accuracy is obtained: the errors are less than a

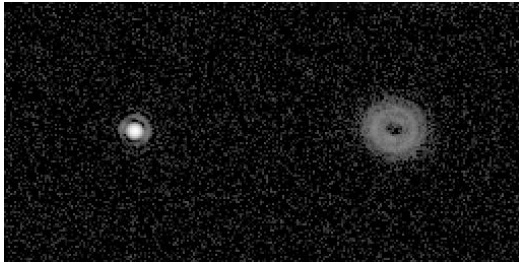


Fig. 2. Simulated focused and defocused images (logarithmic scale). Noise is added to obtain a SNR of 200 in the focused image.

Zernike number	true (nm)	estimated (nm)	error (nm)
4	60.5	61.5	1
5	-39.3	-46.7	7.4
6	58.1	61.7	3.6
7	-16.2	-15.9	0.3
8	-14.1	-10.9	3.1
9	-2.5	-3.6	1.1
10	13.7	12.9	0.7
11	-24.3	-26.	1.7
12	0.5	1.2	0.7
13	-3.2	-4.9	1.7
14	2.8	3.3	0.4
15	-2.4	-2.2	0.2

Table 2. Comparison between true and estimated Zernike coefficients applying phase diversity to simulated data with an image SNR equal to 200. The absolute value of the error is given for each coefficient. The total error is equal to 9.3 nm RMS.

few nanometers. The maximum error is for the wave aberrations. The influence of SNR on the simulation results is analyzed in Section 5.

The slight over-illumination of the two input images ($a_2=7$ in RMS and $a=-1.3$ nm RMS) are simulated with a high precision too. The error amounts to less than 1% (typical values are also shown in Table 2). In the next section, we focus on the possible sources for losses of the simulation accuracy.

5. Limitations

The PD algorithm is based on several assumptions which must be verified to obtain a good accuracy on the results. A list of possible sources is given below. The global procedure of data reduction can be found in Section 6. The sources can be decomposed in the part where the algorithm has a non-prior knowledge of the system (calibration, noise uncertainty), the loss due to the image acquisition and processing (noise, sidereal background, etc.) and the loss due to the limitation of the algorithm (spectral bandwidth of the simulation, etc.).

5.1. System limitations

Let us first focus on the loss due to the imperfect system knowledge.

5.1.1. Defocus distance

The main assumption of the PD principle is the addition of a known defocus (defocus in our case) between two images. In the case of defocus induced on the object, the coefficients of radial symmetric aberrations, with a main part, the image defocus is sufficient (see Figure 4). The difference between the maximal and the minimal simulated value of the defocus coefficient is 55 nm. For the spherical aberration, it is 1 nm and for the other ones, it is less than 5 nm.

In addition one can show that Figure 5 has a focus at a large distance on the known defocus, the propagation coefficient is equal to one. Hence, the uncertainty on the known defocus distance leads directly to uncertainty on the simulated defocus distance.

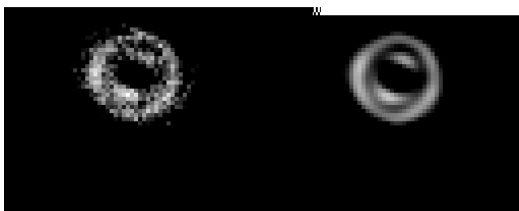


Fig. 3. Comparison between images (left) and reconstructed PSFs (right) from estimated aberrations. On top the focused images, at the bottom the defocused images (logarithmic scale and zoom $\times 2$ are considered for each image).

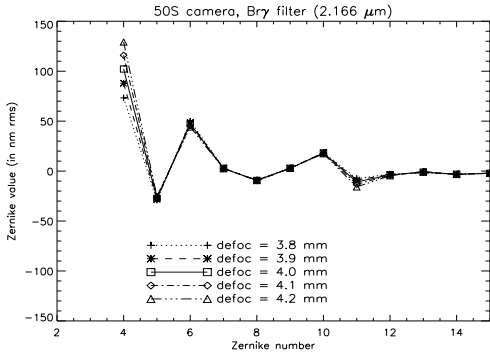


Fig. 4. Influence of the error on defocus distance on estimated aberration by phase diversity. The CONICA camera C50S and the narrow band filter Br_γ are used. The focused and defocused images are obtained using the 0-4 pinhole pair (that is a theoretical defocus of 4 mm between the two images). For the same couple of images several defocus distances (from 3.8 to 4.2 mm) serve as input parameters.

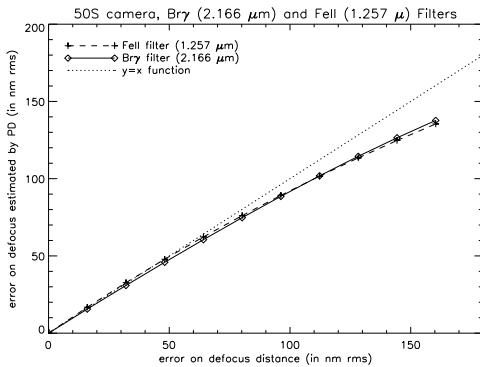


Fig. 5. Influence of the error on defocus distance on the estimation of the defocus. Experimental data have been used. The results are given for the CONICA camera C50S and two narrow band filters: FeII ($1.257 \mu\text{m}$) and Br_γ ($2.166 \mu\text{m}$). The pinhole pair 0-2 (that is a theoretical defocus of 2 mm) and 0-4 (that is a theoretical defocus of 4 mm) are respectively used for FeII and Br_γ filters. For each couple of images several defocus distances (respectively from 1.5 to 2.5 mm and from 3.5 to 4.5 mm for FeII and Br_γ filters) serve as input parameters.

This uncertainty on the defocus distance can be due to:

- an uncertainty on the physical position of the CONICA pinholes in the entrance focal plane. This uncertainty is estimated to ± 0.15 mm corresponding to $\Delta a_4 = \pm 24$ nm RMS,
- a systematic uncertainty on the F/D ratio (estimated to be less than two percent),

- the precision of a focus adjustment with the NAOS deformable mirror for the dichroic aberration estimation. This error is estimated to less than a few percent.

All these items lead to a precision of the estimated defocus roughly equal to a ± 30 nm. It will be shown in the following that the error on the defocus distance is by far the dominant error for our application.

5.1.2. Camera pixel scale

The camera pixel scale is needed to calculate the oversampling factor. An error on this factor induces an error on the coefficients of all radially symmetric aberrations (defocus, spherical aberration ...) as shown in Figure 6.

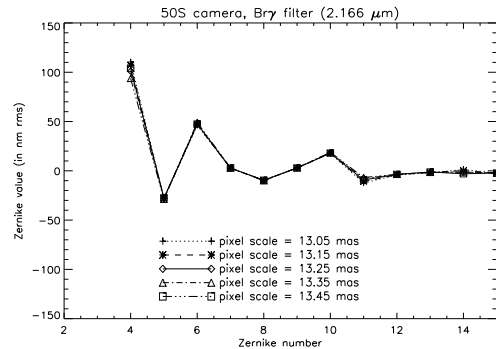
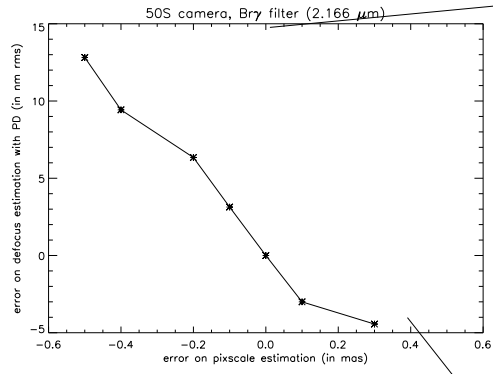


Fig. 6. Influence of the pixel scale error on estimated aberration by phase diversity. Experimental data have been used. The CONICA camera C50S and the narrow band filter Br_γ are used. The focused and defocused images are obtained using the 0-4 pinhole pair. For the same couple of images several pixel scales (from 13.05 to 13.45 mas) serve as input parameter for the PD algorithm.

An error of the pixel scale is essentially propagated to the defocus aberration estimation. The difference between the maximal and the minimal estimated value of the defocus coefficient is 17 nm. A slight error of 6 nm can be seen on spherical aberration but remains negligible in comparison to the one on the defocus. In Figure 7 the evolution of the estimation error of the defocus coefficient is plotted as a function of a pixel scale measurement error. It is assumed that the true value is 13.25 mas measured during the first on-sky tests of the AO system. Since the accuracy on the pixel scale measurement is better than 0.2 mas, one can estimate the wavefront error (WFE) due to this uncertainty to be less than a few nanometers and therefore to remain negligible.



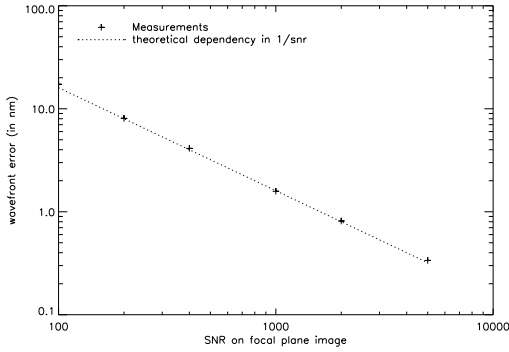


Fig. 9. Evolution of the wavefront error as function of SNR on focal plane image. Simulation presented in Section 4 have been used. A $1/\text{SNR}$ theoretical behavior is plotted for comparison.

Four pairs of focused-defocused images have been recorded sequentially in the same CONICA configuration (FeI filter and camera C50S). The SNR on these images is 400. Aberrations are then estimated for each couple and the WFE fluctuation on this four set of coefficients is equal to 2.2 nm RMS, which is in perfect agreement with Figure 9 (obtained on simulation).

5.2.2. Residual background features

The principle of the PD method is the minimization of a criterion (Equation 10) which is based on a convolution image model (Equation 1). Thus the image should be perfectly corrected for all instrumental features (background, dead pixels, etc.) in order to match the model. In practice, residual features are still present. In particular, in the case of CONICA images, a background fluctuation due to pick-up noise can induce residual features on the images even after a proper background calibration (see Subsection 6.2). These features are interpreted as signal by the phase-diversity algorithm. Therefore they induce bias on the aberration estimation. The effect of such fluctuations is highlighted in Figure 10 on experimental data. The difference of the PD results obtained with and without residual background features yield the WFE which is plotted as a function of the image size. This can be understood as a function of the residual background influence, too, because it obviously depends on the image size: the smaller the images, the less important the residual background in comparison to the signal. Nevertheless, the image size should be large enough to contain the whole signal. Furthermore, the modelisation of the pupil shape (see Subsection 5.3.3) must also be taken into account to choose the right image size. In Subsection 6.2 we describe

a pre-processing algorithm which allows to remove these residual background features.

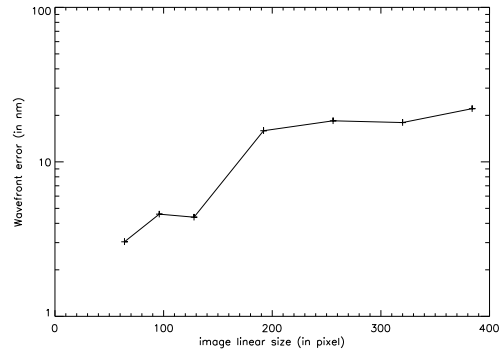


Fig. 10. Evolution of the wavefront error as a function of the image size. Experimental data have been used.

5.2.3. Number of estimated Zernike polynomials

As presented in Subsection 2.3, the phase regularization in our algorithm is provided by a truncation of the solution space through the use of a finite (and small) number of unknowns (typically the first twenty Zernike coefficients). Figure 11 shows, on experimental data, the influence of the number of estimated Zernike on the reconstruction quality. Note that, in the case of measurements of CONICA stand-alone aberrations, the pupil is unobscured and thus the Zernike polynomials are strictly orthogonal.

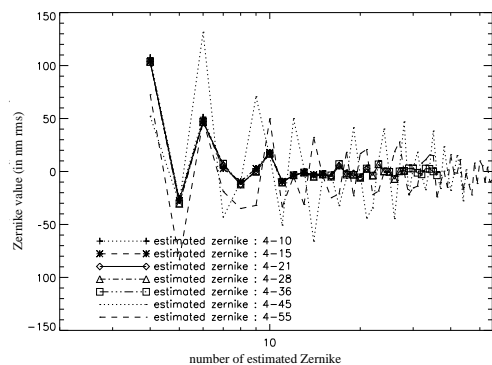


Fig. 11. Evolution of the aberration estimation as a function of the number of Zernike in PD algorithm. Experimental data have been used.

There exists a limit to the number of Zernike polynomials that can be estimated with a reasonable accuracy. Of course this limit depends on the signal to noise ratio on the images. In the present case, this number is equal to 36. Note that if a more sophisticated regularization term is introduced in the PD algorithm both on the object and the aberrations this limitation should be overcome. Nevertheless such a regularization is not needed here since the aberration amplitudes are negligible (less than a few nanometers) for Z_i above $i = 11$ (i.e the spherical aberration). The WFE between estimated Zernike coefficients 4-15 and estimated Zernike coefficients 4-36 is about 1.3nm RMS. This WFE is very small and thus shows that the aliasing of the Zernike polynomials above 15 on the estimated coefficients 4-15 is negligible.

5.3. Algorithm limitations

5.3.1. Spectral bandwidth

The phase diversity concept proposed here is a monochromatic wave-front sensor (theoretically the concept can be applied on polychromatic images but it induces an important modification of the algorithm to model the data (Seldin and Paxman, 2000)). Nevertheless it has been shown (Meynadier et al., 1999) that the use of broad-band filters does not significantly degrade the accuracy as long as $\frac{\Delta\lambda}{\lambda}$ is lower than a few tens of percents (typically $\frac{\Delta\lambda}{\lambda} \leq 0.15$).

5.3.2. Image centering

As mentioned above, the PD algorithm can not estimate a tip-tilt between the two images larger than 2π . Therefore, a fine centering between focused and defocused images must be done before the aberration estimation (see Subsection 6.2).

5.3.3. Pupil model

Since we consider here experimental data (see Section 6), the modelisation of the pupil shape in the algorithm is critical, in particular the pixelisation effects. Indeed, in PD algorithm the pupil definition depends on the image size and on the oversampling factor. For example, images of camera C50S in K band oversample with a factor of 2 and a 32×32 image will lead to a pupil diameter of 8 pixels (see Figure 12). In this case, the pixelisation effects on the shape of the pupil will induce aberration estimation error. These effects are illustrated in Figure 13.

Therefore, large images are recommended to well model the pupil and to obtain accurate results. Nevertheless, two problems may occur with the processing of large images:

- a residual background problem (see Subsection 5.2.2). The larger the image, the more important the effects



Fig. 12. Pupil shape in the PD algorithm for three numbers of pixel used in the pupil sampling [8, 32, 64 and 256]. The oversampling factor of 2 in K band leads to corresponding image sizes equal to [32, 128, 256 and 1024].

of residual background unless the images are tapered outside their central region.

- a computation time problem. Because of the iterative resolution of criterion defined in Equation 7, the increase of images size will lead to a increase of computation time.

This highlights the choice of a good compromise between pupil model in the PD algorithm to avoid phase reconstruction error and a reasonable image size.

The evolution of the reconstruction error as a function of pupil sampling is presented in Figure 13. To minimize the residual background effects, all the background pixels (that is pixels with no PSF signal) has been put to zero in the images.

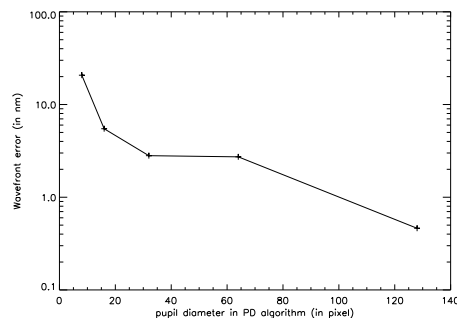


Fig. 13. Evolution of the wavefront reconstruction error as function of pupil sampling in PD algorithm. The x axis gives the pupil diameter in pixel in the PD algorithm.

Considering the results shown in Figures 13 and 10 along with the computation load lead us to choose an image size equal to 128×128 pixels for the K band and 64×64 pixel for the J band.

5.4. Conclusion

In this part, we have analyzed and quantified, on experimental and simulated data, the possible sources of errors in the static aberration estimation for NAOS-CONICA. It is shown that the main source of errors is due to an imperfect knowledge of the system (that is calibration errors).

In particular a precise knowledge of the defocus distance between focused and defocused planes is essential.

If very high precisions are required on the estimation and on the correction of static aberrations (for instance in the case of a future very high SR AO for exo-planet detection), the PD must be taken into account in the early design of the system in order to optimize with respect to the constraints and error sources listed above.

6. Practical example

In this section we give a detailed description on a practical example of the global procedure used to estimate NAOS-CONICA static aberrations. This procedure is quite general but the particularities of the CONICA and NAOS dichroic aberration measurements are underlined. All the illustrations are obtained for the following configuration of CONICA: objective C50S (pixel size equal to 13.25 mas) and Br₇ filter.

6.1. Input data

The input data are a focused and a defocused image ($i_{1,2}$) with their associated backgrounds ($b_{1,2}$) and a known defocus distance expressed in mm in the entrance focal plane of CONICA. This distance is given by the pinhole choice in the case of CONICA measurements (see Table 1) or by the defocus introduced by the DM in the case of the NAOS dichroic aberration measurements. In the example we consider the first approach and introduce the defocus by the pinhole choice.

6.2. Pre-processing

The pre-processing of the images is required before the wavefront can be estimated. We split the pre-processing in several steps:

- Conventional background subtraction in order to remove the main part of detector defects (bad pixels, background level, possible background features, etc.):

$$i_{1,2}^{corr} = i_{1,2} - b_{1,2} \quad (14)$$

The division by a flat-field pattern is recommended to increase the accuracy of the results even if it is not done in this example.

- The focused and defocused images are re-centered by a correlation procedure. For each filter we computed the relative shifts of the PSFs against each other. The median of these shifts was determined and serves to re-center the defocused images. This procedure ensures that re-centering is accurate enough to obtain a relative tip-tilt between the two images lower than 2π .
- Removing of the residual background feature: the most important feature is a residual sine function in vertical direction due to pick-up noise. Its amplitude is

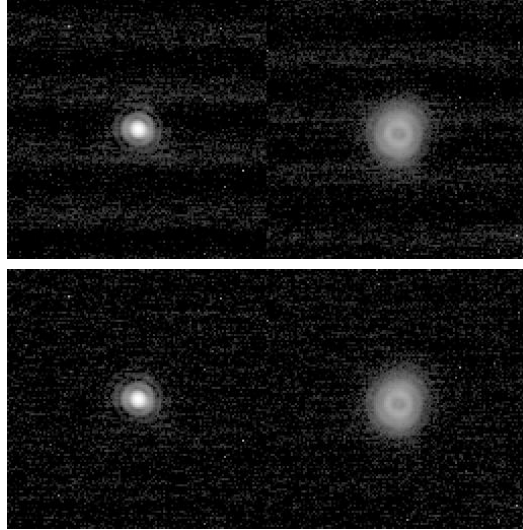


Fig. 14. Comparison of focused and defocused images before (on top) and after (at the bottom) the application of an algorithm which removes residual background features (log scale). The estimated SNR is equal to 400.

greater than the noise level. An estimation of the residual background features is necessary and performed directly on the images using a median filter in horizontal direction applied to each image column. A comparison of the images before and after this residual feature removal is presented in Figure 14.

- Image windowing: the image size is a trade-off between a good numerical pupil modeling and a reasonable computation time. We achieve reliable results by using 128×128 frames.

Note that we have assumed that all the bad pixels have been removed by the background subtraction. If some of them are still present in the pre-processed images, they must be removed by hand to ensure that they do not induce reconstruction errors in the PD algorithm.

6.3. CONICA aberration estimation

When both focused and defocused images have been pre-processed as described above, the PD algorithm can be applied. The inputs of the PD algorithm are:

- pixel scale: 13.25 mas (camera C50S),
- central obstruction given by the fraction of the pupil diameter: 0 (full pupil),
- wavelength (in μm): 2.166 (Br₇ filter),
- highest estimated Zernike number: 15 (estimation of Zernike polynomials from 4 up to 15),

- defocus coefficient a_4^d (in nm RMS): in the present example, $d = 4$ mm which leads to $a_4^d = -641.5$ nm,
- the focused and defocused images obtained after pre-processing.

Zernike	4	5	6	7	8	9
aberration (nm) raw	91	-27	48	3	-9	-4
aberration (nm) corrected	112	-24	47	1	-5	-1

Zernike	10	11	12	13	14	15
aberration (nm) raw	19	-20	-5	-1	1	-2
aberration (nm) corrected	17	-19	-3	-2	-3	-3

Table 3. Measured aberrations (in nm RMS) for the CONICA camera C50S and the Br₇ filter. The defocus distance between the two images is 4 mm and the estimated SNR is 400. Only the 12 first Zernike 4-15 are given. The raw values are obtained without residual background subtraction. The corrected ones are obtained after subtraction of the residual background features.

The results obtained are summarized in Table 3. A bad background correction leads to an important error on the defocus ($\simeq 20$ nm). A comparison between focused and defocused images and reconstructed PSFs from the estimated Zernike is proposed in Figure 15. The estimated SR on the 12 estimated Zernike is equal to 87%. It compares nicely to the SR directly computed on the focal plane image, which is equal to 85%.

7. Conclusion

We have given a precise description of the phase diversity algorithm developed at ONERA and of its use for the calibration of NAOS-CONICA static aberrations. The concept of phase diversity has first been recalled, then the expression of the criterion to be minimized has been given. Guidelines for a practical implementation have been proposed. The essential parameters of the algorithm have been studied. This has allowed us to highlight the essential verifications and calibrations needed to obtain accurate results with PD.

First, a simulated example of utilization has been proposed in order to demonstrate the precision of the estimation under typical SNR conditions of acquisition of NAOS-CONICA images. It shows that for a SNR of 200, the error on the aberration estimation is less than few nanometers per polynomial. This simulation gives the ultimate performance for a given SNR since it does not account for experimental uncertainties and bias.

The limitations of the experimental approach have been listed and studied. It has been shown that the main

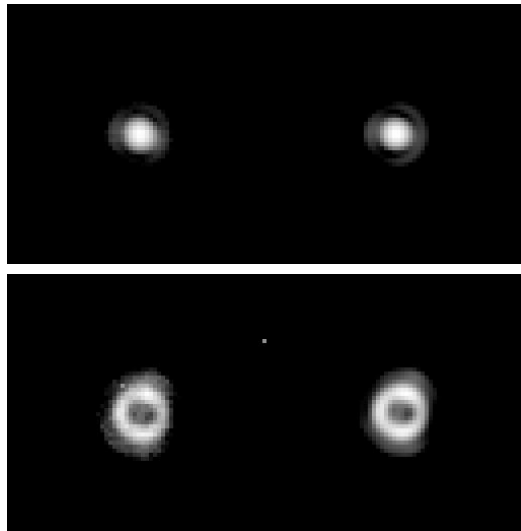


Fig. 15. Comparison between images (left) and reconstructed PSF from estimated aberrations (right). [up] focused image, [down] defocused image (log scale are considered for each image).

source of error is induced by the uncertainty on the pinhole distances which leads to an error on the defocus estimated to ± 24 nm. An other source of degradation is due to the bad knowledge of the pixel scale which contributes for a few nanometers on the defocus. Furthermore, an exact knowledge of the pupil alignment is required to ensure a good estimation of the aberrations. The residual background after compensation in the images is another important source of degradation, estimated to a few tens nanometers. In the case of the images of NAOS-CONICA, the SNR is high enough to make this error source negligible in comparison to the others.

All these error sources lead to an uncertainty on the aberration estimation: ± 35 nanometers on the defocus and around few nanometers for the high order Zernike polynomials.

Finally, an example of processing of experimental data obtained on CONICA during commissioning at the VLT has been given. Pre-processing required to minimized the Camera defects has been described and a comparison of focal plane PSFs with reconstructed PSFs using estimated aberrations has been made. It shows good agreements both in term of PSF shape and SR measurements.

It has been shown that Phase Diversity is a simple and powerful approach to compute unseen aberration of an AO system. This is of a great interest for future very high SR system in which an accurate estimation and correction of such aberrations is essential to achieve the ultimate

performance and to reach the scientific goals (exo-planet detection for instance)

Acknowledgements. We wish to thank V. Michau and M.-T. Velluet for fruitful discussions and helpful comments on this work. And we are thankful to F. Lacombe for his help during the acquisition and the interpretation of NAOS-CONICA data. This research was partially supported through the Marie Curie Fellowship Association of the European Community.

References

- Blanc, A., Idier, J., and Mugnier, L. M., 2000, in J. B. Breckinridge and P. Jakobsen (eds.), *UV, Optical, and IR Space Telescopes and Instruments*, Vol. 4013, pp 728–736, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE, Bellingham, Washington
- Carreras, R. A., Restaino, S., and Duneman, D., 1994, in *Image Reconstruction and restoration*, Vol. 2302, pp 323–328, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE
- Demoment, G., 1989, *IEEE Trans. Acoust. Speech Signal Process.* 37(12), 2024
- Ekstrom, M. P. and Rhoads, R. L., 1974, *J. Comput. Phys.* 14, 319
- Gates, E. L., Restaino, S. R., Carreras, R. A., Dymale, R. C., and Loos, G. C., 1994, in *Image Reconstruction and Restoration*, Vol. 2302, pp 330–339, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE
- Gonsalves, R. A., 1982, *Optical Engineering* 21(5), 829
- Hartung, M., Bizenberger, P., Boehm, A., Laun, W., Lenzen, R., and Wagner, K., 2000, in M. Iye and A. F. Moorwood (eds.), *Optical and IR Telescope Instrumentation and Detectors*, Vol. 4008, pp 830–841, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE
- Hartung, M., Blanc, A., Fusco, T., Lacombe, F., Mugnier, L. M., Rousset, G., and Lenzen, R., In this issue, *Astron. Astrophys.*
- Kendrick, R. L., Acton, D. S., and Duncan, A. L., 1994, *Appl. Opt.* 33(27), 6533
- Lee, D., Welsh, B., and Roggemann, M., 1997, *Opt. Lett.* 22(13), 952
- Lee, D. J., Roggemann, M. C., and Welsh, B. M., 1999, *J. Opt. Soc. Am. A* 16(5), 1005
- Lenzen, R., Hofmann, R., Bizenberger, P., and Tusche, A., 1998, in *Infrared Astronomical Instrumentation*, Vol. 3354, pp 606–614, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE
- Löfdahl, M. G. and Scharmer, G. B., 1994, *Astron. Astrophys.* 107, 243
- Löfdahl, M. G., Scharmer, G. B., and Wei, W., 2000, *Appl. Opt.* 39(1), 94
- Little, R. J. A. and Rubin, D. B., 1983, *The American Statistician* 37(3), 218
- Meynadier, L., Michau, V., Velluet, M.-T., Conan, J.-M., Mugnier, L. M., and Rousset, G., 1999, *Appl. Opt.* 38(23), 4967
- Noll, R. J., 1976, *J. Opt. Soc. Am.* 66(3), 207
- Paxman, R. G., Schulz, T. J., and Fienup, J. R., 1992, *Journal of the Optical Society of America A* 9(7), 1072
- Restaino, S., 1992, *Appl. Opt.* 31(35), 7442
- Rousset, G., Lacombe, F., Puget, P., Gendron, E., Arsenault, R., Kern, P., Rabaud, D., Madec, P.-Y., Hubin, N., Zins, G., Stadler, E., Charton, J., Gigan, P., and Feautrier, P., 2000, in P. L. Wizinowich (ed.), *Adaptive Optical Systems Technology*, Vol. 4007, pp 72–81, Proc. Soc. Photo-Opt. Instrum. Eng., SPIE, Bellingham, Washington
- Rousset, G., Lacombe, F., Puget, P., Hubin, N., Gendron, E., Conan, J.-M., Kern, P., Madec, P.-Y., Rabaud, D., Mouillet, D., Lagrange, A.-M., and Rigaut, F., 1998, in D. Bonaccini and R. K. Tyson (eds.), *Astronomical Telescopes & Instrumentation*, Vol. 3353, Proc. Soc. Photo-Opt. Instrum. Eng., Kona, Hawaii
- Seldin, J. and Paxman, R., 2000, in *Imaging Technology and Telescopes*, Vol. 4091-07, SPIE
- Seldin, J. H. and Paxman, R. G., 1994, in *Image Reconstruction and Restoration*, Vol. 2302, SPIE
- Thelen, B. J., Paxman, R. G., Carrara, D. A., and Seldin, J. H., 1999, *J. Opt. Soc. Am. A* 16(5), 1016

Calibration of NAOS and CONICA static aberrations. Experimental results.

Cet article a été accepté pour publication dans le journal *Astronomy&Astrophysics*.

Calibration of NAOS and CONICA static aberrations

Experimental results

M. Hartung¹, A. Blanc^{2,3}, T. Fusco², F. Lacombe⁴, L. M. Mugnier², G. Rousset², and R. Lenzen¹

¹ Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
e-mail: hartung@mpia.de

² Office National d'Études et de Recherches Aérospatiales, Département d'Optique Théorique et Appliquée,
BP 72, F-92322 Châtillon Cedex, France

³ Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Électricité, Plateau de Moulon, F-91192 Gif-sur-Yvette,
France

⁴ Observatoire Paris-Meudon, LESIA, Place Jules Janssen, F-92195 Meudon Cedex, France

Abstract. NAOS is the first adaptive optics (AO) system of the Very Large Telescope and will provide CONICA with diffraction-limited images. CONICA is a near infrared camera that offers a variety of imaging and spectroscopic observing modes. A technique will be described to benefit of the AO system to correct not only for atmospheric turbulence but also for the internal optical aberrations of the high-resolution camera and the beam splitters of NAOS. The aberrant optical components in the light path of CONICA as well as the beam splitters are outside of the AO loop and therefore no self-acting correction is possible. Independently of the AO wavefront sensor, a separate measurement of these aberrations using a method called phase diversity allows one to predict for a certain instrument configuration the corresponding aberrations. They are quantified by sets of Zernike coefficients that are rendered to the adaptive optics. This technique turns out to be very flexible and results in a further improvement of the optical overall performance. The application of phase diversity to the instrument is investigated in a preceding paper (Blanc et al. 2002). In the present paper we present in detail the instrumental implementation of phase diversity, the obtained calibration results, and the achieved gain in optical performance.

Key words. adaptive optics – static aberrations – phase diversity – wavefront sensing – NAOS-CONICA – VLT

1. Introduction

NAOS-CONICA is the first adaptive optics (AO) system of the Very Large Telescope (VLT) and saw its first light in November 2001 (Brandner et al. 2002). The Nasmyth Adaptive Optics System NAOS (Rousset et al. 1998, Rousset et al. 2000) delivers diffraction-limited images to the Coudé¹ Near Infrared CAmera CONICA (Lenzen et al. 1998, Hartung et al. 2000). To retrieve the maximum possible performance of the system in terms of Strehl ratio² (SR) a method has been developed to calibrate the remaining degradation of the image quality induced by its optical components. Defaults of the wavefront attributed to any degradation within the AO loop (common path) are seen directly by the AO wavefront sensor (WFS) and thus the AO system can correct for these aberrations au-

¹ The notation Coudé has historical origins as CONICA originally was planned for the Coudé focus.

² The SR ratio is a common way to describe the quality of the point spread function. It is given by the ratio of the measured and the theoretical diffraction-limited peak intensity.

tomatically. This is not the case for a degradation of image quality induced by components outside the AO loop. An experimental setup has been applied which allows one to sense the wavefront of the light which has passed the whole system *without* making use of the AO wavefront sensor. Therefore we draw on a well-known method called phase diversity (Gonsalves 1982, Paxman et al. 1992). It turns out that a number of theoretical and experimental constraints have to be examined before reliable results can be obtained in sensing the wavefront via phase diversity (PD). We focused on this in a precedent paper (Blanc et al. 2002), hereafter Paper I. In this second paper we first give a brief description of the instrument (Section 2). Then we focus on the experimental setup which enables us to calibrate the variety of beam splitters, filters and camera objectives. The design constraints for the implementation of PD are illustrated, and the resulting setup as well as the procedure to obtain the appropriate input data for PD are described (Section 3).

Because of the huge number of instrument modes it is not feasible to perform the PD calibration for each possible configuration. We explain how the wavefront degradations of the different optical components are disentangled. Then, the individual parts of the optical train can be calibrated separately and it is no longer required to do this for every possible combination. In detail, we will allocate the wavefront error to the dichroic mirrors of NAOS (beam splitter between wavefront sensor and imaging path), to the CONICA filters and camera objectives (Section 4).

Thereafter, the sensed wavefront errors are used to calculate the corresponding SRs. These are compared to the SRs directly determined from the images and the consistency is verified (Section 5). Finally, after presentation of the complete calibration procedure and its results, the measured wavefront errors are rendered in terms of Zernike coefficients to the AO system to demonstrate the gain in overall performance after closed loop correction (Section 6).

2. The instrument

This section is dedicated to a brief description of the instrument NAOS-CONICA. In particular, we emphasize the sub-devices as much as the aspects of design which are important with respect to the static wavefront error estimation by phase diversity. Fig. 1 gives an overview of the VLT instrument. NAOS is installed at one of the Nasmyth foci. It picks up an $f/15$ beam, corrects for atmospheric turbulence and hands on again an $f/15$ beam providing CONICA with diffraction-limited images. Having passed the Nasmyth focal plane, the beam is led to a first collimating parabola. Then it is reflected successively onto the tip-tilt (TTM) and the deformable mirror (DM). The following dichroic mirror separates the optical train into the imaging path and the wavefront sensing path. NAOS offers five different dichroic beam splitters to adapt for the flux and the spectral characteristics of the guide star. In the imaging path the light is refocused onto the entrance focal plane of CONICA, which is located behind the entrance window in the cold cryostat. Between NAOS and CONICA an atmospheric dispersion compensator (ADC) can be slid in in case of a high zenith angle.

The wavefront sensing path consists of a field selector (Spanoudakis et al. 2000) and two wavefront sensors. They are located between the dichroic mirror and the WFS input focus. For the sake of clarity these components are not shown in Fig. 1. The two wavefront sensors, one in the visible and one in the near infrared spectral range, enhance the sky coverage of possible guide stars. The field selector chooses the guide star in a 2 arcmin field of view and allows differential object tracking, pre-calibrated flexure compensation and counter-chopping. In combination with the deformable mirror it is also able to correct for a certain amount of defocus, as needed when the prisms of the atmospheric dispersion compensator are shifted into the beam. Note that this ability of focus correction offers

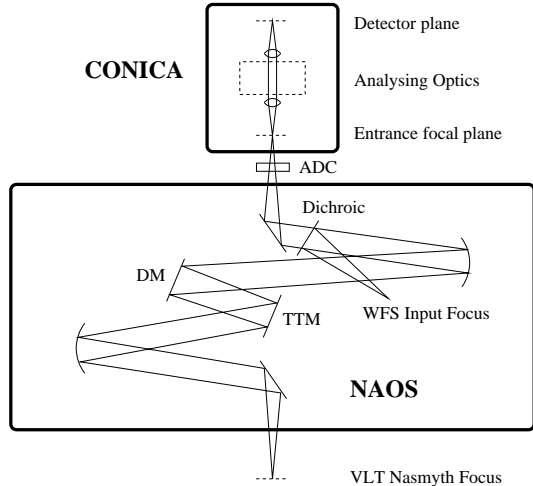


Fig. 1. Outline of the VLT instrument NAOS-CONICA

Table 1. The required defocus distances for a phase diversity of 2π rad (peak to valley) ($\lambda = 1\mu\text{m}$) are listed corresponding to the f-ratios (pixel scales).

Camera	f/D	Pixel scale	d (mm)
C50S	51	13.3 mas/pixel	20.8
C25S/C25L	25.5	27.1 mas/pixel	5.2
C12S/C12L	12.8	54.6 mas/pixel	1.3
C06S/C06L	6.38	109 mas/pixel	0.33
Entrance focal plane	15	1.72 mas/ μm	1.8

a possibility to perform PD measurements that we will refer to later on (Section 3.2).

The high angular resolution camera CONICA is equipped with an Aladdin array (1Kx1K) covering the 1-5 μm spectral region. Splitting the wavelength region into two parts (1 to 2.5 μm and 2.0 to 5 μm) allows us to keep the light path achromatic. Therefore the four different pixel scales are realized by seven cameras (Table 1). To each pixel scale a camera is associated with the short wavelengths region (S-camera) and another one with the long wavelengths region (L-camera). The only exception is the camera with the highest magnification (C50S). There is no long wavelength counterpart needed.³

A variety of different observing modes is provided by the analyzing optics: chronography, low resolution long slit spectroscopy, imaging spectroscopy by a tunable cold Fabry-Perot, polarimetry by wire-grids or Wollaston prisms, and about 40 broad- and narrow-band filters can be chosen.

³ The Nyquist criterion for L and M-band is already fulfilled for a camera with lower magnification (C25L).

3. Phase diversity setup

Our input for phase diversity wavefront estimation are two images: one of them in focus and the other one out of focus. In this manner we introduce the well-known phase diversity which is an obligatory input parameter for PD. One should recall that best phase diversity estimates are to be expected applying a peak to valley phase diversity $\Delta\phi$ between 1π and 3π and the input images must be at least Nyquist sampled⁴ (Paper I). The corresponding defocus distance d depending on the applied wavelength is obtained by

$$d = \frac{4\lambda}{\pi} (f/D)^2 \Delta\phi. \quad (1)$$

In the next two subsections we describe in detail two ways of introducing this phase diversity. Both ways are essential to enable us to separate the wavefront error and to assign it to different contributors. This disentanglement is described in Section 4.2.

3.1. CONICA stand-alone: focus shift by object

First, we regard the possibilities to obtain the necessary input images with CONICA stand-alone. The CONICA detector is mounted on a tunable stage which is software controlled and can be driven in the cold environment. This allows us in principal to obtain a defocused image but the focus drive spans only a region of 2 mm. Using Equation 1 we compile for all available camera objectives the necessary defocus distances in the detector plane corresponding to a diversity of 2π at a wavelength of $1\mu\text{m}$ in Table 1. Only for the low magnification cameras (C06, C12) is the defocus distance sufficient. But these very cameras under-sample in K and at shorter wavelengths so that the focus stage mechanism finally fails in every case. For that reason we swerve to the entrance focal plane. Here a phase diversity of 2π corresponds to 1.8 mm at a wavelength of $1\mu\text{m}$ or 3.6 mm at $2\mu\text{m}$, which is small enough to be implemented in the entrance focal plane.

In this plane a wheel is located carrying different field limiting masks, coronagraphic masks and the slits for spectroscopy. On the wheel we implement four different pinholes at different axial positions. The pinhole diameter is $10\mu\text{m}$. One pinhole is placed exactly in the entrance focal plane and yields a focused image on the detector (0 mm), and three other pinholes are located 1 mm, 2 mm and 4 mm out of the entrance focal plane. The four pinholes are mounted onto a plate fitting in a socket of the mask wheel. This device will be referred to as a Zernike tool later on. Impacts on the PD estimation due to the

⁴ Indeed, it is possible to model the wavefront error even at undersampled images (Jefferies et al. 2002). In our case there is no cost to be constrained by the Nyquist criterion because we can rather select the appropriate filter or else we deal with a camera used in the L- and M-band where static wavefront aberrations are negligible.

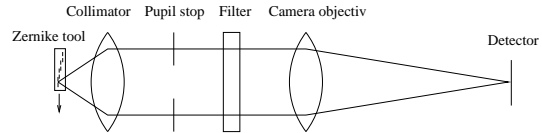


Fig. 2. The Zernike tool with its pinholes in the light path of CONICA.

mechanical precision of the pinhole positions and possible deviations of their shape are investigated in Paper I.

Fig. 2 depicts the setup for the CONICA internal phase diversity measurements. The leftmost component carries the four pinholes, which are shifted against each other, with the values given above. In rotating the wheel holding the Zernike tool we are able to select a pinhole in the field of view. After a collimating lens and a pupil stop, a filter selects the wavelength range and finally the camera objective forms the object image on the detector. The chosen camera objective determines the f-ratio and the pixel scale.

To center the image of the pinholes on the detector, the whole pinhole mount is shifted by turning the mask wheel. In principal PD needs the input images to be on the same spot to ensure that the same aberrations are sensed. The horizontal position of the pinholes can be controlled by adjusting the rotation angle of the wheel. In vertical direction there is no degree of freedom, but the four pinholes are mounted circularly to compensate for the circular movement. By this means a vertical precision of 50 mas (C50S) can be reached⁵. This is easily sufficient not to see any influence due to field aberration effects. PD measurements taken at different detector positions and calculations performed with an optical design software showed that even at the corner of the field of view (13 arcsec) the field aberration is negligible (Paper I). Note that for some measurements in Paper I an earlier version of the Zernike tool was used with a design not optimized for the circular movement of the pinholes. The worst separation that could occur with the former Zernike tool was about 1.3 arcsec. But even with this tool no relevant impact on the precision of wavefront sensing was detected.

Apart from the fact that the Zernike tool with its pinholes at the entrance focal plane provides the required focus shifts, it is convenient that the required focus shifts do not depend on the camera objective (pixel scale) anymore. But note: defocusing by moving an object in the entrance focal plane does not correspond exactly to a defocus due to a shifted detector plane. An investigation of this effect is done in Paper I and turns out to be negligible.

To summarise this section: the PD input data to derive the total CONICA internal aberrations are obtained by object defocusing in the CONICA entrance focal plane. The object defocusing is realized by four $10\mu\text{m}$ pinholes at different axial positions. Note that since the entrance focal plane of CONICA is located inside the cold cryostat,

⁵ 1 pixel corresponds to 13.3 mas (C50S)

aberrations accrued from the CONICA entrance window are not included in this wavefront estimation.

3.2. NAOS-CONICA: focus shift by the deformable mirror

Now, we describe how the PD input images are obtained which are used to sense the wavefront aberrations of the whole instrument, i.e., the adaptive optics NAOS together with its infrared camera CONICA. In this case we can take advantage of the AO system's capabilities to itself introduce an adequate focus shift and thus there is no need for the implementation of a special tool or a modification of the design.

In the entrance focal plane of NAOS, which coincides with the VLT Nasmyth focal plane, a calibration point source can be slid in and imaged by CONICA. This point source is realized by the output of a fiber with a diameter of $10\ \mu\text{m}$ fixed on a movable stage. On the same stage a second source much larger in diameter ($400\ \mu\text{m}$) is mounted. It is only seen by the WFS and serves as a reference source to close the loop. This extended source is needed for technical reasons. In the case of no atmospheric turbulence the more extended source provides a much better feed-back signal to the WFS than the small one. By this means the AO control loop is adjusted for any aberrations emerging in the common path. To obtain the focus shift affecting the entire instrument, we introduce the desired amount of defocusing in the WFS path by moving the mirrors of the field selector. During this process the loop is kept closed. Instantly, the arising focus shift is detected by the WFS. Correspondingly, the real time computer commands the DM to compensate for the detected defocus. Finally, the spots on the Shack-Hartmann WFS are centered again, but the defocus of the DM takes effect in the imaging path. For a pure defocus the DM will take a parabolic shape. The maximum achievable defocus by this method is limited by the DM's stroke and turns out to be about 20mm. Referring to Table 1, this is enough to introduce the needed diversity for an $f/15$ beam.

The procedure is shown in Fig. 3 and provides us with the PD input data to estimate the NAOS-CONICA overall wavefront errors. In comparison with the procedure described in Section 3.1 we deal with the same object now, and we must not care about any deviations in the position of the image pairs. This simplifies data acquisition for the measurement and diminishes the number of possible error sources.

4. Calibration of NAOS and CONICA static aberrations

4.1. Introduction

An appropriate way to describe the shape of a wavefront in the telescope pupil is by using Zernike polynomials (Noll 1976). A set of Zernike coefficients indicates the linear combination describing the present wavefront. As a matter

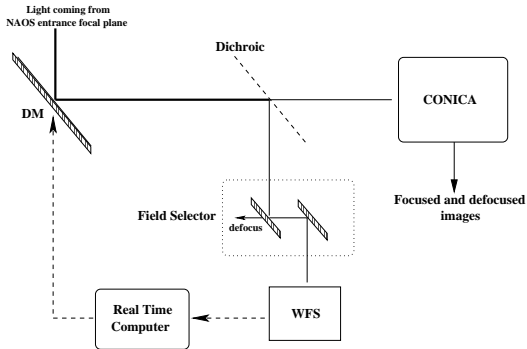


Fig. 3. Defocusing in closed loop using the NAOS Field Selector

of course, we can regard this set of coefficients as a vector. We refer to the Noll notation (Noll 1976) which labels the focus with 4, the tangential and sagittal astigmatism with 5 and 6, coma with 7 and 8, and so on. Since the coefficients for piston (1), and tip-tilt (2, 3) are extraneous to the image quality they are dropped.

An extensive examination of the variety of error sources due to the practical and instrumental constraints was done in the preceding Paper I. The induced aberrations due to defocusing by a shifted object in the CONICA stand-alone case have been simulated and proven to be negligible. The influence of the pupil shape and its numerization have been evaluated, errors taken in account with regard to the camera pixel scale and the defocus distance deviations have been simulated and the problem of different object structure was considered. Furthermore we focused in detail on the handling of data reduction, e.g. the influence of the different noise sources such as read-out noise or pickup noise. In Paper I we state that all these error sources accumulate to $\pm 35\ \text{nm}$ RMS for the focus coefficient (4). Since the presented calibration data of this paper are acquired with an optimized Zernike tool, the expected error should be well below this number. The accuracy of the higher order coefficients has not changed and amounts to about $\pm 5\ \text{nm}$ RMS.

In this section we describe how the overall wavefront error can be decomposed and assigned to its corresponding optical components. Then we present the experimental results for one camera objective and some selected filters of CONICA as well as the results for the dichroics of NAOS.

4.2. Disentanglement

In the preceding sections we described how the static non-common-path wavefront error can be measured, whether for CONICA stand-alone or for the entire instrument NAOS-CONICA. However, each determination of the wavefront error is only valid for the particular instrument configuration in which it was measured. The tremendous

number of instrument configurations⁶ makes it impractical to perform these calibrations for any possible instrument setup. For a practical application we need to split up the measured wavefront aberrations and assign the corresponding contributions to the divers optical components. This allows the construction of a configuration table⁷ with entries for each optical component of the instrument. When a special instrument configuration is selected, the corresponding wavefront error contributions can be read out, added together and delivered to the AO system. This enables the DM to pre-correct for the current static wavefront aberrations.

In principal, we have to differentiate between three categories of optical components in the imaging path: the NAOS dichroics, the CONICA filters and the camera objectives.

The contribution of the NAOS dichroics a_i^{dichro} can be determined by subtracting the overall NAOS-CONICA instrument aberrations a_i^{NCtot} from the total CONICA instrument aberrations a_i^{Ctot} :

$$a_i^{\text{dichro}} = a_i^{\text{NCtot}} - a_i^{\text{Ctot}} \quad (2)$$

The vector components are labeled by the Zernike number i running in our case from 4 to 15. Regarding the PD estimations of the different CONICA filters for one camera objective (see Fig. 4 and Section 4.3) we ascertain that generally the filter aberrations a_i^{fil} are small and mainly the achromatic camera objective aberrations are seen. In any case the filter aberrations are not correlated with each other nor with the camera ones. This suggests that we can deduce the camera objective contribution by applying the median to the total CONICA internal aberrations $a_i^{\text{Ctot,fil}i}$. We prefer the median instead of the mean to avoid taking into account highly aberrant filters. The filter i which was used to determine the corresponding total CONICA internal aberration a_i^{Ctot} is indicated by $\text{fil}i$. The camera objective and the residual filter contributions are obtained by these relations:

$$a_i^{\text{cam}} = \text{median}(a_i^{\text{Ctot,fil}1}, a_i^{\text{Ctot,fil}2}, \dots, a_i^{\text{Ctot,fil}n}) \quad (3)$$

$$a_i^{\text{fil}} = a_i^{\text{Ctot}} - a_i^{\text{cam}} \quad (4)$$

The separation of the wavefront aberration into the contributions associated with the three categories of optical components (NAOS dichroics, CONICA filters and camera objectives) is only possible when we make use of *both* ways to introduce a focus shift, i.e. the DM to determine the NAOS-CONICA overall aberrations and the Zernike tool to determine the total CONICA aberrations. In addition, we note that even if we refer to these three categories by the notation dichroics, filters and objectives,

⁶ Given the combinations of 5 NAOS dichroics, 40 narrow and broad band filters and 7 camera objectives!

⁷ In fact, a configuration file is generated. The instrument control software takes care of what coefficients have to be applied for the selected instrument setup. These processes are hidden and completely automatic.

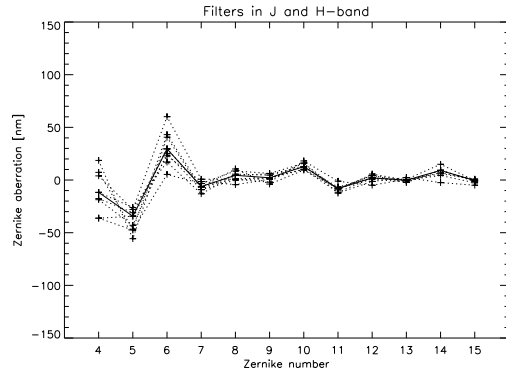


Fig. 4. CONICA internal aberrations measured by 8 NB filters in J- and H-band with camera objective C50S and pinhole pair 0/2 mm. The thick line indicates the median representing the camera aberrations.

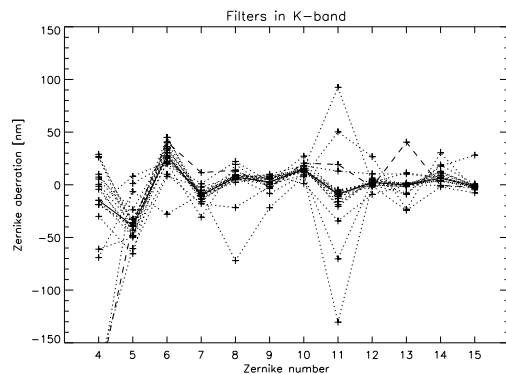


Fig. 5. CONICA internal aberrations measured by 19 NB filters in K-band with camera objective C50S and pinhole pair 0/4 mm. The thick line indicates the median representing the camera aberrations. The dashed line highlights the aberrant filter NB2.09 which is picked out for the demonstration images in Fig. 6.

the other components in the optical path are included, as well, even when they are not mentioned explicitly. E.g., the aberrations of the CONICA entrance window are included in the dichro aberrations and the aberrations of the CONICA collimator are an inextricable part of the camera aberrations.

4.3. Calibration of CONICA: camera and filters

Fig. 4 shows the aberrations for all eight narrow band filters in J- and H-band of CONICA. The camera objective C50S and the pinhole pair (0/2 mm) is used to obtain the calibration data. The fourth coefficient $a_4^{\text{Ctot,fil}i}$ expressing the defocus shows a peak-to-peak variation of

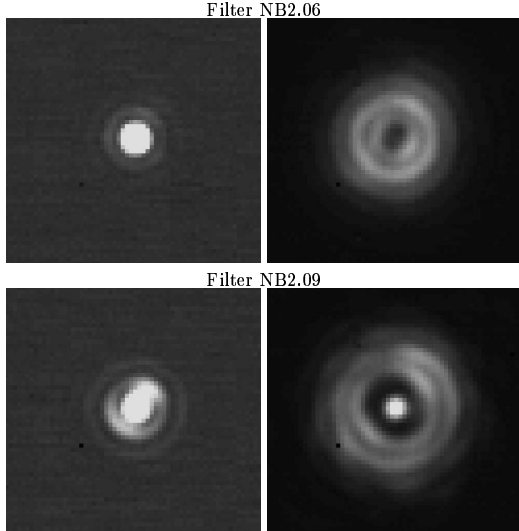


Fig. 6. Comparison of PD input images of a filter with small aberrations (NB2.06, on the top) and a filter with high aberrations (NB2.09 at the bottom). The in-focus images are placed on the left side, the out of focus images on the right side. The defocus distance is 4 mm for both filters ($f/15$).

up to 60 nm. This implies a slight imprecision of coplanarity of the filters in the cold environment. The other measured coefficients associated with the different filters noticeably resemble each other. This is evidence that these narrow-band filters contribute little to the total aberration of the system and mainly the camera objective aberration is seen.

Fig. 5 displays the calibration results in the K-band. In total, 19 filters have been calibrated using the pinhole pair $0/4$ mm. One of the strongly aberrant filters (NB2.09) is highlighted by a dashed line. A large defocus in comparison to the other ones is detected. This filter is expected to have a striking error of coplanarity. It is not surprising that the strong defocus comes along with a particularly high spherical aberration ($i = 11$). The other highly aberrant filters show the same behaviour in comparison with the common filters of minor aberrations. The spherical aberration expresses the next order of a radial symmetric Zernike mode. The probability that a strong default of coplanarity induces only a defocus and does not concern higher orders is small. The PD input images of this aberrant filter is depicted at the bottom of Fig. 6. The right image shows the PSF registered in focus, and the left image a PSF having introduced a defocus of 4 mm. Already the in-focus image reveals a strong degradation, but especially the phase inversion due to the high defocus can be clearly seen in the out of focus image. A bright spot emerges in the center of the “donut”. On the top of this couple of images another couple of images is de-

picted. These are the PD input data of a filter (NB2.06) with normal behaviour and without strong aberrations.

As described in Section 4.2 the median of each Zernike number of the whole set of vectors yield the vector describing the camera contribution.

The accuracy of separating the camera aberrations from the raw aberrations (filters including camera) by the method described above is striking. The median aberrations for the filters of the two different wavelength regions plotted in Fig. 4 and Fig. 5 are compiled in Table 2. The deviations of both median values are clearly below the expected error (see Subsection 4.1). Table 2 lists these median coefficients taken from all NB filters in J-, H- and K-band. Keeping in mind that the achievable precision is a few nm we state that the camera aberrations are very small. The highest contributions arise from the focus term (4) and the astigmatism (5, 6). Section 5.3 gives an idea of the impact on the image quality dealing with Zernike mode aberrations in this order of magnitude. The residual filter aberrations are obtained by Equation 4. In general, besides the focus coefficient and a few deviating filters these values are close to zero, too.

4.4. Calibration of NAOS: dichroics

The calibration data are obtained with the fiber at the entrance focal plane of NAOS using the adaptive optic system itself for defocusing (see Subsection 3.2). Since the Zernike coefficients for the NAOS dichros are determined differentially, i.e. by subtraction of the total CONICA aberrations from the NAOS-CONICA overall aberrations, we can choose any reference camera and filter to perform the measurements as long as the components stay the same. A good choice is camera objective C50S and filter FeII1257. This objective oversamples even in the J-band and the filter has a small wavelength and therefore yields a higher accuracy in sensing wavefront errors. A suitable distance for the focus shift at this filter wavelength in the $f/15$ -beam is 2 mm. We can calculate suitable defocus distances using Equation 1.

In the following the properties of the five NAOS dichroics are itemized:

- VIS: Visible light to WFS; J, H, K, L and M to CONICA
- N20C80: 20% of the incoming light to WFS; 80% to CONICA (J, H, K)
- N90C80: 90% of the incoming light to WFS; 10% to CONICA (J, H, K)
- K: K to NAOS; J and H to CONICA
- JHK: J, H and K to NAOS; L and M to CONICA

Four of these five dichroics have been calibrated. The dichroic JHK is omitted since only light in L and M band reaches CONICA. It is unreliable to sense the small wavefront errors of NAOS-CONICA at these wavelengths. Furthermore there is no need to, because the small static aberrations become completely negligible in L and M.

Table 2. Camera aberrations in nm RMS by the median over the filter + camera aberrations in the bands J, H (pinholes 0/2 mm), K (pinholes 0/4 mm) and all bands (J, H, K).

Bands for Median	4	5	6	7	8	9	10	11	12	13	14	15
J, H	-12	-34	30	-6	4	2	13	-8	2	0	9	-1
K	-15	-39	27	-10	7	3	15	-9	2	0	6	-1
J, H, K	-15	-39	27	-9	6	2	13	-9	2	0	7	-1

The calibration results are compiled in Table 3 and Table 4. The first table lists the direct PD results. Any correction performed by these coefficients would only apply to the instrument configuration that was used to obtain the calibration data. The second table lists the aberrations directly assigned to the dichroics. These were obtained by subtracting the total CONICA aberrations that have been measured with the same filter and camera objective using the Zernike tool. It is noteworthy that the sensed astigmatism (Zernike number 5, 6) in the separated case is higher than in the overall case. Obviously a part of the camera astigmatism is compensated by the dichroics.

It is noteworthy that this tendency applies for all dichroics. Different reasons can cause this behaviour. First, the inclination of the dichroics artificially introduce an astigmatism. Even if the NAOS dichroics are designed for prism shape and do correct for this effect, a residual error cannot be excluded. Furthermore a certain amount of astigmatism can be introduced by components other than the dichroics lying in the same part of the light path, e.g. the parabolic output folding mirror or the CONICA entrance window (see Fig. 1). Nevertheless, it is not a limitation of the calibration method but only a question of assigning the contribution of the wavefront errors to the different optical components. In the end, only the sum of all aberrations has to be correct.

5. Image quality versus estimated aberrations

5.1. Strehl ratios by PD and focal plane image

The PD calibration data can be used to investigate the available image quality in different ways. First, the knowledge of the wavefront allows us to calculate a SR. After the reduction of the calibration data the wavefront is described by a set of Zernike coefficients. Furthermore, we can just refer to the in-focus image and calculate a SR with the measured point spread function. In the following we give a more detailed explanation of how these SRs are obtained.

Strehl by PD

For small wavefront deviations the SR can be determined via the coherent energy referring to the wavefront vari-

ance σ^2 in radian. The PD estimation yield the wavefront expanded in terms of Zernike coefficients a_i . For small σ^2

$$\text{SR} \simeq e^{-\sigma^2} \simeq 1 - \sigma^2 \simeq 1 - \sum_{i=4}^m a_i^2 \quad (5)$$

allows us to calculate the SR directly by the output of PD estimation. In principal the sum runs to infinity ($m = \infty$) but for our purpose we stop at $m = 15$. We can compare these SR numbers to the ones that are directly determined by the in-focus images.

Strehl on image

A straight-forward way to calculate a SR on the focus image (PSF) is to construct a theoretical diffraction-limited image PSF_{diff} taking into account the wavelength, the pixel scale, the aperture and the central obscuration.⁸ Having normalized the total intensity of the PSF and PSF_{diff} to 1, the fraction of these values yields the SR (see Equation 6).

In particular, in the case of the PSF sampling being close to the Nyquist criterion this approach has the disadvantage of being sensitive to the exact position of the PSF peak with respect to the pixel center. Furthermore, since the total intensity has to be determined by the integrated signal over a wider region around the PSF, the reliability of the SR value depends on a precise background correction. If the background is overestimated, then the SR will be overestimated, too, and vice versa. The reliability of the SR values can be enhanced when we switch from the image space to the Fourier space by

$$\text{SR} = \frac{\text{PSF}(\alpha = 0)}{\text{PSF}_{\text{theo}}(\alpha = 0)} = \frac{\int \text{OTF}(f)}{\int \text{OTF}_{\text{theo}}(f)}, \quad (6)$$

where OTF is the optical transfer function. Since in Fourier space only spatial frequencies are considered, a shift of the PSF is of no importance anymore. Aside from that, an elegant and reliable background correction can be performed using the zero spatial frequency. We calculate the SRs by the following procedure⁹:

- The image is corrected by its corresponding background.

⁸ The central obscuration which is caused by the secondary mirror will decrease slightly the central peak intensity and raise the side lobes of the Airy function.

⁹ This routine was developed by T. Fusco, ONERA

Table 3. NAOS dichros, overall NAOS-CONICA aberrations in nm RMS, reference filter: FeII1257.

Dichro	4	5	6	7	8	9	10	11	12	13	14	15
VIS	15	-5	24	-6	23	5	-8	-9	7	-13	-7	3
N20C80	2	-1	42	-2	30	5	-4	14	-1	-19	-8	4
N90C10	-7	-3	36	-3	19	6	-5	-28	1	-9	-9	1
K	-8	14	-17	-4	18	3	-5	-6	7	-14	-10	2

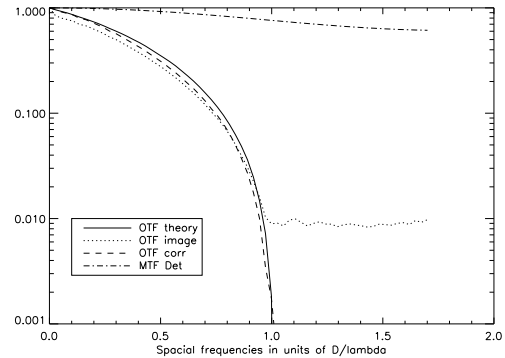
Table 4. Separate NAOS dichro aberrations in nm RMS.

Dichro	4	5	6	7	8	9	10	11	12	13	14	15
VIS	-18	37	-5	3	16	2	-21	-44	6	-13	-14	6
N20C80	-32	41	13	8	23	1	-17	-21	-3	-18	-15	6
N90C10	-41	38	7	7	12	3	-18	-64	-1	-9	-16	4
K	-42	56	-47	5	11	-1	-18	-42	5	-14	-17	5

- The OTF is calculated. It is given by the real part of the Fourier transform of the image.
- The residual background is corrected by the zero frequency. A fit of the very first spatial frequencies is used to extrapolate the true zero frequency value. The difference of the measured and the extrapolated value for the zero frequency yields the residual background.
- The noise level is subtracted using the high frequencies beyond the diffraction limit.
- A theoretical telescope OTF is constructed and multiplied by a Bessel function to account for the spatial spread due to the object size.
- The detector response is taken into account. This is done by a further multiplication of the theoretical telescope OTF with the Fourier transform of the detector response.
- The SR is obtained by the division of the normalized integrals from the measured and the theoretical OTF. All points with spatial frequencies higher than the diffraction limit are excluded.

Fig. 7 gives an example of how the described OTFs look for a PSF taken through the filter H2(1.0)S(7). To allow for a one dimensional representation, the circular mean of the two dimensional OTF is calculated. The raw, untreated OTF of the image is labeled “OTF image”. “OTF corr” displays the image OTF which is corrected for the residual background at zero frequency and for the noise level. The noise level of the uncorrected image OTF can be seen as plateau beyond the cutoff frequency D/λ and averages 1 % of the maximum value. The theoretical diffraction-limited telescope OTF includes the reduction due to the object size¹⁰ and the correction for the detector modulation transfer function (MTF). It is labeled “MTF Det” and is located at the top of the plot. The detector response is constructed by the assumption that roughly 5% of the total intensity is contained at each of the adja-

¹⁰ Since a $10\ \mu\text{m}$ pinhole is used which is barely resolved, the influence is small and amounts for the smallest wavelength ($1\ \mu\text{m}$) at most to 5%.

**Fig. 7.** Visualization of the theoretical and measured OTF at the example of filter FeII1644.

cent pixels and 2.5 % in the corner pixels. Here we refer to Finger et al. (2000). In this paper a measurement of the response of a comparable infrared array is described. Due to the lack of precise knowledge of the detector response it is constructed by a linear scaling in relation to the different pixel size.¹¹ A linear scaling is implied by a linear behaviour of the diffusion length of the minority carriers in the detector material.

5.2. Comparison of Strehl ratios

The resulting SRs for the narrow-band filters in J, H, and K are presented in Fig. 8 and Fig. 9. For each filter two SRs are given: the SR by PD and the SR on the image.

A number of error sources contribute to the error of the SR values on image. Beside of small error contributions due to uncertainties of the pixel scale and the flatfield, the remaining uncertainty of the background correction

¹¹ The pixel size of the CONICA detector (Aladdin 1Kx1K) is $27.0\ \mu\text{m}$.

and the detector response lead us to estimate an absolute error of $\pm 4\%$. The expected wavelength dependency of the MTF error is minor with respect to the remaining background error. Therefore it is neglected and we use the constant value given above derived from experience in reducing the experimental data.

Recall that the SR by PD has a *maximal* wavelength-dependent error of $\pm 5\%$ at $1\ \mu\text{m}$ and $\pm 1\%$ at $2\ \mu\text{m}$ taking into account an error of $\pm 35\ \text{nm}$ RMS for the focus estimation ($i = 4$). The main contributor to this error is a systematic error in the precision of the pinhole positions in the Zernike tool (see Section 4.1).

In general the PD SRs exceed the other SR values. This reflects the fact that the wavefront is expanded by a limited number of Zernike coefficients and the higher order aberrations are cut off. Note that it is not astonishing that in the case of very low SR values (worse than 50%) the PD SR value may lie below the image SRs (Fig. 9). Such strong wavefront errors violate the condition under which Equation 5 is valid. Thus, we expect Equation 5 to yield underestimated values.

The comparison of the SR values determined by the different methods turns out to be consistent. The longer the wavelength, the more the image and PD SR values approach each other. This shows that the influence of aberrations scale with the wavelength. In other words, the fact that we cut off at a certain Zernike number ($i = 15$) has a greater impact at short wavelengths.

5.3. Focus adjustment

Having in mind the small estimated wavefront errors that we presented in the previous sections we become conscious of the required precision of the most trivial aberration we regard: the focus. It is striking that even in the focus determination we depend on the precision of PD calibration. This becomes evident when we look at the conventional procedure of focus tuning and regard the loss of SR caused by the detected aberrations.

To tune the focus of CONICA, the in-focus pinhole of the Zernike tool is imaged on the detector. Now, a focus curve is obtained by taking images at different axial positions of the detector stage (see Section 3.1). The maximum of the SRs indicate the proper focus position of the stage. The maximum of the obtained focus curves for the different cameras can be located with an accuracy of about 50 nm RMS. For this wavefront error, Equation 5 yields a loss of SR of 2.5% at a wavelength of $2\ \mu\text{m}$ and almost 10% at $1\ \mu\text{m}$. Thus, in particular in the J- and H-band, the inaccuracy of determining the focus only by moving the detector stage gives reason for a significant loss of SR. Furthermore the whole effort of fine-tuning for the remaining static wavefront aberrations becomes irrelevant when the remaining focus error is in the regime of the highest higher-level aberrations (Zernike number $i \geq 5$). Compare the aberrations for focus with astigmatism in Fig. 4 and 5. The only way to achieve a significant improvement of the

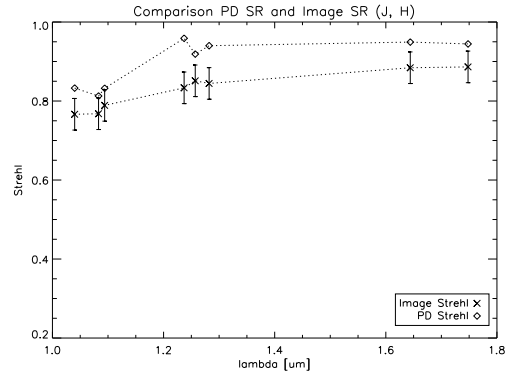


Fig. 8. Comparison of SR versus wavelength in J and H band calculated directly and derived from PD results.

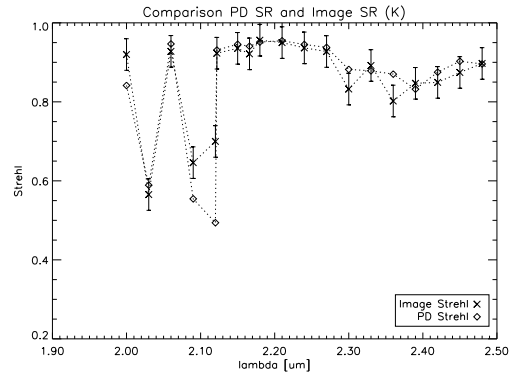


Fig. 9. Comparison of SR versus wavelength in K band calculated directly and derived from PD results.

wavefront error, and therefore of the SR after closed loop compensation, is to ensure that the residual focus deviation is corrected properly, too. This is guaranteed by following the procedure:

- Determination of the rough nominal focus position of the CONICA detector for each camera with one reference filter. The in-focus pinhole of the Zernike tool serves as a reference.
- Determination of the nominal focus for the whole instrument. The calibration point source in the NAOS entrance focal plane serves as a reference. The data points for the focus curve are obtained by moving the field selector in closed loop. This has to be done for every NAOS dichroic.
- The corresponding data base entries are updated by the nominal focus positions (CONICA internal and NAOS). The nominal focus deviations are included in the data base. They are compensated for by moving the field selector in the case of switching the NAOS

dichroics and by moving the CONICA detector stage in the case of switching the camera objectives.

- Then the PD estimation reveals the residual focus error for each configuration, in particular for each filter. They are entered into the data base together with the higher order aberrations. For a certain instrument configuration the corresponding values are fetched automatically and delivered to the AO system. The DM corrects for the residual focus deviations.

6. Closed loop compensation of NAOS-CONICA static aberrations

6.1. Rendering of aberrations

Having explored in detail the application of phase diversity to calibrate NAOS and CONICA static aberrations in Paper I, we presented above the experimental results applying PD as a wavefront sensor. We described how the contributions of the different optical components in the light path are separated to create a complete calibration configuration table. For each possible configuration of the instrument the corresponding correction coefficients are rendered to NAOS and are used to adjust the AO system. In this manner the DM will take the shape needed for compensation of the static wavefront 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. The gain will be quantified in terms of SR numbers.

6.2. Full AO correction

The $10\ \mu\text{m}$ calibration source in the entrance focal plane of NAOS simulates a star without turbulence. The visible WFS is used to correct for the common path aberrations. Therefore, the loop is closed on the $400\ \mu\text{m}$ source as described in Section 3.2. The light is separated by the dichroic VIS, thus the WFS sees the visible part and the near-infrared is directed towards CONICA.

Fig. 10 shows two extreme cases of applying AO compensation. The upper pictures demonstrate the correction for a filter in J-band, the pictures below in K-band. In accordance with Fig. 4 and 5 the sensed aberrations in J and K band are very similar - recall that the main contribution arises from the achromatic camera objective and the NAOS dichroic. But even if similar correction coefficients are rendered to the AO system, the effect on the image is strongly wavelength dependent. This is due to the fact that the influence of the applied Zernike coefficients scales with the wavelength. Thus, we achieve a striking correction in J-band visible with the naked eye on the images before and after correction. The most important aberration, the astigmatism, vanishes and the PSF is contracted. In K-band the non-corrected image is already very close to the optimum and the improvement is hard to see directly on the image. But calculating the SRs shows that even in K-

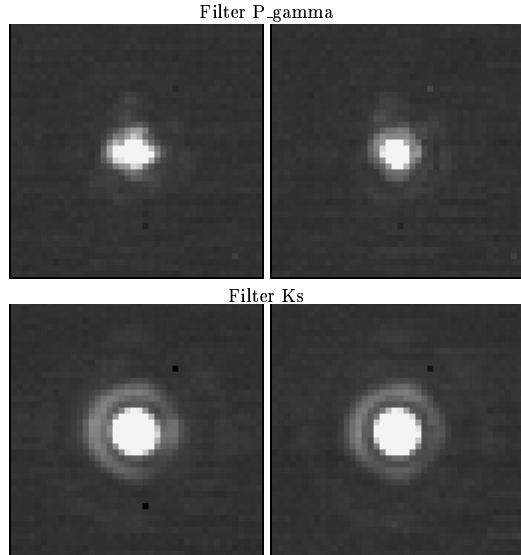


Fig. 10. Comparison of PSFs before and after closed loop compensation. Above a couple of J-band images and at the bottom a couple of K-band images are shown. Left side: without AO correction. Right side: with AO correction. Especially in J-band, the sharpening of the PSF can be clearly seen.

Table 5. Comparison of SRs for two selected filters before and after AO compensation for static aberrations. The maximum mainly systematic error is given. Statistical errors are significantly smaller.

Filter	SR no corr (%)	SR with corr (%)
Pgamma	60 ± 4	70 ± 4
Ks	91 ± 4	93 ± 4

band the performed correction is still significant (Table 5). Note that the given error arises from a maximum estimate of all error sources as described in Section 5.2. The nature of the error is mainly systematic (e.g., caused by background correction) and affects the calculated SRs for the image pairs in the same way. SRs determined on experimental data are intrinsically afflicted by rather high error bars, but a direct inspection of the images (central intensity, shape of the diffraction rings) shows the relative gain of 2 to 3% in K-band to be true. Even this rather small appearing gain in K is of high importance. On the way to scientific goals such as e.g. planet detection, the total error budget must be tackled to eliminate every percentage point of loss in SR.

7. Conclusion

In Paper I we presented a guideline for a PD approach to calibrate static wavefront aberrations. An extensive inves-

tigation of system limitations and error sources has been carried out. This approach was shown to be very flexible and powerful for precise wavefront sensing using experimental data of the first VLT AO system NAOS-CONICA. In Paper II we have given a detailed description of its implementation at the instrument and presented the experimental results of the calibration data for a variety of observing configurations. Especially, we turned our attention to the disentanglement of the measured overall wavefront errors which allows a convenient allocation to the diverse optical components and makes the calibration procedure feasible for an instrument with a huge number of possible configurations. The sensed wavefront errors expanded in Zernike coefficients have been used to quantify the image quality in terms of SR and be proven to be consistent with the SRs directly determined with the image data. Finally we gave a striking example of the acquired improvement of optical performance in comparing images with and without AO correction for static wavefront errors. It turned out that even starting with a very good image quality, we still could achieve a significant gain in terms of the SR ratio.

The design of the instrument control software is harmonized with this calibration procedure and the AO loop parameters are automatically updated when the instrument setup changes. This ensures that the utmost optical performance is provided for all the configurations. Additionally, this implementation of wavefront sensing can be used to monitor the optical quality and to alert for small degradations of optical performance. In general, for future high performance AO systems, the presented technique is of great importance to achieve the challenging science goals of the astronomical community. It should be included in the instrumental design at a very early phase.

Acknowledgements. We would like to thank Eric Gendron and Wolfgang Brandner for their patience and assistance in the fine tuning of the instrument and interpretation of the measurements during the commissioning runs. Furthermore, we thank Gert Finger for the fruitful discussions about the detector characteristics. We are thankful to Norbert Hubin for the assistance to this work on the part of the European Southern Observatory. This research was partially supported through a European Southern Observatory contract and the Marie Curie Fellowship Association of the European Community.

References

- Blanc, A., Fusco, T., Hartung, M., Mugnier, L., Rousset, G., A&A, **submitted (Paper I)**
- Brandner, W., Rousset, G., Lenzen, R., et. al., 2002, *The Messenger*, 107, p. 1-6
- Finger, G., Mehrgan, H., Meyer, M., Moorwood, A. F., Nicolini, G., Stegmeier, J., 1998, in A. Fowler (ed.), *Infrared Astronomical Instrumentation*, SPIE, 3354, p. 87-98
- Gonsalves, R. A., 1982, in *Optical Engineering*, Vol. 21-5, p. 829-832
- Hartung, M., Bizenberger, P., Boehm, A., Laun, W., Lenzen, R., Wagner, K., 2000, in M. Iye and A. Moorwood (eds), *Optical and IR Telescope Instrumentation and Detectors*, SPIE, 4008, 830
- Jefferies, S. M., Lloyd-Hart, M., Keith Hege, E. and Georges, J., *Applied Optics*, 2002, 41, 2095-2102
- Lenzen, R., Hofmann, R., Bizenberger, P., Tuschke, A., 1998, in A. Fowler (ed.), *Infrared Astronomical Instrumentation*, SPIE, 3354, 606
- Noll, R. J., 1976, *J. Opt. Soc. Am.*, Vol. 66-3, p. 207-211
- Paxman, R. G., Schulz, T. J., and Fienup, J.R.: 1992, *Journal of the Optical Society of America A* **9(7)**, 1072
- Rousset, G., Lacombe, F., Puget, P., Hubin, N., Gendron, E., Conan, J.-M., Kern, P., Madec, P.Y., Rabaud, D., Mouillet, D., Lagrange, A.-M., Rigaut, F., 1998, in D. Bonaccini and R. K. Tyson (eds.), *Adaptive Optical System Technologies*, SPIE, 3353, 508
- Rousset, G., Lacombe, F., Puget, P., Gendron, E., Arsenault, R., Kern, P., Rabaud, D., Madec, P.-Y., Hubin, N., Zins, G., Stadler, E., Charton, J., Gigan, P., Feautrier, P., 2000, in P. L. Wizinowich (ed.), *Adaptive Optical Systems Technology*, SPIE, 4007, 72
- Spanoudakis, P., Zago, L., Chételet, O., Gentsch, R., Mato Mira, F., in P. L. Wizinowich (ed.), *Adaptive Optical Systems Technology*, SPIE, Vol. 4007, p. 408

Annexe H

Article GRETSI

Estimation des aberrations d'un télescope optique par diversité de phase

Article paru dans « 17ième Colloque sur le Traitement du Signal et des Images », volume 1, pages 63-66, GRETSI

Estimation des aberrations d'un télescope optique par diversité de phase

Amandine BLANC^{1,2}, Laurent MUGNIER¹, Jérôme IDIER²

¹O.N.E.R.A,
B.P. 72, 92322 Châtillon Cedex, France

²Laboratoire des Signaux et Systèmes(CNRS-SUPELEC-UPS),
École Supérieure d'Électricité
Plateau de Moulon, 91192 Gif-sur-Yvette, France
blanc@onera.fr, mugnier@onera.fr, idier@lss.supelec.fr

Résumé – Cet article est consacré à l'estimation des aberrations optiques d'un télescope spatial d'observation de la Terre par diversité de phase. Deux estimateurs sont présentés et comparés à l'aide de simulations. Le premier, déjà utilisé dans la littérature, repose sur la recherche conjointe des aberrations et de l'objet observé par maximisation de la vraisemblance généralisée. Le second est un nouvel estimateur que nous proposons, basé sur la recherche des aberrations seules par maximum de vraisemblance.

Abstract – This paper deals with the estimation of the optical aberrations of a space telescope by phase diversity. Two estimators are proposed and compared by means of simulations. The first one, which was already implemented in the literature, is based on a joint estimation of the aberrated phase and the observed object through maximization of the generalized likelihood; the other one is a new approach, based on the sole estimation of the aberrated phase through maximum likelihood.

1 Introduction

La présence d'aberrations optiques dans un système d'imagerie (télescope par exemple), qu'elles soient induites par la turbulence atmosphérique ou par les défauts optiques de l'instrument, dégrade la qualité des images enregistrées. Lors de la formation de ces images, ces aberrations donnent naissance à des variations de la phase optique ϕ dans la pupille de l'instrument qui atténuent la fonction de transfert. On cherche à déterminer ces aberrations (c'est-à-dire à reconstruire la carte de phase) à partir de l'image acquise, afin de les corriger. Cette correction peut être effectuée *a posteriori* (restauration d'images) ou en temps réel (optique active ou adaptative).

La réponse impulsionnelle du système ainsi que l'objet observé étant inconnus, retrouver ces aberrations à partir de l'image enregistrée est un problème de déconvolution myope. Or, en l'absence d'hypothèses structurelles fortes sur la phase et sur l'objet, le déficit d'information limite fortement les performances de ce type de méthodes de déconvolution. Une technique simple et facile à mettre en œuvre, pour apporter plus d'information sur la phase, est d'acquérir simultanément plusieurs images liées entre elles par des relations parfaitement connues. Cette méthode s'appelle la diversité de phase[?, ?]. Dans la suite, nous nous proposons d'utiliser la diversité de phase pour estimer les aberrations optiques d'un télescope spatial d'observation de la Terre. Le modèle d'imagerie utilisé, ainsi que la diversité de phase seront présentés dans le paragraphe 2. Les travaux menés par différents auteurs [?, ?, ?] sur la reconstruction de surface d'onde par diversité de

phase sont tous basés sur la recherche conjointe de l'objet observé et des aberrations par maximisation de la vraisemblance généralisée. Une approche similaire est présentée dans le paragraphe 3. Mais comme nous le verrons, cet estimateur n'offre pas la cohérence statistique souhaitée. C'est pourquoi un nouvel estimateur marginalisé de type maximum de vraisemblance est proposé et présenté dans le paragraphe 4. Une comparaison des performances de ces deux estimateurs est faite, sur la base de simulations, dans le paragraphe 5.

2 Modèle d'imagerie et diversité de phase

Dans le domaine isoplanétique (c'est-à-dire le domaine où la réponse impulsionnelle demeure identique en tout point du champ), l'image acquise i est liée à la réponse impulsionnelle du système h et à l'objet observé o par la relation de convolution:

$$i = h * o + n \quad (1)$$

où n est un bruit additif. De plus, dans l'approximation du champ proche (scintillation négligeable), la réponse impulsionnelle s'écrit :

$$h = |TF(\exp(i\phi))|^2 \quad (2)$$

avec ϕ la phase dans la pupille du télescope et TF l'opérateur de transformée de Fourier. En pratique, les quantités accessibles sont discrètes du fait de l'échantillonnage des images. On peut donc récrire l'équation (1) sous forme

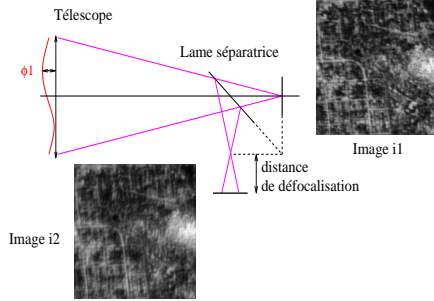


FIG. 1: *Principe de la diversité de phase*

où H est une matrice de Toeplitz à blocs Toeplitz correspondant à la convolution par la réponse impulsionnelle h et où i , o et n sont désormais les formes discrètes des grandeurs définies précédemment.

Le principe de la diversité de phase repose sur la mesure simultanée de plusieurs images (ici nous nous limiterons à deux). La première image i_1 est acquise dans le plan focal du système optique et la deuxième i_2 dans un plan défocalisé d'une distance d connue (voir figure 1). Les phases associées à chacun de ces plans d'observation sont liées par la relation :

$$\phi_2 = \phi_1 + \phi_d \quad (4)$$

avec ϕ_1 la phase dans le plan focal, ϕ_2 celle associée au plan défocalisé et ϕ_d la phase induite par la défocalisation (ϕ_d est parfaitement connue).

Le problème à résoudre est le suivant: à partir des deux images i_1 , i_2 et de la distance de défocalisation d et sans connaître l'objet o , on souhaite estimer les aberrations optiques c'est à dire reconstruire la phase ϕ_1 . Comme il a été indiqué dans l'introduction, la connaissance de ces aberrations permet, selon le dispositif expérimental, soit de les corriger en temps réel par un miroir déformable, soit de connaître la réponse impulsionnelle pour effectuer une déconvolution ultérieure de l'image dégradée.

3 Estimation conjointe

Nous nous intéressons, tout d'abord, à une reconstruction conjointe de l'objet et des aberrations. Cette approche est basée sur la recherche du Maximum de Vraisemblance Généralisé (MVG) défini ainsi :

$$(\hat{o}, \hat{\phi})_{\text{MVG}} = \arg \max_{o, \phi} f(i_1, i_2, o; \phi, \theta). \quad (5)$$

où θ représente les hyperparamètres du problème, et avec $f(i_1, i_2, o; \phi, \theta) = f(i_1|o; \phi, \theta)f(i_2|o; \phi, \theta)f(o; \theta)$. Le bruit est supposé blanc gaussien de moyenne nulle et de variance σ^2 (identique pour les deux images). On associe un *a priori* gaussien à l'objet de moyenne o_m et de matrice de covariance R_o . Nous supposons la loi de l'objet stationnaire donc que R_o est une matrice de Toeplitz.

Les matrices de Toeplitz peuvent être approximées par des matrices circulantes (cette approximation correspondant à une périodisation). Sous cette hypothèse, R_o , H_1 et H_2 sont diagonalisables dans la base des exponentielles de Fourier discrètes. On peut alors écrire le critère à minimiser $L_{\text{MVG}}(o, \phi) = -\ln(f(i_1, i_2, o; \phi, \theta))$ dans le domaine de Fourier :

$$L_{\text{MVG}}(o, \phi) \propto \sum_f \left[\frac{|\tilde{h}_1(f)\tilde{o}(f) - \tilde{i}_1(f)|^2}{\sigma^2} + \frac{|\tilde{h}_2(f)\tilde{o}(f) - \tilde{i}_2(f)|^2}{\sigma^2} + \frac{|\tilde{o}(f) - \tilde{o}_m(f)|^2}{S_o(f)} \right] \quad (6)$$

où le tilde représente la transformée de Fourier discrète de la grandeur, f la fréquence spatiale et S_o la densité spectrale de puissance de l'objet.

L'annulation de la dérivée partielle du critère L_{MVG} par rapport à l'objet fournit une expression analytique $\tilde{o}(\phi)$ qui minimise le critère à ϕ fixée. En injectant \tilde{o} dans le critère, on fait disparaître la dépendance vis-à-vis de l'objet. L'avantage de ce critère $L'_{\text{MVG}}(\phi) = L_{\text{MVG}}(\tilde{o}(\phi), \phi)$ est de réduire l'espace des solutions à explorer. En revanche, L'_{MVG} et L_{MVG} ne sont pas convexes et possèdent le même nombre de minima locaux. En effet on montre que $\frac{\partial L'}{\partial \phi} = \frac{\partial L}{\partial \phi} \Big|_{o=\tilde{o}(\phi)}$ ainsi ϕ est un extremum de L'_{MVG} si et seulement si $(\tilde{o}(\phi), \phi)$ est un extremum de L_{MVG} .

Il reste à préciser le modèle de densité spectrale de puissance de l'objet que nous avons choisi[5] :

$$S_o(f) = E[|\tilde{o}(f) - \tilde{o}_m(f)|^2] = k/[1 + (f/f_o)^p] - |\tilde{o}_m(f)|^2 \quad (7)$$

avec E l'espérance mathématique. Ce modèle est bien adapté aux images terrestres que nous traitons. Il requiert cependant le choix de 3 hyperparamètres k , f_o et p .

Intéressons-nous aux propriétés de l'estimateur du Maximum de Vraisemblance Généralisé. Il est connu comme étant asymptotiquement biaisé et possédant une variance asymptotique non nulle[6]. Par ailleurs, nous avons cité plus haut le problème de la non-convexité du critère MVG. Enfin, la mise en œuvre de cet estimateur nécessite le choix d'hyperparamètres $\theta = (\sigma, o_m, k, f_o, p)$; or θ ne peut pas être estimé en même temps que o et ϕ (problème de dégénérescence du critère). Ce qui signifie que ces hyperparamètres devront être choisis de façon arbitraire. Pour résoudre ces problèmes, nous nous tournons vers un autre estimateur qui présente de meilleures propriétés.

4 Estimation marginale

L'objet observé peut être considéré comme un paramètre de nuisance puisqu'il ne nous importe pas de le reconstruire. Nous proposons, dans la suite, de l'intégrer hors du problème et d'estimer la phase seule ou bien la phase et les hyperparamètres.

L'estimateur du maximum de vraisemblance pour ϕ s'écrit :

$$\begin{aligned} \hat{\phi}_{\text{MV}} &= \arg \max_{\phi} f(i_1, i_2; \phi, \theta) \\ &= \arg \max_{\phi} \int f(i_1, i_2, o; \phi, \theta) do \end{aligned} \quad (8)$$

L'estimateur du maximum de vraisemblance pour ϕ s'écrit :

$$\begin{aligned}\hat{\phi}_{\text{MV}} &= \arg \max_{\phi} f(i_1, i_2; \phi, \theta) \\ &= \arg \max_{\phi} \int f(i_1, i_2, o; \phi, \theta) do\end{aligned}\quad (8)$$

Nous avons montré que maximiser $f(i_1, i_2; \phi, \theta)$ revient à minimiser le critère suivant :

$$L_{\text{MV}}(\phi) \propto \ln |R_I| + (I - m_I)^T R_I^{-1} (I - m_I) \quad (9)$$

Avec $I = (i_1 \ i_2)^T$, vecteur concaténant les données, $m_I = (H_1 o_m \ H_2 o_m)^T$, celui qui contient les images moyennes, R_I la matrice de covariance de I et $|R_I|$ son déterminant.

Les deux critères L_{MV} et L'_{MVG} sont liés par l'identité remarquable [?] :

$$L_{\text{MV}} \propto \ln |R_I| + L'_{\text{MVG}} \quad (10)$$

Ils ne diffèrent donc que par un terme, qui dans l'espace de Fourier, s'écrit :

$$\begin{aligned}\ln |R_I| &= \sum_f \ln S_o(f) + N^2 \ln \sigma^2 \\ &+ \sum_f \ln \left(|\tilde{h}_1(f)|^2 + |\tilde{h}_2(f)|^2 + \frac{N^2 \sigma^2}{S_o(f)} \right)\end{aligned}\quad (11)$$

avec N^2 le nombre de pixels dans l'image.

Sous des conditions mathématiques peu restrictives, l'estimateur du maximum de vraisemblance présente un biais et une variance asymptotiques nuls et d'après un théorème de type centrale limite, la nappe de vraisemblance devient de plus en plus gaussienne (donc présente de moins en moins de minima locaux) lorsqu'on augmente la taille des images traitées. En supposant que ces conditions sont réalisées, ce qui fera l'objet d'une étude formelle spécifique, cet estimateur résout asymptotiquement les problèmes que présentent l'estimateur MVG. De plus les hyperparamètres peuvent être estimés en même temps que la phase. En effet, on peut écrire :

$$(\hat{\phi}, \hat{\theta})_{\text{MV}} = \arg \max_{\phi, \theta} f(i_1, i_2; \phi, \theta)$$

Les hyperparamètres à estimer sont σ^2 , k , f_o et p (nous avons supposé l'objet moyen nul). En effectuant le changement de variables $\mu = \frac{\sigma^2}{k}$, on peut écrire l'expression analytique de \hat{k} qui est tel que $\frac{\partial L_{\text{MV}}}{\partial k} \Big|_{k=\hat{k}} = 0$. Le critère à minimiser s'écrit alors $L_{\text{MV}}(\phi, \mu, f_o, p)$.

L'estimateur marginalisé que nous proposons a donc, *a priori*, des propriétés asymptotiques intéressantes et permet une estimation «non supervisée» de la phase. Il reste à mettre en œuvre ces deux estimateurs sur la base de traitements de données simulées.

5 Simulations

Dans la suite, nous avons décomposé la phase sur la base des polynômes de Zernike, qui forment une base orthonormée sur un support circulaire et qui sont bien adaptés pour décrire les aberrations optiques. La phase a été simulée à partir des 21 premiers polynômes de Zernike (et nous chercherons à reconstruire ces 21 premiers modes); en effet, on

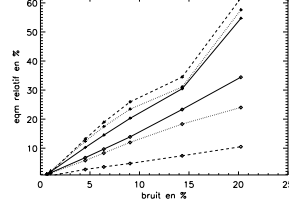


FIG. 2 – comparaison de l'erreur quadratique relative en % sur l'estimation des modes avec les hyperparamètres vrais, commise avec l'estimateur MV [losanges] et avec l'estimateur MVG [croix]. En [traits pleins] pour une image 32×32 pixels, [pointillés] pour une image 64×64 , [tirets] pour une image 128×128 .

peut considérer que les seuls premiers polynômes suffisent à décrire les aberrations qui affectent un télescope spatial. L'amplitude de la phase aberrante ainsi simulée est de $\frac{\lambda}{8}$ rms. La phase correspondant au plan défocalisé est calculée à partir de la phase plan focal à laquelle on ajoute une défocalisation telle que l'amplitude crête à crête de la défocalisation ϕ_d soit égale à 2π radians. Les images simulées (présentées sur la figure ??) sont obtenues par convolution d'une scène terrestre avec les réponses impulsionnelles associées à chaque plan d'observation et bruitées avec deux réalisations différentes d'un bruit blanc gaussien stationnaire. Quant à la minimisation des critères, elle est assurée par un algorithme de type gradient à pas adaptatif.

5.1 Comparaison à hyperparamètres «vrais»

Nous présentons, tout d'abord, (voir figure ??) une comparaison des performances des deux estimateurs, pour différents niveaux de bruit et différentes tailles d'images, avec les «vrais» hyperparamètres, c'est à dire ceux obtenus en ajustant le modèle de densité spectrale de puissance S_o à l'aide de l'objet vrai.

Quel que soit le bruit et la taille de l'image traitée, l'estimateur MV se révèle bien plus performant que l'estimateur MVG pour la reconstruction des modes. Lorsque l'on augmente le nombre de données traitées, l'erreur sur la reconstruction des modes diminue pour l'estimateur MV alors qu'elle reste stable voire augmente avec le MVG; c'est la différence de comportement asymptotique des deux estimateurs que nous avons annoncée dans les paragraphes précédents.

5.2 Influence de l'hyperparamètre global

Nous allons à présent faire varier l'hyperparamètre «global» λ défini par $S'_o = S_o/\lambda$, et étudier son influence sur la reconstruction des modes et de l'objet (faire varier λ autour de 1 revient à faire varier k autour de sa valeur «vraie» définie plus haut). Pour être exhaustif, il aurait fallu également faire varier les autres hyperparamètres. La figure ?? présente les erreurs résiduelles sur l'objet et sur les modes après reconstruction par MVG et celles après reconstruction par MV, pour un bruit de 14%.

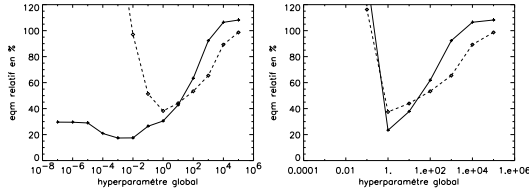


FIG. 3 – Comparaison de l’erreur quadratique relative en % sur l’estimation de l’objet [losanges-tirets] et sur l’estimation des modes [croix-trait plein] pour différentes valeurs de λ et un bruit de 14%. Le graphique de gauche correspond à l’estimateur MVG, celui de droite à l’estimateur MV.

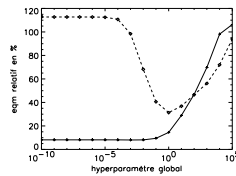


FIG. 4 – Comparaison de l’erreur quadratique relative en % sur l’estimation de l’objet [losanges-tirets] et sur l’estimation des modes [croix-trait plein] par l’estimateur MVG pour différentes valeurs de λ et un bruit de 6%.

L’estimateur MV présente une bonne cohérence puisque le meilleur réglage de l’hyperparamètre pour la reconstruction de l’objet λ_{opt}^o correspond au meilleur réglage pour les modes λ_{opt}^ϕ , de plus nous verrons dans la sous-section suivante qu’il est possible d’estimer cet hyperparamètre optimal en même temps que les modes. Pour l’estimateur MVG, λ_{opt}^o est égal à 1 alors que λ_{opt}^ϕ est approximativement égal à $1/250$. Comme nous cherchons à estimer les aberrations sans nous intéresser à l’objet reconstruit, avec l’estimateur conjoint, au vu de nos simulations, il semble que l’on obtienne une meilleure reconstruction des modes en sous-régularisant. Notons que l’erreur résiduelle sur les modes reconstruits au «meilleur réglage» de l’hyperparamètre pour l’estimation conjointe est légèrement inférieure à celle obtenue avec l’estimation marginale, mais contrairement au λ_{opt} du MV, celui associé au MVG ne peut être estimé de façon réaliste. La figure ?? présente la même comparaison erreur-objet / erreur-modes avec l’estimateur MVG mais pour un bruit plus faible (6%). On observe toujours $\lambda_{opt}^o \neq \lambda_{opt}^\phi$ mais le meilleur réglage pour les modes correspond à n’importe quel $\lambda_{opt}^\phi < 1/100$, c’est-à-dire qu’une sous-régularisation «extrême» donne de bons résultats sur la reconstruction des modes (il est à noter qu’il n’y a plus alors que du bruit dans l’objet reconstruit). Nous avons observé ce comportement surprenant dès lors que le niveau de bruit passe sous un certain seuil (dans nos simulations 10% pour une image 64×64 pixels).

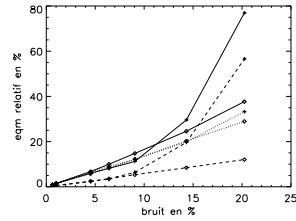


FIG. 5 – comparaison de l’erreur quadratique relative en % sur l’estimation des modes par l’estimateur MV [losanges] et par l’estimateur MVG [croix], en non-supervisé. En [trait plein] pour une image 32×32 pixels, [pointillés] pour une image 64×64 , [tirets] pour une image 128×128 .

5.3 Estimation non supervisée

Afin de se placer dans des conditions réalisables en pratique, nous allons comparer l’estimateur du MVG pour une sous-régularisation «extrême» ($\lambda = 10^{-10}$), puisque c’est ainsi que le MVG fournit la meilleure estimation des modes, et nous allons estimer les hyperparamètres en même temps que les modes pour l’estimateur MV. Les résultats, pour différentes tailles d’images et de niveaux de bruit sont présentés figure ??.

L’estimation MV non supervisée donne des résultats presque aussi bons que ceux que l’on avait obtenus avec l’estimateur MV à hyperparamètres «vrais» (voir figure ??). Ces courbes montrent, de plus, que l’estimation marginale «non supervisée» permet une bien meilleure reconstruction des modes que l’estimation conjointe avec régularisation quasi-nulle, lorsque le bruit est important et d’aussi bons résultats pour des niveaux de bruit plus faibles. Pour un coût de calcul quasiment identique, les résultats obtenus avec l’estimateur marginalisé sont également bien cohérents d’un point de vue méthodologique.

6 Conclusion

Les propriétés des deux estimateurs ont été vérifiées à l’aide des simulations. Le bon comportement du MVG pour une régularisation quasi-nulle reste à expliquer, mais justifie, en partie, ce qui est souvent fait dans la littérature. Toutefois l’estimateur marginalisé que nous avons proposé possède des propriétés statistiques plus intéressantes, facilitant, entre autres, la mise au point de méthodes non supervisées.

Références

- [1] Gonsalves R. A. *Phase retrieval and diversity in adaptive optics*. Optical Engineering, 21 (5) (1982), pp. 829–832.
- [2] Paxman R. G., Schulz T. J. et Fienup J. R. *Joint estimation of object and aberrations by using phase diversity*. Journal of the Optical Society of America A, 9 (7) (1992), pp. 1072–1085.
- [3] Meynadier L. *Analyse de surface d’onde pour le contrôle actif d’un télescope spatial*. Thèse de doctorat, Université de Nice-Sophia Antipolis (1997).

- [4] Lee D. J., Roggemann M. C., Welsh B. M. et Crosby E. R. *Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing*. Appl. Opt. (décembre 1997), pp. 9186–9197.
- [5] Kattinig A. P. et Primot J. *Model of the second-order statistic of the radiance field of natural scenes, adapted to system conceiving*. Dans *Aeorosense'97 - Visual information processing VI*, vol. 3074, Washington (1997). SPIE, Soc. Photo-Opt. Instrum. Eng.
- [6] Champagnat F. et Idier J. *An alternative to standard maximum likelihood for Gaussian mixtures*. Dans *ICASSP*, pp. 2020–2023, Detroit, U.S.A (mai 1995).
- [7] Goussard Y., Demoment G. et Idier J. *A new algorithm for iterative deconvolution of sparse spike*. Dans *ICASSP*, pp. 1547–1550, Albuquerque, U.S.A. (avril 1990).

Annexe I

Article SPIE

NOVEL ESTIMATOR FOR THE ABERRATIONS OF A SPACE TELESCOPE BY PHASE DIVERSITY

Article paru dans « UV, Optical, and IR Space Telescopes and Instruments », volume 4013, pages 728-736, SPIE

Novel estimator for the aberrations of a space telescope by phase diversity

Amandine Blanc^{1,2}, Jérôme Idier², Laurent M. Mugnier¹

¹Office National d'Études et de Recherches Aérospatiales
Division Imagerie Optique Haute Résolution
BP 72, F-92322 Châtillon cedex
E-mail: <author>@onera.fr

²Laboratoire des Signaux et Systèmes
École Supérieure d'Électricité
Plateau de Moulon, 91192 Gif-sur-Yvette, France
E-mail: <author>@lss.supelec.fr

ABSTRACT

In this communication, we propose a novel method for estimating the aberrations of a space telescope from phase diversity data. The images recorded by such a telescope can be degraded by optical aberrations due to design, fabrication or misalignments. Phase diversity is a technique that allows the estimation of aberrations. The only estimator found in the relevant literature is based on a joint estimation of the aberrated phase and the observed object. By means of simulations, we study the behavior of this estimator. We propose a novel marginal estimator of the sole phase by Maximum Likelihood. It is obtained by integrating the observed object out of the problem; indeed, this object is a nuisance parameter in our problem. This reduces drastically the number of unknown and provides better asymptotic properties. This estimator is implemented and its properties are validated by simulation. Its performance is equal or even better than that of the joint estimator for the same computing cost.

Keywords: phase diversity, wavefront sensing, Bayes estimation, telescopes

1. INTRODUCTION

The images recorded by a telescope are degraded by aberrations. Indeed, aberrations lead to phase variations in the pupil plane, which severely reduces the optical transfer function (OTF). They must be estimated in order to be corrected, either in real-time (active optics) or *a posteriori* (image restoration). In the case of a space telescope, these aberrations may originate in thermal and mechanical disturbances as well as in the design, fabrication and assembly of the instrument. The method of phase diversity, first proposed by Gonsalves¹, makes it possible to measure the wavefront^{2,3,4}. This technique uses a low-cost, optically simple wavefront sensor and it can work with an extended scene. It consists in simultaneous acquisition of the usual focal plane image and of an additional image with a known defocus. The aberrations are estimated numerically from these two images, by solving an optimization problem. So far, the estimation used in the literature is based on a joint estimation of the aberrations and the observed object. In this communication, we propose a novel marginal estimator of the sole interesting parameters: the aberrations. It is obtained by integrating the observed object out of the problem. This paper is organized as follows. In Section 2, we present the mathematical formulation of the problem and a brief presentation of our simulation method, in Section 3 we recall the usual joint estimation scheme based on Generalized Maximum Likelihood. We present the new marginal estimator based on the Maximum Likelihood and its properties in Section 4. In the last section, we compare the performance of the two estimators.

2. DATA MODEL AND SIMULATION METHOD

In the isoplanatic patch of the telescope, the image is the noisy convolution of the point-spread function h in the observation plane and the object o :

$$i(r) = (h * o)(r) + n(r) \quad (1)$$

where r is a two-dimensional vector in the image plane and n is an additive noise. Under near field approximation, the point-spread function associated with the focused image is given by:

$$h_1(r) = |FT[A(u) \cdot \exp(i\phi(u))]|^2 \quad (2)$$

where u is a two-dimensional vector in the pupil plane, ϕ is the unknown aberrated phase function, A is the binary aperture function and FT denotes the Fourier transform. In the defocused plane:

$$h_2(r) = |FT[A(u) \cdot \exp(i(\phi(u) + \phi_d(u)))]|^2 \quad (3)$$

where ϕ_d is the known diversity phase function. The aberrated phase function is expanded onto a set of Zernike polynomials. Indeed, the presence of aberration in an optical system can be mathematically represented by Zernike polynomials⁽⁵⁾.

$$\phi(u) = \sum_{i=1}^n a_i Z_i(u) \quad (4)$$

In the case of a space telescope, only the first polynomials are enough to describe the aberrations. The coefficients a_{1-3} will not be estimated: the piston coefficient a_1 is a constant added to the phase and has no influence on the point-spread function and the tilts coefficients a_2, a_3 introduce a shift in the image that is of no importance for extended objects.

In practice, data are discrete arrays because of the spatial sampling of the images. Equation 1 takes the form:

$$i = Ho + n \quad (5)$$

where H is the Toeplitz matrix corresponding to the convolution by h , and where i , o and n are the discrete forms of the previous variables.

The problem is to estimate the aberration parameters (the set of a_i) from the data (focused i_1 and defocused i_2 images) and the defocused distance, without knowing the object o .

In the following, we shall study the properties of estimators by means of simulations. The simulations have been obtained in the following way: our object is an Earth view. The phase is generated with the first 21 Zernike polynomials (their values are listed in Table 1) and the estimated phase will be expanded on the same polynomials. The defocused amplitude for the second observation plane is 2π rad ($a_4 = 1.814$ rad). The simulated images are monochromatic and are sampled at the Shannon rate. They are obtained by the convolution (made by FFT) of the point-spread function and the object (see Figure 1) and corrupted by a stationary white Gaussian noise.

a_i	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
(rad)	-0.2	0.3	-0.45	0.4	0.3	-0.25	0.35	0.2	0.1
a_i	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}
(rad)	0.05	-0.05	0.05	0.02	0.01	-0.01	-0.02	0.01	0.01

Table 1. Values of the coefficients used for simulations

3. JOINT ESTIMATION

3.1. Criterion

We recall the general approach for jointly estimating the object and the aberrations, which is sometimes called the Generalized Maximum Likelihood (GML) approach. The GML estimator is defined by

$$(\hat{o}, \hat{\phi})_{\text{GML}} = \arg \max_{o, \phi} f(i_1, i_2, o; \phi, \theta) \quad (6)$$

$$= \arg \max_{o, \phi} f(i_1|o; \phi, \theta) f(i_2|o; \phi, \theta) f(o; \theta) \quad (7)$$

where $f(i_1, i_2, o; \phi, \theta)$ is the joint probability density function of the data (i_1, i_2) and of the object o . It also depends on the deterministic aberrated phase ϕ and on the set of hyperparameters θ . $f(i_1|o; \phi, \theta)$ and $f(i_2|o; \phi, \theta)$ denote the

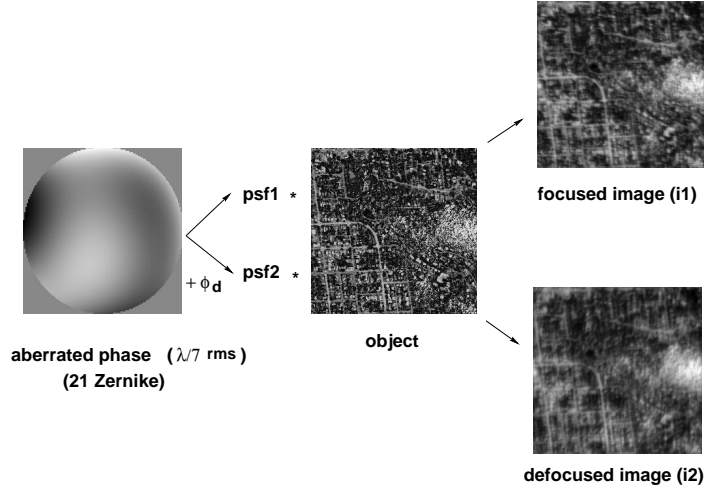


Figure 1. Schematic diagram of simulations

likelihood of the data i_1 and i_2 and $f(o; \theta)$ is the *a priori* probability density function of o . We do not assume a prior probability function on ϕ but the phase is expanded on a finite linear combination of the Z_i functions, which is an implicit regularization. The noise is a stationary white Gaussian noise with a variance σ^2 (the same for the two images). We assume that the *a priori* statistics on the object are Gaussian with a mean of o_m and a covariance matrix R_o . Hence, we have

$$f(i_1, i_2, o; \phi, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \|i_1 - H_1 o\|^2\right) \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \|i_2 - H_2 o\|^2\right) \quad (8)$$

$$\times \frac{1}{\sqrt{2\pi|R_o|}} \exp\left(-\frac{1}{2}(o - o_m)^t R_o^{-1} (o - o_m)\right)$$

where $|R_o|$ denotes the determinant of R_o . We assume that $(o - o_m)$ is stationary, so that the covariance matrix R_o is Toeplitz. Toeplitz matrices can be approximated by circulant matrices with the approximation corresponding to a periodization. Under this assumption, the covariance matrix R_o and the convolution matrices H_1 and H_2 are diagonalized by a discrete Fourier transform (DFT). We can write

$$R_o = F^{-1} \text{diag}[S_o] F$$

$$H_1 = F^{-1} \text{diag}[\tilde{h}_1] F$$

$$H_2 = F^{-1} \text{diag}[\tilde{h}_2] F$$

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 power spectral density of the object. Maximizing $f(i_1, i_2, o; \phi, \theta)$ is equivalent to minimizing $L_{\text{GML}}(o, \phi) = -\ln(f(i_1, i_2, o; \phi, \theta))$ which, in the discrete Fourier domain, is given by:

$$L_{\text{GML}}(o, \phi) \propto \sum_v \left[\frac{|\tilde{h}_1(v)\tilde{o}(v) - \tilde{i}_1(v)|^2}{\sigma^2} + \frac{|\tilde{h}_2(v)\tilde{o}(v) - \tilde{i}_2(v)|^2}{\sigma^2} + \frac{|\tilde{o}(v) - \tilde{o}_m(v)|^2}{S_o(v)} \right] \quad (9)$$

where v denotes the spatial frequency. The derivative of the criterion L_{GML} with respect to the object gives⁶ a closed-form expression for the object $\hat{o}(\phi)$ that minimizes the criterion for a given ϕ . Substituting \hat{o} into the criterion yields a that does not depend explicitly on the object, $L'_{\text{GML}}(\phi) = L_{\text{GML}}(\hat{o}(\phi), \phi)$. This new criterion reduces the

dimension of the parameter space to explore. On the other hand, L'_{GML} and L_{GML} are not convex and have the same number of critical points. Indeed, it can be shown that $\frac{\partial L'}{\partial \phi} = \frac{\partial L}{\partial \phi} \Big|_{\sigma=\hat{\sigma}(\phi)}$ so ϕ is a critical point of L'_{GML} if and only if $(\hat{\sigma}(\phi), \phi)$ is a critical point of L_{GML} and on the other hand, every critical point of L_{GML} takes the form $(\hat{\sigma}(\phi), \phi)$. The *a priori* information required on the object consists of the power spectral density of the object. We choose the following model

$$\begin{aligned} S_o(v) &\triangleq E[|\hat{\sigma}(v) - \bar{\sigma}_m(v)|^2] \\ &= k/[v_o^p + v^p] - |\bar{\sigma}_m(v)|^2 \end{aligned} \quad (10)$$

where E stands for the mathematical expectation. This heuristic model and similar ones have been quite widely used⁷. Notice that this model requires the choice of three hyperparameters k , v_o and p .

3.2. Properties

One acknowledged problem with GML estimation is that it does not ensure that bias and variance asymptotically vanish, i.e. the estimate may not converge towards the true value as the data size tends to infinity^{8,9}. Figure 2, obtained with simulations, illustrates this asymptotic behavior. (Notice that we use a gradient-conjugate method to minimize the criterion and that these curves, like all the following ones, are the average on 20 realizations of noisy monochromatic image pairs). Before minimizing the criterion, we have to set the value of the hyperparameters: σ^2 is equal to the number of photons/pixels and we fit the power spectral density model with the true object to obtain the hyperparameters (k , v_o and p), which will call the "true" hyperparameters. When the number of data increases, the root mean square error (RMSE) of the aberrations estimates does not decrease. An intuitive explanation of this phenomenon is that, if a bigger image is used to estimate the aberrations, the size of the object, which is jointly reconstructed, increases too. Another problem with the GML criterion is that it may present local minima. Yet another problem is that this estimator calls for the estimation of the hyperparameters $\theta = (\sigma, o_m, k, v_o, p)$ and that θ can not be jointly estimated with o and ϕ . Indeed, the criterion degenerates when one seeks θ together with o and ϕ ; in particular it is possible to find a pair (θ_e, o_e) that does not depend on the data such that the criterion tends to minus infinity. So before minimizing L_{GML} , these hyperparameters must be chosen by the user.

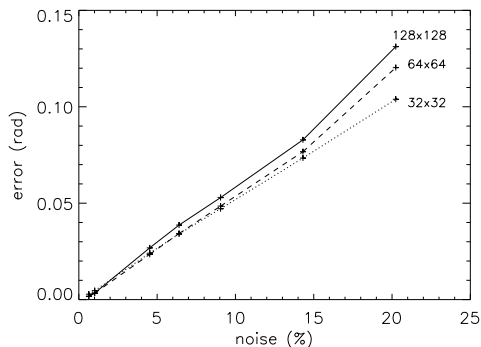


Figure 2. RMSE (root mean square error) of GML aberrations estimates as a function of noise level for the true hyperparameters. The solid line indicates an image size of 128×128 pixels; the dashed line indicates an image size of 64×64 pixels; and the dotted line indicates an image size of 32×32 pixels.

3.3. Influence of the hyperparameters

We shall now study the influence of the global hyperparameter on aberrations estimates and object estimate. This global hyperparameter quantifies the trade-off between goodness of fit to the data and the *a priori*. In our problem, this hyperparameter is $1/k$ as seen from equations 9, 10. We varied $1/k$ around its true value (corresponding, in figure, to 1). Figure 3 shows the RMSE on the aberrations estimates and on object estimate as a function of the

value of this hyperparameter. We see that the true value of this hyperparameter is not the same for the object and for the aberrations. For our problem (the estimation of the aberrations), the reconstruction is better when the object is under-regularized. The behavior of the curve corresponding to the aberrations depends on the noise level: for a high noise level (14% here), there is one optimal hyperparameter but for lower noise levels (4%), any value under 1, including a null regularization, leads to an accurate estimation of the aberrations even though the jointly estimated object is totally noisy. This observation sheds some light on the fact that unregularized phase estimation has been successfully used in the literature^{10,11}. This observation also leads us to study the asymptotic behavior of the GML with no regularization. Figure 4 shows the results. In this case, when the number of data increases, the RMSE on the aberrations estimates decreases. The number of data samples over the number of unknowns is the same as for the estimation with the true hyperparameters but here, the estimator behaves, surprisingly, as if the object were not being estimated.

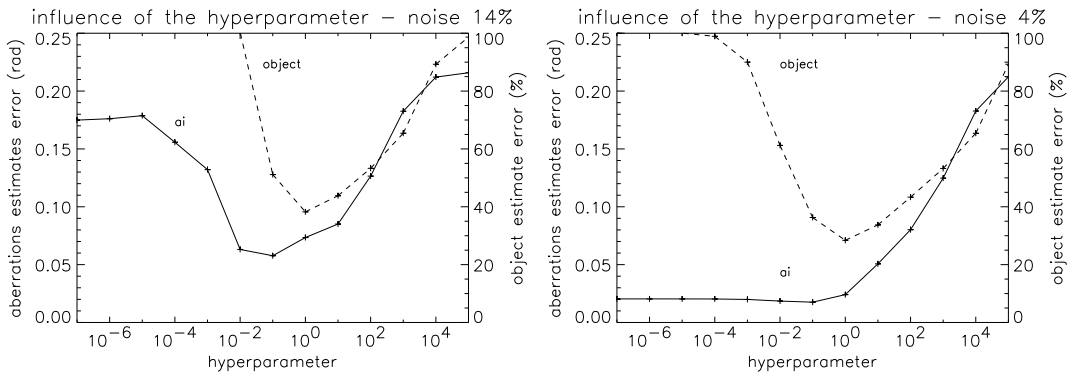


Figure 3. Plots of RMSE for aberrations estimates (solid line) and object estimate (dashed line) as a function of the value of the hyperparameter $1/k$ for an image size of 32×32 pixels. Left figure is for a noise level of 14%, right figure is for 4%.

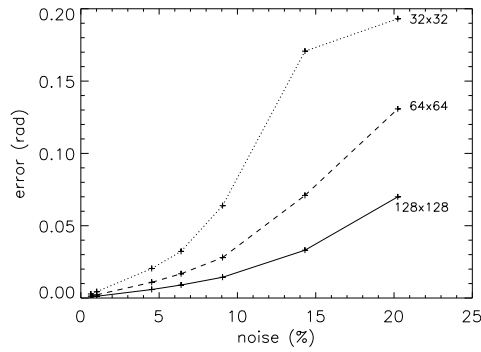


Figure 4. RMSE of GML aberrations estimates as a function of noise level for a null regularization. The solid line indicates an image size of 128×128 pixels; the dashed line indicates an image size of 64×64 pixels; and the dotted line indicates an image size of 32×32 pixels.

4. MARGINAL ESTIMATION

4.1. Criterion

The observed object is a nuisance parameter in our problem. We propose to estimate the sole phase or the phase and the hyperparameters by integrating the object out of the problem. This novel estimator is obtained by integrating the probability density function used for the joint estimation and it is based on the Maximum Likelihood estimator for ϕ , which is defined by

$$\begin{aligned}\hat{\phi}_{\text{ML}} &= \arg \max_{\phi} f(i_1, i_2; \phi, \theta) \\ &= \arg \max_{\phi} \int f(i_1, i_2, o; \phi, \theta) do\end{aligned}\quad (11)$$

We denote $I = (i_1 \ i_2)^T$ the vector which concatenates the data, so maximizing $f(i_1, i_2; \phi, \theta) = f(I; \phi, \theta)$ is equivalent to minimizing the following criterion :

$$L_{\text{ML}}(\phi) \propto \ln |R_I| + (I - m_I)^T R_I^{-1} (I - m_I) \quad (12)$$

where $m_I = (H_1 o_m \ H_2 o_m)^T$ is the vector which contains the mean images, R_I is the covariance matrix of I and $|R_I|$ its determinant. R_I has the following expression ($\mathbf{1}$ is the identity matrix):

$$R_I = \begin{bmatrix} H_1 R_o H_1^t + \sigma^2 \mathbf{1} & H_1 R_o H_2^t \\ H_2 R_o H_1^t & H_2 R_o H_2^t + \sigma^2 \mathbf{1} \end{bmatrix} \quad (13)$$

4.2. Relationship between ML and GML

It can be shown¹² that the two criterions L_{ML} and L'_{GML} are related by the following relationship:

$$L_{\text{ML}} \propto \ln |R_I| + L'_{\text{GML}} \quad (14)$$

Thus, the difference consists of a single additional term between the marginal estimator and the joint estimator, which in the Fourier domain is given by:

$$\ln |R_I| = \sum_f \ln S_o(v) + N^2 \ln \sigma^2 + \sum_v \ln \left(|\tilde{h}_1(v)|^2 + |\tilde{h}_2(v)|^2 + \frac{N^2 \sigma^2}{S_o(v)} \right) \quad (15)$$

where N^2 is the number of pixels in the image. Although the two estimators differ only by one term, we shall see that their properties differ considerably. Before that, let us see how the hyperparameters can be estimated.

4.3. Estimation of the hyperparameters

We have implemented the ML estimator and noticed that the quality of reconstruction depends strongly on a good tuning of the hyperparameters. This is not a problem because they can be estimated jointly with the aberrations. Indeed, they are given by

$$(\hat{\phi}, \hat{\theta})_{\text{ML}} = \arg \max_{\phi, \theta} f(i_1, i_2; \phi, \theta) \quad (16)$$

In our problem, there are four hyperparameters σ^2 , k , v_o and p . By doing the following change of variable $\mu = \sigma^2/k$, we can find a closed-form expression for \hat{k} such that $\frac{\partial L_{\text{ML}}}{\partial k} \Big|_{k=\hat{k}} = 0$. The criterion becomes $L_{\text{ML}}(\phi, \mu, v_o, p)$. The performance of this estimator is shown in Figure 5. We compare the quality of the aberrations reconstruction obtained by minimizing L_{ML} with the true hyperparameters and by minimizing $L_{\text{ML}}(\phi, \mu, v_o, p)$ which we call unsupervised estimation, for two different image sizes. From the plots, we see that for low noise levels, the agreement is very good. For 128×128 pixels, it is quite good for any noise level. For 32×32 pixels and high noise level, the reconstruction is seriously degraded because of the lack of information contained in the noisy data.

4.4. Properties

Under not very restrictive conditions, the ML estimator has a null asymptotic bias and variance^{13,14}. We show this property with simulations in figure 6. When the number of data increases the RMSE of estimated aberrations decreases. Furthermore, according to a Central Limit Theorem, when the number of data increases (i.e. the image size), the shape of the likelihood function tends to become more and more Gaussian so that it presents less and less local minima. We will now compare, with simulations, the performance of the GML and ML estimators for the reconstruction of the aberrations.

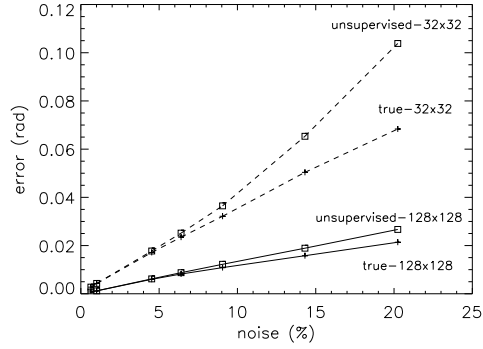


Figure 5. Performance of ML estimation with the true hyperparameters (pluses +) and for an unsupervised estimation (square \square) as measured by RMSE of aberrations estimates vs. noise level. The solid line indicates an image size of 128×128 pixels; and the dashed line indicates an image size of 32×32 pixels

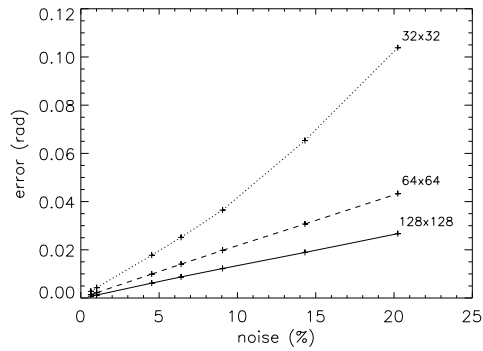


Figure 6. RMSE of ML aberrations estimates as a function of noise level for an unsupervised estimation. The solid line indicates an image size of 128×128 pixels; the dashed line indicates an image size of 64×64 pixels; and the dotted line indicates an image size of 32×32 pixels.

5. COMPARISON

5.1. Estimation with the true hyperparameters

First, we compare the performance of these two estimators with the true hyperparameters. Figure 7 demonstrates that the reconstruction of the aberrated parameters, with the true hyperparameters, is much better with the marginal estimator. The difference is all the more important as the image size is bigger. This result is due to the different asymptotic behavior of these two estimators.

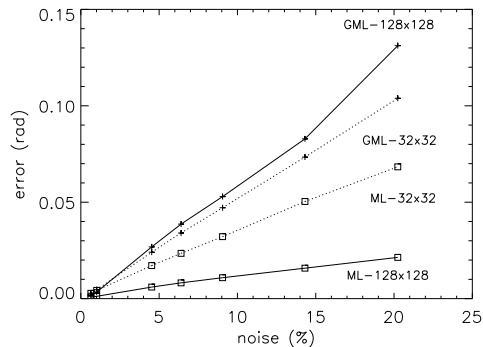


Figure 7. RMSE of ML (square \square) and GML (pluses $+$) with the true hyperparameters, aberrations estimates as a function of noise level. The solid line indicates an image size of 128×128 pixels; and the dashed line indicates an image size of 32×32 pixels

5.2. Unsupervised estimation

Here we present results of the comparison between the joint estimator with a null regularization and the unsupervised marginal estimator (i.e. hyperparameters are reconstructed jointly with the aberrations). We compared the root mean square error of the aberration 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 8. We can see at least two different domains: when the noise level $n < 5\%$, the two estimators give approximately the same results. For $5\% < n < 20\%$, the ML estimation is better. We have noticed that the results, for very high noise level $n > 15\%$ and an image size of 32×32 pixels, depends on the value of the aberrations coefficients i.e. the joint estimator could give better results than the marginal estimator. Anyway, in practice, this noise level is never reached. In terms of computational time, these two methods are quite equivalent.

6. CONCLUSION

We have presented two estimators. The first one is based on the Generalized Maximum Likelihood. It is the usual estimator in the field of phase diversity. We have shown its properties with simulations and pointed out its surprisingly good behavior for null regularization. We have also proposed a novel marginal estimator based on Maximum Likelihood, which estimates the sole aberration parameters and allows unsupervised estimation of the hyperparameters. We have shown, with simulations, its good asymptotic behavior. Finally, we have compared the performance of these estimators, for the reconstruction of the aberrations, for different image sizes and different noise levels. It appears that the marginal estimator gives equal or better reconstructions than the joint estimator while presenting a better coherence.

Acknowledgments

The authors wish to thank Vincent Michau for fruitful discussions.

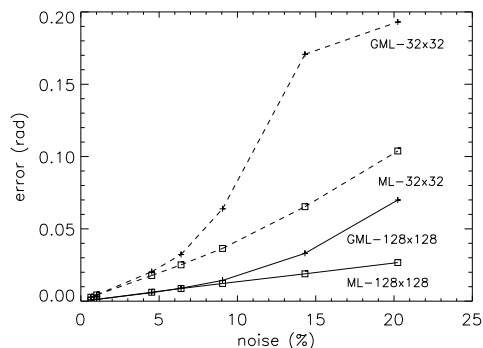


Figure 8. RMSE of unsupervised ML (square \square) and GML with null regularization (pluses $+$) aberrations estimates as a function of noise level. The solid line indicates an image size of 128×128 pixels; and the dashed line indicates an image size of 32×32 pixels

REFERENCES

1. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Optical Engineering* **21**(5), pp. 829–832, 1982.
2. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," *Applied Optics* **33**(27), pp. 6533–6546, 1994.
3. 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.*, pp. 9186–9197, Dec. 1997.
4. M. G. Lofdahl and A. L. Duncan, "Fast phase diversity wavefront sensor for mirror control," in *Adaptive Optical System Technologies*, D. Bonaccini and R. K. Tyson, eds., vol. 3353, pp. 952–963, SPIE, Mar. 1998.
5. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**(3), pp. 207–211, 1976.
6. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," *Journal of the Optical Society of America A* **9**(7), pp. 1072–1085, 1992.
7. A. P. Kattinig and J. Primot, "Model of the second-order statistic of the radiance field of natural scenes, adapted to system conceiving," in *Aerosense'97 - Visual information processing VI*, vol. 3074, SPIE, Soc. Photo-Opt. Instrum. Eng., (Washington), 1997.
8. F. Champagnat and J. Idier, "An alternative to standard maximum likelihood for gaussian mixtures," in *ICASSP*, pp. 2020–2023, (Detroit, U.S.A.), May 1995.
9. R. J. Little and D. B. Rubin, "On jointly estimating parameters and missing data by maximizing the complete-data likelihood," *The American Statistician* **37**, pp. 218–220, August 1983.
10. R. G. Paxman and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," *Journal of the Optical Society of America A* **5**(6), pp. 914–923, 1988.
11. 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**, pp. 4967–4979, Aug. 1999.
12. Y. Goussard, G. Demoment, and J. Idier, "A new algorithm for iterative deconvolution of sparse spike," in *ICASSP*, pp. 1547–1550, (Albuquerque, U.S.A.), avril 1990.
13. E. Lehmann, *Theory of point estimation*, Wiley, 1983.
14. E. D. Carvalho and D. Slock, "Maximum-likelihood blind FIR multi-channel estimation with gaussian prior for the symbols," in *ICASSP*, pp. 3593–3596, (Munich), Apr. 1997.

Bibliographie

- ACTON, D., SOLTAU, D., et SCHMIDT, W. (1996). « Full-field wavefront measurements with phase diversity ». *Astron. Astrophys.*, 309:661–672.
- BABA, N. et MUTOH, K. (2001). « Measurement of telescope aberrations through atmospheric turbulence by use of phase diversity ». *Appl. Opt.*, 40(4):544–552.
- BARON, F., CASSAING, F., BLANC, A., et LAUBIER, D. (2002). « Cophasing a wide field multi-aperture array by phase-diversity: influence of aperture redundancy and dilution. ». volume 4852. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- BIRAUD, Y. (1969). « A New Approach for Increasing the Resolving Power by Data Processing ». *Astron. Astrophys.*, 1(1):124–127.
- BLANC, A., IDIER, J., et MUGNIER, L. M. (2000a). « Novel estimator for the aberrations of a space telescope by phase diversity ». Dans BRECKINRIDGE, J. B. et JAKOBSEN, P., éditeurs, *UV, Optical, and IR Space Telescopes and Instruments*, volume 4013, pages 728–736, Bellingham, Washington. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- BLANC, A., MUGNIER, L. M., et IDIER, J. (1999). « Estimation des aberrations d'un telescope optique par diversité de phase ». Dans *17ième Colloque sur le Traitement du Signal et des Images*, volume 1, pages 63–66. GRETSI.
- BLANC, A., MUGNIER, L. M., et IDIER, J. (2000b). « Marginal estimator for the aberrations of a space telescope by phase diversity ». Dans *4th International Conference On Space Optics*, pages 77–86, Toulouse, France. CNES, CNES.
- BORN, M. et WOLF, E. (1983). *Principles of Optics*. Pergamon Press.
- BUCCI, O. M., CAPOZZOLI, A., et D'ELIA, G. (1999). « Regularizing strategy for image restoration and wave-front sensing by phase diversity ». *J. Opt. Soc. Am. A*, 16(7):1759–1768.
- CARRARA, D. A., THELEN, B. J., et PAXMAN, R. G. (2000). « Aberration correction of segmented-aperture telescopes by using phase diversity ».
- CARVALHO, E. D. et SLOCK, D. (1997). « Maximum-likelihood blind FIR multi-channel estimation with Gaussian prior for the symbols ». Dans *ICASSP*, pages 3593–3596, Munich.
- CHAMPAGNAT, F. et IDIER, J. (1995). « An alternative to standard maximum likelihood for Gaussian mixtures ». Dans *ICASSP*, pages 2020–2023, Detroit, U.S.A.
- COMMENGES, D. (1984). « The deconvolution problem: fast algorithms including the pre-conditioned conjugate-gradient to compute a MAP estimator ». *IEEE Transactions on automatic control*, 29:229–243.
- CONAN, J.-M., MUGNIER, L. M., FUSCO, T., MICHAU, V., et ROUSSET, G. (1998). « Myopic Deconvolution of Adaptive Optics Images using Object and Point Spread Function Power Spectra ». *Appl. Opt.*, 37(21):4614–4622.

- DEMOMENT, G. (1989). « Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems ». *IEEE Trans. Acoust. Speech Signal Process.*, 37(12):2024–2036.
- ELLERBROEK, B. L., THELEN, B. J., LEE, D. J., et PAXMAN, R. G. (1997). « Comparison of Shack-Hartmann wavefront sensing and phase-diverse phase retrieval ». Dans TYSON, R. K. et FUGATE, R. Q., éditeurs, *Adaptive Optics and Applications*, volume 3126, pages 307–320, Washington. SPIE, Soc. Photo-Opt. Instrum. Eng.
- FIENUP, J., THELEN, B., PAXMAN, R., et CARRARA, D. (1998). « Comparison of phase diversity and curvature wavefront sensing ». Dans BONACCINI, D. et TYSON, R., éditeurs, *Adaptive Optical System Technologies*, volume 3353, pages 930–940, Bellingham, Washington. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- FONTANELLA, J.-C. (1985). « Analyse de surface d'onde, déconvolution et optique active ». *J. of Optics (Paris)*, 16(6):257–268.
- FRIED, D. (1966). « Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures ». *J. Opt. Soc. Am.*, 56:1372–1379.
- FUSCO, T., BLANC, A., MUGNIER, L., et ROUSSET, G. (2002). « Development of a Phase Diversity approach to calibrate NAOS and CONICA static aberrations ». Rapport Technique de Synthèse RTS 1/06394 DOTA, ONERA. Contrat ESO.
- GANTMACHER, F. (1966). « *Théorie des Matrices, Tome I* », Chapitre L'algorithme de Gauss et quelques-unes de ses applications, pages 42–50. Dunod.
- GHIGLIA, D. et ROMERO, L. (1994). « Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods ». *J. Opt. Soc. Am.*, 11:107–117.
- GHIGLIA, D. C. et PRITT, M. D. (1998). *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley-Interscience.
- GONSALVES, R. A. (1976). « Phase retrieval from modulus data ». *Journal of the Optical Society of America*, 66(9):961–964.
- GONSALVES, R. A. (1982). « Phase retrieval and diversity in adaptive optics ». *Optical Engineering*, 21(5):829–832.
- GOODMAN, J. W. (1968). *Introduction to Fourier optics*. McGraw-Hill.
- GOUSSARD, Y., DEMOMENT, G., et IDIER, J. (1990). « A new algorithm for iterative deconvolution of sparse spike ». Dans *ICASSP*, pages 1547–1550, Albuquerque, U.S.A.
- JEFFERIES, S. M., LLOYD-HART, M., KEITH HEGE, E., et GEORGES, J. (2002). « Sensing Wave-Front Amplitude and Phase with Phase Diversity ». *Appl. Opt.*, 41(11):2095–2102.
- KATTNIG, A. P. et PRIMOT, J. (1997). « Model of the second-order statistic of the radiance field of natural scenes, adapted to system conceiving ». Dans *Aeorosense'97 - Visual information processing VI*, volume 3074, Washington. SPIE, Soc. Photo-Opt. Instrum. Eng.
- KENDRICK, R. L., ACTON, D. S., et DUNCAN, A. L. (1994). « Phase-diversity wave-front sensor for imaging systems ». *Appl. Opt.*, 33(27):6533–6546.
- KOLMOGOROV, A. N. (1941). « Local structure of turbulence in incompressible fluids with very high Reynolds number ». *Dokl. Akad. Nauk. SSSR*, 30(4):301–305.
- KUPKE, R., RODDIER, F., et MICKEY, D. (1998). « Wavefront Curvature sensing on extended arbitrary scenes: simulation results ». Dans *Proc. Soc. Photo-Opt. Instrum. Eng.*, volume 3353, pages 918–929.

- LEE, D., WELSH, B., et ROGGEMANN, M. (1997a). « Diagnosing unknown aberrations in an adaptive optics system by use of phase diversity ». *Opt. Lett.*, 22(13):952–954.
- LEE, D. J., ROGGEMANN, M. C., et WELSH, B. M. (1999). « Cramer-Rao analysis of phase-diverse wave-front sensing ». *J. Opt. Soc. Am. A*, 16(5):1005–1015.
- LEE, D. J., ROGGEMANN, M. C., WELSH, B. M., et CROSBY, E. R. (1997b). « Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing ». *Appl. Opt.*, pages 9186–9197.
- LEHMANN, E. (1983). *Theory of point estimation*. Wiley.
- LÖFDAHL, M. G. et DUNCAN, A. L. (1998). « Fast phase diversity wavefront sensor for mirror control ». Dans BONACCINI, D. et TYSON, R. K., éditeurs, *Adaptive Optical System Technologies*, volume 3353, pages 952–963. SPIE.
- LÖFDAHL, M. G., KENDRICK, R. L., HARWIT, A., MITCHELL, K. E., et DUNCAN, A. L. (1998). « A phase diversity experiment to measure piston misalignment on the segmented primary mirror of the Keck II Telescope ». Dans *Proc. Soc. Photo-Opt. Instrum. Eng.*, volume 3356, pages 1190–1201.
- LÖFDAHL, M. G. et SCHARMER, G. B. (1994). « Wavefront sensing and image restoration from focused and defocused solar images ». *Astron. Astrophys.*, 107:243–264.
- LÖFDAHL, M. G., SCHARMER, G. B., et WEI, W. (2000). « Calibration of a deformable mirror and Strehl ratio measurements by use of phase diversity ». *Appl. Opt.*, 39(1):94–103.
- LITTLE, R. J. A. et RUBIN, D. B. (1983). « On Jointly Estimating Parameters and Missing Data by Maximizing the Complete-Data Likelihood ». *The American Statistician*, 37(3):218–220.
- MARÉCHAL, A. et FRANÇON, M. (1970). *Diffraction, Structures des images*. Masson & Cie.
- MEYNADIER, L. (1997). « *Analyse de surface d'onde pour le contrôle actif d'un télescope spatial* ». PhD thesis, Université de Nice-Sophia Antipolis.
- MEYNADIER, L., MICHAU, V., VELLUET, M.-T., CONAN, J.-M., MUGNIER, L. M., et ROUSSET, G. (1999). « Noise propagation in wave-front sensing with phase diversity ». *Appl. Opt.*, 38(23):4967–4979.
- MICHAU, V., ROUSSET, G., et FONTANELLA, J.-C. (1992). « Wavefront sensing from extended sources ». Dans *Real Time and Post Facto Solar Image Correction*, pages 124–128, Sunspot, New Mexico, United States of America. NSO.
- MUGNIER, L., VELLUET, M.-T., BLANC, A., et MICHAU, V. (2001a). « Analyseur de front d'onde pour le système Pléiades – Tranche optionnelle ». Rapport Technique de Synthèse 105331 DOTA, ONERA. Contrat CNES.
- MUGNIER, L., VELLUET, M.-T., ROUSSET-ROUVIÈRE, L., BLANC, A., et MICHAU, V. (2001b). « Analyseur de front d'onde pour le système Pléiades ». Rapport Technique de Synthèse 105330 DOTA, ONERA. Contrat CNES.
- NAKAMURA, O., KAWATA, S., et MINAMI, S. (1988). « Optical microscope tomography. II. Nonnegative constraint by a gradient-projection method ». *J. Opt. Soc. Am. A*, 5(4):554–561.
- NOLL, R. J. (1976). « Zernike polynomials and atmospheric turbulence ». *J. Opt. Soc. Am.*, 66(3):207–211.
- OBUKHOV, A. M. (1949). « Structure of the temperature field in a turbulent current ». *Izv. Aka. Nauk SSSR, Ser. Geograf. Geofiz.*, 13(1):58–69.

- PAXMAN, R. G. et CRIPPEN, S. L. (1990). « Aberration correction for phased-array telescopes using phase diversity ». Dans *Digital Image Synthesis and Inverse Optics*, volume 1351, pages 787–797, Bellingham, Washington. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- PAXMAN, R. G., SCHULZ, T. J., et FIENUP, J. R. (1992). « Joint estimation of object and aberrations by using phase diversity ». *Journal of the Optical Society of America A*, 9(7):1072–1085.
- PAXMAN, R. G., THELEN, B. J., et SELDIN, J. H. (1994). « Phase-diversity correction of turbulence-induced space-variant blur ». *Opt. Lett.*, 19(16):1231–1233.
- RODDIER, C. et RODDIER, F. (1993). « Wave-front reconstruction from defocused images and testing of ground based optical telescope ». *J. Opt. Soc. Am. A*, 10(5):957–965.
- RODDIER, F. (1981). The effects of atmospheric turbulence in optical astronomy. Dans WOLF, E., éditeur, *Progress in Optics*, volume XIX, pages 281–376. North Holland, Amsterdam.
- RODDIER, F. (1988). « Passive versus active methods in optical interferometry ». Dans MERKLE, F., éditeur, *High-Resolution Imaging by Interferometry*, numéro 29 dans ESO Conference and Workshop Proceedings, pages 565–574, Garching Germany. NOAO-ESO, ESO.
- RODDIER, F., éditeur (1999). *Adaptive Optics in Astronomy*. Cambridge University Press, Cambridge.
- ROUSSET, G. (1999). Wave-front sensors. Dans (Roddier, 1999), Chapitre 5, pages 91–130.
- SELDIN, J. et PAXMAN, R. (2000). « Closed-loop wavefront sensing for a sparse-aperture, phased-array telescope using broadband phase diversity ». Dans *Imaging Technology and Telescopes*, volume 4091-07. SPIE.
- SELDIN, J. H. et PAXMAN, R. G. (1994). « Phase-diverse speckle reconstruction of solar data ». Dans *Image Reconstruction and Restoration*, volume 2302. SPIE.
- SHACK, R. B. et PLATT, B. C. (1971). « Production and use of a lenticular Hartmann Screen ». *J. Opt. Soc. Am.*, 61:656.
- TATARSKI, V. (1961). *Wave Propagation In a Turbulent Medium*. Dover Publications, Inc. New York.
- THELEN, B., CARRARA, D., et PAXMAN, R. (2000). « Pre- and post-detection correction of turbulence-induced space-variant blur ». Dans *Propagation and imaging through the atmosphere IV*. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- THELEN, B. J., PAXMAN, R. G., CARRARA, D. A., et SELDIN, J. H. (1999). « 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.
- THIÉBAUT, E. et CONAN, J.-M. (1995). « Strict *a priori* constraints for maximum-likelihood blind deconvolution ». *J. Opt. Soc. Am. A*, 12(3):485–492.
- TIKHONOV, A. et ARSENIN, V. (1977). *Solutions of Ill-Posed Problems*. Winston, DC.
- TYLER, D., FORD, S., PAXMAN, R., ROGGEMANN, M., ROUNTREE, J., SCHULZ, T., SCHILZE, K., SELDIN, J., SHEPPARD, D., STRIBLING, B., VAN KAMPEN, W., et WELSH, B. (1998). « Comparison of image reconstruction algorithms using adaptive optics instrumentation ». volume 3353, pages 160–171. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.
- VAN KAMPEN, W. et PAXMAN, R. (1998). « Multi-frame blind deconvolution of infinite-extent Objects ». Dans *Propagation and imaging through the atmosphere II*, volume 3433, pages 296–307. Proc. Soc. Photo-Opt. Instrum. Eng., SPIE.

- VOGEL, C. R., CHAN, T., et PLEMMONS, R. (1998). « Fast algorithms for phase-diversity-based blind deconvolution ». Dans BONACCINI, D. et TYSON, R. K., éditeurs, *Adaptive Optical System Technologies*, volume 3353, pages 994–1005. SPIE.