



Etude d'un nouveau concept d'analyse de front d'onde en plan focal appliqué à l'optique adaptative en astronomie

C. Plantet

► **To cite this version:**

C. Plantet. Etude d'un nouveau concept d'analyse de front d'onde en plan focal appliqué à l'optique adaptative en astronomie. Optique / photonique. OBSERVATOIRE DE PARIS, 2014. Français. <tel-01151683>

HAL Id: tel-01151683

<https://hal.archives-ouvertes.fr/tel-01151683>

Submitted on 13 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OBSERVATOIRE DE PARIS

**ECOLE DOCTORALE
ASTRONOMIE ET ASTROPHYSIQUE D' ILE-DE-FRANCE**

**Doctorat
ASTRONOMIE ET ASTROPHYSIQUE**

AUTEUR : Cédric Plantet

**ETUDE D'UN NOUVEAU CONCEPT D'ANALYSE DE FRONT
D'ONDE EN PLAN FOCAL APPLIQUÉ À L'OPTIQUE ADAPTATIVE EN ASTRONOMIE**

*Thèse dirigée par Thierry Fusco,
Encadrée par Serge Meimon et Jean-Marc Conan.*

Soutenue publiquement le 29/09/2014

Jury :

Sylvie ROQUES	(Rapporteur)
Simone ESPOSITO	(Rapporteur)
Gérard ROUSSET	(Examineur)
Benoît NEICHEL	(Examineur)
Michel TALLON	(Examineur)
Serge MEIMON	(Co-directeur de thèse)
Thierry FUSCO	(Directeur de thèse)

ONERA

THE FRENCH AEROSPACE LAB



Remerciements

J'aimerais tout d'abord remercier ma famille, avec laquelle j'ai baigné dans une ambiance de curiosité scientifique, qui m'a finalement menée jusque-là. Merci à Jean-Marc et Serge, qui m'ont encadré efficacement pendant cette thèse, en me laissant suffisamment d'autonomie sans pour autant m'abandonner à mon triste sort ! Bien évidemment, je remercie aussi Thierry, mon directeur de thèse, avec qui les réunions, bien que rares, ont été fructueuses et agréables. Je suis content de pouvoir continuer à travailler avec toi pour mon premier post-doc. Merci à Gérard Rousset, qui m'a conseillé d'aller faire mon stage de fin d'études à l'Onera après nous avoir donné un cours excellent à SupOptique. Merci aussi d'avoir accepté de présider le jury de ma soutenance. Je remercie d'ailleurs chaleureusement les membres de ce jury, qui ont tous montré un intérêt certain pour ma thèse, ainsi que toutes les personnes qui ont sacrifié un lundi après-midi pour assister à ma soutenance (j'apprécie grandement) !

Bien sûr, pour qu'une thèse se passe bien, il faut aussi que le moral soit présent. Pour cela, j'ai beaucoup de monde à remercier. Merci aux permanents du groupe Haute Résolution Angulaire de l'Onera, qui sont tous très sympathiques, toujours prêts à aider, et ouverts aux discussions, qu'elles soient intelligentes ou non. Merci aux (ex-)thésards : Baptiste, Jessica, Silvère (Silvène pour les intimes), Faustine (je ne m'étendrai pas dessus !), Jonathan (pas thésard mais tout comme), Rémy, Gaetano, Juliette, Bertrand, Léo... J'ai apprécié chaque moment passé avec vous ! Merci aux stagiaires qui sont resté plus ou moins de temps, mais qui ont toujours contribué à une bonne ambiance. Merci aux gens de ma chère résidence étudiante des Rives, où je me trouvais encore en stage : Virginie, JB, Thibault, Jonathan, et bien sûr ceux qui étaient là avant, Romain, Sophie, Guillaume, Bruce, François... L'atmosphère de cette résidence était clairement ce qu'il y avait de mieux pour décompresser après une journée de travail. Merci aux personnes de l'association d'astronomes amateurs de l'université d'Orsay, ALCOR, avec qui j'ai passé des super moments (pas de liste, il y a trop de monde !). Je remercie quand même mon collègue guitariste préféré, Matthieu, avec qui c'est toujours un plaisir de jouer. En parlant de musique, j'ai eu l'honneur et le plaisir de jouer avec Lucile en duo guitare acoustique/chant, sous le nom Antares (cherchez « Antares acoustic » sur Youtube pour écouter la belle voix de Lucile et moi qui essaye de grattouiller derrière). Merci pour tout ce qu'on a réussi à faire, malgré le temps limité par la fin de nos thèses respectives ! Merci à l'équipe d'organisation d'Elbereth (conférence par et pour les doctorants en astronomie d'Ile-de-France) de 2012, dont j'ai fait partie. C'était une belle expérience (oui j'avoue, je n'ai servi à rien), et c'est grâce à ça que j'ai rencontré Lucile ! Merci aux autres amis que je n'ai pas cité plus haut, mais qui ont quand même une place dans mon coeur (oh la chance !) : Julien, Agnès, Clément, Quentin, Charlène, Lucien, Amir (on ne s'est pas beaucoup revu pendant la thèse mais quand même !)... Je sais que j'en oublie, et j'en suis vraiment désolé ! Enfin, je voudrais remercier ma prof de chant, Christine, et mon prof de guitare, Emmanuel, qui m'ont beaucoup apporté sur le plan musical.

Voilà, je crois que tout est dit. Pour les thésards qui liraient ceci (et avec un peu de chance la suite) : bon courage !

Table des matières

Introduction	17
1 Systèmes d'optique adaptative pour l'astronomie et analyse de front d'onde	21
1.1 Introduction	21
1.2 Observation depuis le sol en astronomie : formation des images	21
1.2.1 Relation objet - image	21
1.2.1.1 Bruit sur l'image	22
1.2.2 Télescope idéal	23
1.2.2.1 FEP et FTO d'un télescope idéal	23
1.2.2.2 Échantillonnage	24
1.2.3 Imagerie en présence de turbulence	26
1.2.3.1 Effets optiques de la turbulences atmosphérique	26
1.2.3.2 Courte pose et longue pose en présence d'aberrations	26
1.3 Optique adaptative	27
1.3.1 Principe	27
1.3.2 Qualité de correction	29
1.3.2.1 Variances résiduelles	29
1.3.2.2 Rapport de Strehl	29
1.3.2.3 Choix d'une base pour décrire la phase	30
1.3.2.4 Densité Spectrale de Puissance	33
1.3.3 Budget d'erreur	33
1.3.3.1 Anisoplanétisme	35
1.3.3.2 Erreur d'ajustement	37
1.3.3.3 Erreur de mesure	37
1.3.3.4 Erreur temporelle	38
1.3.3.5 Erreurs d'étalonnage de la boucle d'OA	38
1.3.3.6 Aberrations non communes	39
1.3.3.7 Erreurs exogènes	39
1.3.4 Optique adaptative assistée par laser	39
1.3.4.1 Couverture de ciel	40
1.3.4.2 Étoiles laser	40
1.3.4.3 Systèmes d'optique adaptative assistée par laser	42
1.4 Analyseurs de front d'onde et leur comparaison	46
1.4.1 Analyseurs classiques	46
1.4.1.1 Shack-Hartmann	46
1.4.1.2 Pyramide	50

1.4.1.3	Analyseur de courbure	51
1.4.1.4	Masque de Zernike	54
1.4.1.5	Diversité de phase	55
1.4.2	LIFT	56
1.4.2.1	Historique	56
1.4.2.2	Principe	56
1.4.3	Comparaison des analyseurs de front d'onde	58
1.4.4	Propagation de bruit dans un estimateur Maximum de Vraisemblance	59
1.4.5	Coefficients de propagation du bruit	60
1.4.6	Information de Fisher	60
1.4.7	Lien entre l'information de Fisher et l'estimateur de maximum de vraisemblance	61
1.4.8	Gain pleine pupille	61
1.4.8.1	Division simple du flux	63
1.4.8.2	Division en sous-pupilles	63
1.4.8.3	Application au Shack-Hartmann et à la pyramide	64
2	Optimisation de LIFT et validation en laboratoire	65
2.1	Introduction	65
2.2	LIFT : un analyseur plan focal	66
2.2.1	Formation d'images	66
2.2.2	Algorithme	67
2.3	Optimisation et validation en laboratoire	70
2.3.1	Guide de lecture	70
2.3.2	Article	70
2.3.3	Bilan et discussion	87
2.3.4	Étude de robustesse	88
2.3.4.1	Robustesse à une méconnaissance sur l'échantillonnage	88
2.3.4.2	Robustesse à une méconnaissance sur l'astigmatisme	89
2.4	Comparaison de LIFT avec d'autres analyseurs	89
2.4.1	Paramètres de simulation	90
2.4.1.1	LIFT et diversité de phase	90
2.4.1.2	Shack-Hartmann	90
2.4.1.3	Pyramide	90
2.4.1.4	Masque de Zernike	90
2.4.2	Coefficients de propagation du bruit	91
2.5	Conclusion	92
3	Vers une utilisation de LIFT dans une optique adaptative grand champ	95
3.1	Introduction	95
3.2	Validation sur ciel en boucle ouverte avec le télescope Gemini Sud	96
3.2.1	Présentation du système GeMS	97
3.2.2	Tests sur source de calibration	97
3.2.2.1	Validation de la linéarité	99
3.2.2.2	Répétabilité de l'estimation	99
3.2.3	Validation de la linéarité sur une étoile	100

3.2.3.1	Guide de lecture	100
3.2.3.2	Acte de conférence	101
3.2.3.3	Bilan et discussion	110
3.3	Vers une utilisation de LIFT en boucle fermée	110
3.3.1	Améliorer l'algorithme pour limiter les effets des hauts ordres	110
3.3.1.1	Évaluation des effets de repliement sur l'estimation du tip/tilt et du focus	110
3.3.1.2	Solutions envisagées et amélioration apportée	113
3.3.2	Sensibilité au bruit en présence de hauts ordres résiduels	117
3.3.2.1	Hypothèses et outils de travail	117
3.3.2.2	Impact des hauts ordres sur la performance du Shack-Hartmann 2x2	120
3.3.2.3	Impact des hauts ordres sur la performance de LIFT	121
3.3.3	Simulation d'une boucle fermée	123
3.3.3.1	Paramètres de simulation	123
3.3.3.2	Exemple d'accrochage	124
3.3.4	Discussion autour d'une variation de LIFT	126
3.4	Perspectives sur GeMS et utilisation de LIFT en optique adaptative	128
3.5	Conclusion	129

Mesurer des hauts ordres avec LIFT ? 131

4 Shack-Hartmann LIFTé : une adaptation de LIFT pour estimer les hauts ordres 133

4.1	Introduction	133
4.2	Principe du Shack-Hartmann LIFTé	134
4.3	Première analyse théorique	136
4.3.1	Guide de lecture	136
4.3.2	Article	137
4.3.3	Bilan et discussion	141
4.4	Propagation du bruit dans le Shack-Hartmann LIFTé	141
4.4.1	Propagation du bruit dans une sous-pupille	141
4.4.2	Propagation du bruit sur l'estimation globale	143
4.5	Comparaison de la propagation de bruit du Shack-Hartmann classique et du Shack-Hartmann LIFTé	144
4.5.1	Proposition de dimensionnement	144
4.5.2	Comparaison d'un Shack-Hartmann 20x20 et d'un Shack-Hartmann LIFTé 10x10	144
4.5.2.1	Modèles	145
4.5.2.2	Comparaison	145
4.5.2.3	Discussion	145
4.6	Cas d'application du Shack-Hartmann LIFTé	147
4.7	Conclusion	147

5 Une comparaison équitable entre analyseurs pour l'estimation de hauts ordres ? 149

5.1	Introduction	149
5.2	Analyse de l'état de l'art	150

5.3	Information de Fisher : application au Shack-Hartmann	152
5.3.1	Paramètres de simulation	152
5.3.2	Propagation du bruit dans le Shack-Hartmann	153
5.3.2.1	Pentes	153
5.3.2.2	Image complète	153
5.4	Comparaison du Shack-Hartmann, de la pyramide et du Shack-Hartmann LIFTé	154
5.4.1	Guide de lecture	154
5.4.2	Article	155
5.4.3	Bilan et discussion	173
5.5	Autres comparaisons : premiers résultats et perspectives	173
5.5.1	Paramètres de simulation	173
5.5.2	Propagation du bruit en lumière monochromatique	174
5.5.2.1	Pyramide	174
5.5.2.2	Diversité de phase	174
5.5.2.3	Analyseur de courbure	178
5.5.2.4	Masque de Zernike	179
5.5.2.5	Bilan comparatif	181
5.5.3	Perspectives pour une comparaison en lumière polychromatique	183
5.6	Conclusion	184
	Conclusion et perspectives	185
	A Démonstrations mathématiques des formules utilisées dans LIFT	189
A.1	Expression de $I'_k(0)$	189
A.2	Expression de P_{MV}	189
	B Fonctionnement de LIFT en pratique	191
	C Propagation du bruit au sein d'une sous-pupille du Shack-Hartmann LIFTé	193
	D Information de Fisher pour un modèle de bruit additif gaussien	195
D.1	Expression générale	195
D.2	Lien avec l'estimateur de maximum de vraisemblance	196
	E Calcul de l'information de Fisher en pratique	199
E.1	Analyseur plan focal à image unique : LIFT	199
E.2	Diversité de phase	200
E.3	Shack-Hartmann	201
E.4	Pyramide	202
E.5	Courbure	204
E.6	Filtre de Zernike	205

Table des figures

1.1	Liens entre les différents plans et grandeurs FEP et FTO	23
1.2	Tache d’Airy	24
1.3	Zoom sur la partie centrale de la tache d’Airy	25
1.4	Représentation 3D de la FTO d’un télescope parfait	25
1.5	Effet de la turbulence sur l’image	27
1.6	Schéma de principe d’une optique adaptative.	28
1.7	Exemple de résultats d’OA	28
1.8	Polynômes de Zernike	31
1.9	Principe de la chaîne d’acquisition incluant un système d’OA.	34
1.10	Représentation schématique du phénomène d’anisoplanétisme.	36
1.11	Spectre spatial d’un signal échantillonné et mise en évidence du phénomène de repliement.	38
1.12	Schéma représentatif d’une OA fonctionnant avec une étoile laser Sodium et une étoile laser Rayleigh.	41
1.13	Représentation schématique du problème d’effet de cône avec étoile laser.	42
1.14	Principe d’indétermination du tilt avec une étoile laser	43
1.15	Schéma de principe de la LTAO	44
1.16	Schéma de principe de la MCAO	45
1.17	Schéma de principe de fonctionnement d’un test de Hartmann.	47
1.18	Schéma de principe de fonctionnement d’un SH.	48
1.19	Schéma de principe de l’analyseur à pyramide	50
1.20	Schéma de principe de l’analyseur de courbure	52
1.21	Masque représentant la matrice de photodiodes à avalanches qui constitue le détecteur de l’analyseur de courbure	53
1.22	Levée de l’indétermination du focus grâce à l’astigmatisme	57
1.23	Cartes de pondération pour LIFT en fort flux et faible flux.	58
1.24	Front d’onde tilté arrivant sur une lentille.	62
2.1	Compensation de l’approximation faite par la linéarisation grâce aux itérations	68
2.2	Validation par simulation de l’augmentation du domaine de linéarité avec différents nombres d’itérations pour le tip et le focus	69
2.3	Coefficient de linéarité en fonction de l’échantillonnage entré comme paramètre dans le modèle de LIFT	88
2.4	Coefficient de linéarité en fonction de l’astigmatisme entré comme paramètre dans le modèle de LIFT	89
3.1	Tache présentant des résidus de hauts ordres turbulents.	96

3.2	Le banc Canopus.	98
3.3	Validation de la linéarité de l'estimation du focus sur les sources de calibration de Gemini	99
3.4	Estimation du focus sur source de calibration interne de GeMS avec une turbulence résiduelle simulée grâce aux miroirs déformables.	100
3.5	Superposition des estimation de focus pour trois cycles de turbulence simulée .	101
3.6	Coefficients de repliement r_{ik} sur l'estimation du tip/tilt et du focus en fonction du numéro de mode de Zernike	112
3.7	Coefficients de repliement en fonction du numéro de mode de Zernike et de la taille de la fenêtre de calcul	115
3.8	Coefficients de repliement en fonction du numéro de mode de Zernike avec une estimation MAP	116
3.9	Disposition des étoiles laser	118
3.10	Directions d'analyse par LIFT considérées pour le calcul des DSP	119
3.11	Variance d'erreur d'estimation sur tip/tilt/focus pour différentes directions d'observations et différents seeings en fonction du rapport de Strehl	123
3.12	Variance résiduelle du tip/tilt/focus avec et sans correction en fonction du temps	125
3.13	Erreur résiduelle sur le tip/tilt/focus en fonction du nombre d'itérations numériques	127
3.14	Perte sur le rapport de Strehl et l'énergie encadrée lorsqu'on ajoute un astigmatisme de 0.5 rad pendant 1% du temps de pose, en fonction du rapport de Strehl initial	129
4.1	Cartes de pondération de LIFT pour l'estimation du tip/tilt et de la coma en fort flux et faible flux.	143
4.2	Propagation du bruit dans un SH classique 20x20 par rapport à un SH LIFTÉ 10x10 estimant 9 hyperpentes par sous-pupille	146
5.1	Propagation du bruit dans le Shack-Hartmann lorsqu'on mesure des pentes avec un centre de gravité (COG) ou un centre de gravité pondéré (WCOG).	153
5.2	Propagation du bruit dans le Shack-Hartmann lorsqu'on utilise les données brutes	154
5.3	Propagation du bruit dans la pyramide en fonction de la modulation, en considérant les données brutes ou les pseudo-données	175
5.4	Propagation du bruit dans la diversité de phase, pour une image focale et une image défocalisée de 1,8 rad.	175
5.5	Propagation du bruit dans la diversité de phase, pour différentes amplitudes de diversité.	176
5.6	Propagation du bruit globale dans la diversité de phase en fonction de l'amplitude de défocalisation et du nombre de modes estimés	177
5.7	Propagation du bruit dans la diversité de phase, pour une image focale et une image défocalisée de 1,8 rad (configuration asymétrique), ou deux images défocalisées de $\pm 0,9$ rad (configuration symétrique).	177
5.8	Propagation du bruit dans la diversité de phase pour une répartition du flux image focale/image extra-focale de 50%/50% ou 30%/70%	178
5.9	Propagation du bruit dans l'analyseur de courbure	179
5.10	Propagation globale dans l'analyseur à courbure en fonction de la distance extra-focale l	180

5.11	Propagation du bruit dans le masque de Zernike	180
5.12	Comparaison des propagations de bruit dans divers analyseurs pour l'estimation de 1000 modes de Karhunen-Loève.	181
B.1	Représentation de l'image en entrée dans LIFT, ainsi que la zone utilisée pour l'estimation	192
E.1	Image LIFT et exemples de dérivées.	200
E.2	Images en plan focal et extra-focal et leurs dérivées	201
E.3	Image Shack-Hartmann et exemples de dérivées.	202
E.4	Image d'un analyseur à pyramide et exemples de dérivées	203
E.5	Images d'un analyseur de courbure et exemples de dérivées.	204
E.6	Image d'un analyseur à filtre de Zernike et exemples de dérivées.	206

Liste des tableaux

2.1	Coefficients de propagation du bruit mode à mode des différents analyseurs considérés pour l'estimation du tip/tilt et du focus	91
3.1	Paramètres du système d'optique adaptative considéré dans les simulations. . .	118
3.2	Profils de C_n^2 et de vent	119
3.3	Seeings considérés et performances associées, exprimées en rapport de Strehl (hors tip/tilt).	120
3.4	Variance d'erreur d'estimation du tip/tilt/focus par un Shack-Hartmann 2×2 pour les différents seeings considérés	121
3.5	Coefficients de propagation du bruit de photons et du bruit de détecteur du Shack-Hartmann 2×2 pour l'estimation du tip/tilt et du focus en fonction du rapport de Strehl	121
3.6	Variance d'erreur d'estimation du tip/tilt et focus par LIFT pour les différents seeings	122
3.7	Coefficients de propagation du bruit de photons et du bruit de détecteur dans LIFT pour l'estimation du tip/tilt et du focus en fonction du rapport de Strehl .	122
3.8	Paramètres de la simulation d'une boucle fermée.	125
5.1	Coefficients de propagation du bruit globale des différents analyseurs considérés pour l'estimation de 1000 modes de Karhunen-Loève	182

Liste des acronymes

Acronyme	Signification
ASO	Analyseur de Surface d'Onde
ATLAS	<i>Advanced Tomographic Laser-Assisted adaptive optics System</i>
CCD	<i>Charge-Coupled Device</i>
CdG(P)	Centre de Gravité (Pondéré)
CRLB	<i>Cramer-Rao Lower Bound</i>
COG	<i>Center Of Gravity</i>
DM	<i>Deformable Mirror</i>
DSP	Densité Spectrale de Puissance
EG	Etoile Guide
FEP	Fonction d'Étalement du Point
FTO	Fonction de Transfert Optique
GeMS	<i>Gemini Multiconjugate adaptive optics System</i>
GSAOI	<i>Gemini South Adaptive Optics Imager</i>
ILPD	<i>Iterative Linear Phase Diversity</i>
LBT	<i>Large Binocular Telescope</i>
LIFT	<i>Linearized Focal-plane Technique</i>
LTAO	<i>Laser Tomography Adaptive Optics</i>
MAP	Maximum A Posteriori
MCAO	<i>MultiConjugate Adaptive Optics</i>
MV	Maximum de Vraisemblance
NCPA	<i>Non-Common Path Aberrations</i>
OA	Optique Adaptative
ODGW	<i>On-Detector Guide Window</i>
PSF	<i>Point Spread Function</i> (FEP en français)
RTC	<i>Real-Time Computer</i>
SR	<i>Strehl Ratio</i>
TF	Transformée de Fourier
UV	UltraViolet
WCOG	<i>Weighted Center Of Gravity</i>

Introduction

Contexte

Afin de mieux comprendre les débuts de notre Univers, les astronomes ont besoin d'observer des astres très lointains¹, avec la meilleure **résolution**² possible. L'observation des astres dans les domaines visible et proche infrarouge de la lumière est rendue possible grâce à l'utilisation de télescopes, qui collectent et focalisent la lumière issue de l'objet observé à l'aide d'un miroir concave. Une image se forme alors au point de focalisation ("foyer" du télescope) et peut être acquise sur un détecteur. Dans des conditions idéales, la résolution de cette image est inversement proportionnelle au diamètre du miroir. Malheureusement, plus l'astre se trouve loin, moins on reçoit de lumière de sa part, et plus sa taille angulaire est faible. Dans l'optique d'observer toujours plus loin, on a alors besoin d'augmenter la taille des télescopes, pour gagner en surface collectrice d'une part, et en résolution d'autre part.

Cependant, la lumière issue de l'astre observé traverse l'atmosphère, qui est un milieu turbulent, avant de parvenir jusqu'au sol. Les turbulences atmosphériques provoquent des variations rapides de l'indice de réfraction, donc du chemin optique (produit de l'indice de réfraction et de la distance parcourue) suivi par les rayons lumineux. Les décalages induits par ce phénomène déforment le front d'onde de la lumière, qui est initialement plan à son entrée dans l'atmosphère. Il en résulte une baisse importante de la résolution des images obtenues sur un télescope au sol. Deux solutions peuvent être envisagées pour résoudre ce problème : envoyer le télescope dans l'espace ou compenser les effets de la turbulence atmosphérique. Ces deux solutions sont complémentaires, mais il est plus facile d'atteindre de très grands diamètres sur Terre que dans l'espace. Nous nous intéresserons donc plutôt à la deuxième.

La compensation des turbulences atmosphériques peut être effectuée grâce au concept d'**optique adaptative**. L'optique adaptative est un système qui permet de corriger les perturbations du front d'onde grâce à un miroir qui se déforme en temps réel. La forme à appliquer au miroir est déterminée à partir de la mesure d'un **analyseur de front d'onde**. Il est ainsi possible de retrouver une image à la résolution théorique du télescope. Bien évidemment, un tel système n'est pas parfait et sa performance est limitée par de nombreux aspects. En particulier, la précision sur l'analyse du front d'onde incident dépend du flux lumineux reçu par l'analyseur. Si ce flux est trop faible, le front d'onde est mal estimé et la correction est mauvaise. L'objet d'intérêt étant souvent peu lumineux, le front d'onde est généralement analysé sur une étoile proche, dite "**étoile guide**". Cependant, cette étoile ne doit pas être trop éloignée, car on risque de mesurer des déformations différentes de celles subies par le front d'onde issu de l'objet. Hélas, pour une grande partie du ciel, on ne trouve pas facilement d'étoile brillante à proximité de

1. La lumière ayant une vitesse finie, un objet lointain est vu tel qu'il était à une époque antérieure.

2. On entend ici par résolution la capacité à séparer deux sources ponctuelles voisines.

l'axe d'observation. Les zones de ciel "observables", c'est-à-dire pour lesquelles la correction est satisfaisante, sont donc rares. On dit que l'on a une mauvaise **couverture de ciel**. Cela est évidemment gênant pour les astronomes, qui peuvent justement avoir besoin d'observer des zones où il y a peu d'étoiles brillantes pour éviter les flux parasites.

Pour augmenter la couverture de ciel, on peut essayer d'utiliser un analyseur de front d'onde très peu sensible au bruit de mesure, c'est-à-dire qui peut se contenter de peu de photons. Il est alors possible de garder une bonne estimation du front d'onde en visant des étoiles guides très faibles. Il faut toutefois noter que le front d'onde doit être analysé à haute cadence pour avoir une correction en temps réel. Si aucun photon n'est reçu pendant le temps de pose du détecteur, une analyse est impossible quelle que soit la sensibilité au bruit. L'analyseur habituellement utilisé, le Shack-Hartmann, est bien connu et facile à mettre en oeuvre, mais il est relativement sensible au bruit. Parmi les autres analyseurs envisagés en optique adaptative, on peut notamment citer l'analyseur à pyramide, introduit par Ragazzoni en 1996 [Ragazzoni, 1996]. Cet analyseur est beaucoup moins sensible au bruit que le Shack-Hartmann, et peut donc analyser des étoiles plus faibles [Ragazzoni et Farinato, 1999, Vérinaud, 2004]. Afin d'orienter le choix d'un analyseur, de nombreuses études sur la sensibilité au bruit ont été effectuées. Un article de Guyon [Guyon, 2005] est devenu une référence sur ce sujet. La limite de ces études est qu'elles ne considèrent généralement que le bruit de mesure lié au signal reçu, le bruit de photons, alors qu'il existe un aussi un bruit lié au détecteur lui-même.

L'autre solution est, à défaut d'avoir une étoile brillante à proximité de l'objet d'intérêt, d'en créer une à l'aide d'un laser. C'est ce qu'on a proposé Foy et Labeyrie en 1985 [Foy et Labeyrie, 1985]. Les "**étoiles guides laser**" peuvent par exemple être créées en projetant un laser vers une couche de sodium située dans la haute atmosphère. Le laser, qui doit être à la longueur d'onde d'excitation du sodium, excite alors le volume traversé, qui ré-émet de la lumière en retour. On a ainsi une source brillante à disposition quelle que soit la direction d'observation. Cette solution devrait donc, en théorie, fournir une couverture de ciel de 100%. Cependant, elle présente certaines difficultés. En particulier, l'estimation de certaines déformations du front d'onde, à savoir le basculement (appelé **tip/tilt**) et, dans le cas des étoiles créées avec la couche de sodium, la défocalisation (**focus**) ne peut pas être faite sur les étoiles laser. Ces déformations doivent donc être estimées sur une étoile naturelle avec un analyseur dédié. La couverture de ciel des systèmes assistés par laser est alors limitée par la sensibilité au bruit de l'analyseur sur étoile naturelle. La plupart des solutions actuelles se base sur des analyseurs existants, qui n'ont pas été spécifiquement développés pour cette application. Récemment, Meimon *et al.* ont proposé un nouvel analyseur, appelé LIFT (LInearized Focal-plane Technique), dédié à l'analyse du tip/tilt et du focus sur des étoiles naturelles faibles [Meimon *et al.*, 2010b]. Cet analyseur permettra en théorie de gagner une magnitude par rapport aux solutions actuelles, c'est-à-dire qu'une observation sera possible avec des étoiles guides 2,5 fois moins brillantes.

Objectifs scientifiques

Nous avons pu constater que la couverture de ciel restait limitée, dans tous les cas de figure, par la sensibilité au bruit lors de l'analyse du front d'onde sur une étoile naturelle.

Dans cette thèse, nous étudions la sensibilité au bruit, en portant l'accent sur LIFT. Cet analyseur est prometteur pour l'augmentation de la couverture de ciel des systèmes d'optique adaptative assistée par laser. Cependant, comme tout concept, il est nécessaire de le tester ex-

périmentalement afin de s'assurer de son bon fonctionnement, et de le comparer aux autres solutions pour justifier son utilisation.

Le premier objectif de la thèse est donc d'optimiser LIFT, notamment en termes de sensibilité au bruit, pour qu'il soit le plus compétitif possible, et de le valider expérimentalement. Le second objectif est de comparer les analyseurs existants pour l'estimation sur étoile naturelle du tip/tilt et focus d'une part, et une estimation fine d'autre part. Cette étude permettra de motiver l'utilisation de LIFT et d'un analyseur dérivé, le Shack-Hartmann LIFTé, dans des systèmes d'optique adaptative.

Structure de la thèse

Le manuscrit est constitué de 5 chapitres.

Le chapitre 1 p.21 est consacré à la présentation des différentes notions évoquées dans la thèse. Nous expliquons la formation des images et leur dégradation par la turbulence atmosphérique. Le principe et les limitations de l'optique adaptative, ainsi que les systèmes assistés par laser, sont ensuite abordés. En fin de chapitre, nous décrivons les principaux analyseurs existants, puis nous détaillons le principe de LIFT, ainsi que la méthode utilisée pour évaluer la sensibilité au bruit.

Le chapitre 2 p.65 présente l'étape d'optimisation de LIFT et sa validation expérimentale en laboratoire. L'optimisation s'appuie notamment sur une étude de la sensibilité au bruit. Les tests expérimentaux ont pour but de vérifier la linéarité de l'estimation par LIFT, et de confirmer sa performance en fonction du niveau de bruit. LIFT est ensuite comparé aux autres analyseurs évoqués au chapitre 1 p.21 pour l'estimation du tip/tilt et du focus sur étoile naturelle.

Une fois LIFT validé en laboratoire, nous voulions nous assurer qu'il pouvait fonctionner dans un système opérationnel. C'est le sujet du chapitre 3 p.95, qui montre la première validation de LIFT sur ciel, effectuée grâce au système d'optique adaptative du télescope Gemini Sud, GeMS. Cette validation a été faite en boucle ouverte, c'est-à-dire sans utiliser les estimations de LIFT pour corriger le front d'onde. En vue de tests en boucle fermée, nous nous intéressons, dans la suite du chapitre, aux sources d'erreur d'estimation qui peuvent être rencontrées, et nous proposons une amélioration de LIFT pour limiter ces erreurs. Nous vérifions enfin la performance de LIFT en boucle fermée à l'aide de simulations.

Dans l'attente d'une validation sur ciel en boucle fermée, des questions sur les applications de LIFT ont été soulevées, en particulier sur la possibilité de faire une estimation fine du front d'onde. Si le front d'onde est fortement perturbé, la réponse est négative, mais les réflexions sur cette question ont donné l'idée d'un nouveau concept, le Shack-Hartmann LIFTé.

Le principe du Shack-Hartmann LIFTé est décrit dans le chapitre 4 p.133. Cet analyseur permet une estimation fine du front d'onde en combinant les avantages du Shack-Hartmann et ceux de LIFT. Nous montrons dans ce chapitre l'intérêt d'utiliser un Shack-Hartmann LIFTé plutôt qu'un Shack-Hartmann classique.

Avec l'apparition du Shack-Hartmann LIFTé, le choix d'un analyseur sur étoile naturelle pour un système d'optique adaptative devient très diversifié. Pour guider ce choix, nous proposons dans le chapitre 5 p.149 une comparaison en termes de sensibilité au bruit pour une estimation très fine du front d'onde. Cette étude complète les analyses déjà existantes et permet d'avoir une connaissance du comportement des analyseurs vis-à-vis du bruit de détecteur.

Chapitre 1

Systemes d'optique adaptative pour l'astronomie et analyse de front d'onde

1.1 Introduction

Dans ce chapitre, nous présentons les différentes notions utiles pour la suite de la lecture de la thèse. La première partie décrit la formation des images (§1.2 p.21). Nous y expliquons la relation entre l'objet observé et son image à travers un système optique, ainsi que les effets de la turbulence atmosphérique sur ces images.

Nous montrons ensuite comment ces effets sont compensés par l'optique adaptative (§1.3 p.27), puis nous présentons les systèmes d'optique adaptative qui vont nous intéresser dans cette thèse. Nous détaillons leur fonctionnement ainsi que leurs limites actuelles.

La dernière partie (§1.4 p.46) concerne les analyseurs de front d'onde, qui sont présents dans tout système d'optique adaptative. Nous nous intéressons aux principaux analyseurs qui peuvent être envisagés sur une optique adaptative, ainsi qu'aux moyens de les comparer.

Ce chapitre s'inspire largement du cours "Imagerie à Haute Résolution Angulaire et Applications, Optique Adaptative" de Cyril Petit et Thierry Fusco.

1.2 Observation depuis le sol en astronomie : formation des images

La turbulence atmosphérique induit une dégradation de l'image donnée par un système optique. Après un rappel général sur la formation des images, je décris les effets de la turbulence atmosphérique sur ces images.

1.2.1 Relation objet - image

Dans le cas général, l'image $I(\alpha)$ d'un objet $O(\alpha)$ s'écrit

$$I(\alpha) = \int \int FEP(\beta, \alpha) O(\beta) d\beta \quad (1.1)$$

où FEP désigne la Fonction d'Étalement du Point. La FEP ou "Point Spread Function" (PSF) en anglais caractérise la réponse du système optique considéré au niveau du plan focal d'ima-

gerie pour un point source (réponse du système à une impulsion ou pic de Dirac). α, β représentent des vecteurs positions (exprimés en radian) dans le plan focal. Dans la suite du document, les vecteurs seront représentés en gras, leur module (et plus généralement l'ensemble des variables scalaires) auront un format normal. Ainsi $|\alpha| = \alpha$.

La caractérisation complète du système suppose la connaissance de cette FEP **pour chaque vecteur angulaire** α dans le champ de vue considéré. Soit dans le cas d'une image de $N \times N$ pixels, N^2 FEPs à connaître. Dans certains cas, on peut supposer que cette FEP est invariante par translation, c'est-à-dire que la FEP ne dépend plus de sa position dans le champ, réduisant par là même la connaissance du système à la connaissance **d'une seule** FEP. Le domaine où la FEP est considérée comme invariante est appelé domaine isoplanétique. Sous cette hypothèse, l'équation 1.1 se simplifie en un simple produit de convolution :

$$I(\alpha) = \int \int O(\beta) FEP(\alpha - \beta) d\beta = (O \star FEP)[\alpha] \quad (1.2)$$

Dans le plan de Fourier, cette convolution s'écrit sous la forme d'un produit fréquence à fréquence :

$$\tilde{I}(\kappa) = \tilde{O}(\kappa) \times FTO(\kappa) \quad (1.3)$$

$\tilde{I}, FTO, \tilde{O}$ représentent respectivement les transformées de Fourier de I, FEP, O . Dans la suite la transformée de Fourier (TF) d'une fonction f sera notée \tilde{f} .

On fait ainsi apparaître la fonction de transfert optique FTO . La fonction de transfert optique est donc une représentation de la FEP dans un plan de Fourier conjugué du plan réel (Fig. 1.1). Elle permet de quantifier l'atténuation fréquentielle introduite par la FEP sur l'objet observé.

$$FTO(\kappa) = TF \{FEP(\alpha)\}_{\kappa} \quad (1.4)$$

Ici κ représente une fréquence spatiale (si α est donné en radians, κ est en inverse de radians, ou cycles par radian).

1.2.1.1 Bruit sur l'image

La mesure d'une grandeur physique telle que le nombre de photons, ou flux, reçu par un pixel s'accompagne toujours de ce qu'on appelle un bruit de mesure. Le bruit de mesure est une partie du signal qui s'ajoute au signal vrai, et que l'on ne souhaite pas mesurer. Le bruit est un phénomène aléatoire qu'on caractérise par la loi de probabilité qu'il suit, sa variance et sa moyenne. La variance, ou plutôt sa racine carrée, l'écart-type, donne une idée de l'amplitude de variation du bruit, tandis que la moyenne est la valeur autour de laquelle il varie.

Dans cette thèse, on considère deux sources de bruit sur les images : le bruit de photons et le bruit de détecteur.

Le bruit de photons est un bruit fondamental qui provient de la nature quantique de la lumière. Quelle que soit la source, les photons sont émis de manière aléatoire. La probabilité de l'arrivée d'un photon en un point donné en fonction du temps suit une loi de Poisson. Il s'agit d'un bruit multiplicatif, c'est-à-dire dépendant du signal vrai. La variance et la moyenne du bruit de photons sont toutes deux égales au flux.

Le bruit de détecteur, ou de lecture, est lié à la manière dont on détecte les photons. Les photons sont en effet transformés en électrons à l'aide de semi-conducteurs. On a alors un courant électrique proportionnel au flux reçu. Ce courant, pour être lu, est transféré le long

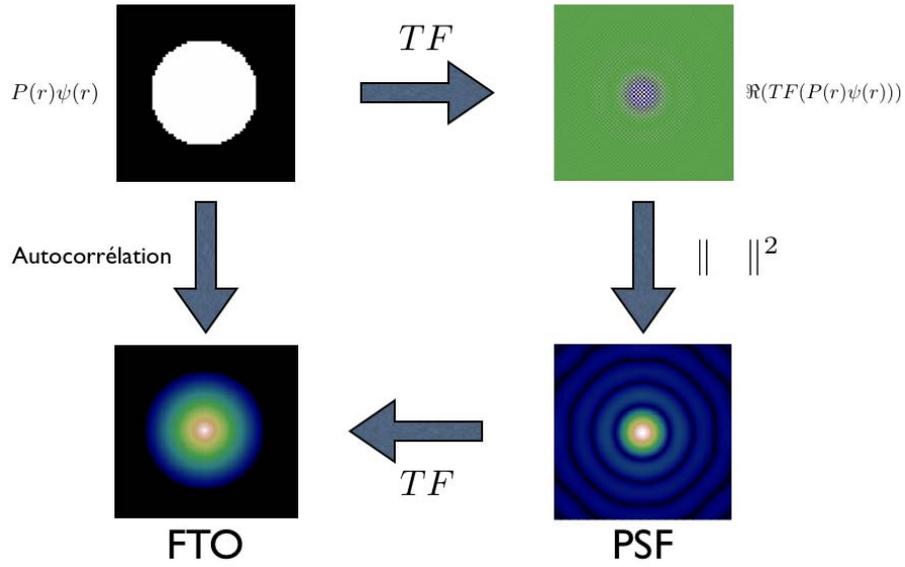


FIGURE 1.1 – Liens entre les différents plans et grandeurs FEP et FTO

de chaque colonne de pixels. Le transfert n'étant pas parfait, quelques électrons provenant de certains pixels sont récupérés par d'autres. On obtient finalement de légères variations de la valeur de chaque pixel. Le bruit de détecteur est un bruit additif, c'est-à-dire indépendant du signal vrai, suivant une loi gaussienne de moyenne nulle et dont la variance dépend du détecteur utilisé.

Ces deux bruit sont combinés dans l'image. Afin de simplifier les modèles, on considère que l'on a un unique bruit additif gaussien de moyenne nulle, et la variance de bruit sur un pixel k s'écrit :

$$\sigma_k^2 = \bar{I}_k + \sigma_e^2 \quad (1.5)$$

avec \bar{I}_k le flux moyen (flux vrai) sur ce pixel, et σ_e^2 la variance du bruit de détecteur. Cette approximation est valide pour les problèmes traités dans cette thèse [Mugnier *et al.*, 2004].

1.2.2 Télescope idéal

1.2.2.1 FEP et FTO d'un télescope idéal

En considérant un télescope comme un système imageur dont la FEP est invariante par translation, la FEP d'un télescope idéal s'exprime alors de la façon suivante :

$$FEP(\alpha, t) = |TF[P]_\alpha|^2 \quad (1.6)$$

où P définit la fonction pupille.

On considère par la suite, et par souci de simplicité, qu'un télescope se décrit par une pupille circulaire, bien qu'il serait préférable de considérer une pupille circulaire présentant

une obturation centrale correspondant à l’empreinte du miroir secondaire (cela a pour effet de déporter une partie de la luminosité du disque central vers les anneaux qui l’entourent).

Avec cette simplification, l’image d’un point source (étoile) observé par un télescope idéal, en l’absence de turbulence, est donnée par la fonction d’Airy à deux dimensions :

$$Airy(\alpha) = \frac{\pi D^2}{4\lambda^2} \left[\frac{2J_1\left(\frac{\pi D}{\lambda}\alpha\right)}{\left(\frac{\pi D}{\lambda}\alpha\right)} \right]^2 \quad (1.7)$$

$J_1(x)$ étant la fonction de Bessel du premier ordre, λ la longueur d’onde d’imagerie et D le diamètre du télescope. La tache d’Airy est illustrée sur les figures 1.2 et 1.3.

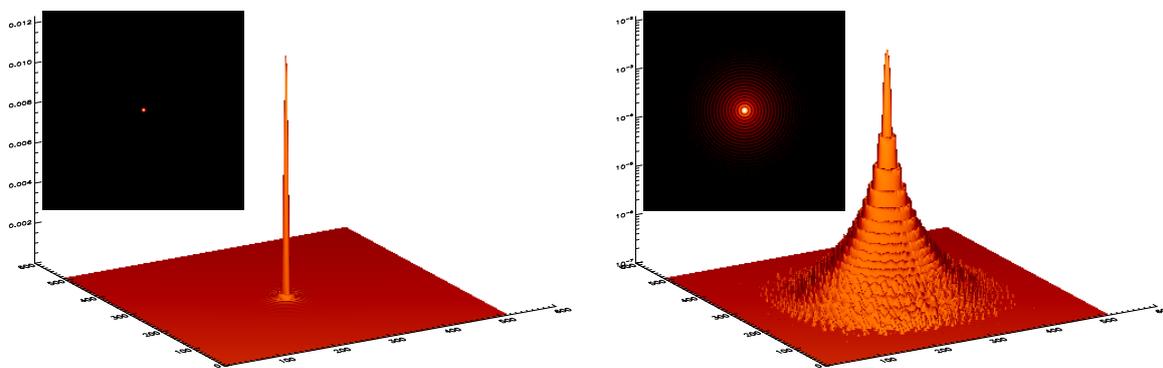


FIGURE 1.2 – Réponse impulsionnelle d’un télescope parfait (sans obturation centrale) = tache d’Airy. A gauche en échelle linéaire à droite la même représentation en échelle logarithmique.

Le premier anneau noir est situé à une distance angulaire de $1.22\lambda/D$ du centre. La largeur à mi-hauteur de la tache d’Airy vaut λ/D et définit la résolution "idéale" du télescope.

On peut en déduire aussi la FTO du télescope idéal par transformée de Fourier. On peut aussi, par utilisation du théorème de Wiener-Khintchine, montrer que la FTO correspond à l’autocorrélation de la pupille, ce qui donne la FTO proposée en figure 1.4.

On remarque ainsi que la FTO du télescope parfait joue le rôle d’un filtre passe bas, la fréquence de coupure du filtre étant directement liée au diamètre du télescope. Les fréquences supérieures à D/λ sont ainsi complètement filtrées par la fonction de transfert du télescope.

1.2.2.2 Échantillonnage

L’acquisition d’une image se fait par l’intermédiaire d’une matrice de pixels. On ne peut donc en avoir qu’une version échantillonnée. On définit la fréquence d’échantillonnage comme

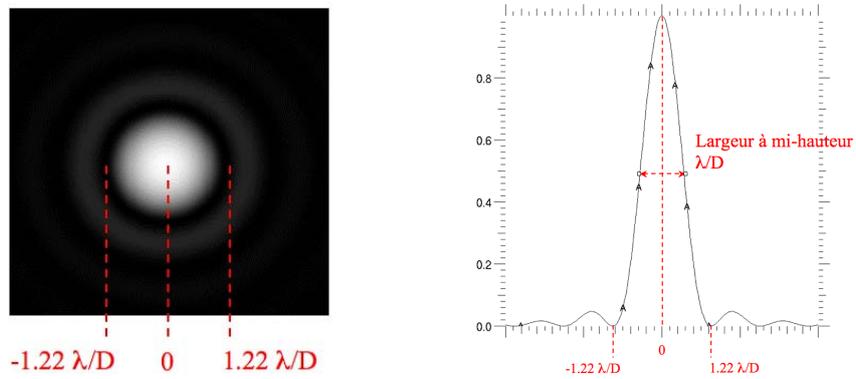


FIGURE 1.3 – Zoom sur la partie centrale de la tache d’Airy (échelle logarithmique). [A gauche] Image à 2 dimensions. [A droite] coupe selon x

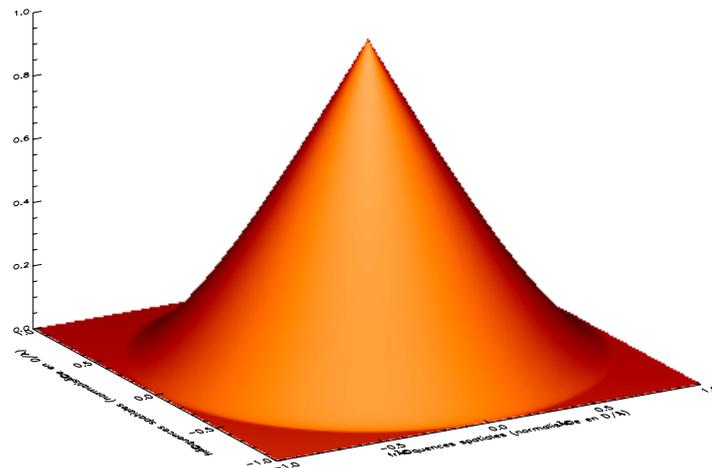


FIGURE 1.4 – Représentation 3D de la FTO d’un télescope parfait. Les axes sont normalisés à $D/\lambda = 1$ par simplicité.

l'inverse de la taille des pixels. Pour respecter le critère de Shannon, la fréquence d'échantillonnage doit être au moins deux fois supérieure à la fréquence maximale présente dans le signal. Si l'on a un système limité par la diffraction, la fréquence minimale d'échantillonnage, ou fréquence de Shannon, vaut $2D/\lambda$. Pour un échantillonnage dit "à Shannon", la dimension angulaire des pixels est donc $\lambda/(2D)$. Dans la suite, un échantillonnage dit "à K fois Shannon" signifie des pixels de dimension angulaire $\lambda/(2K D)$.

1.2.3 Imagerie en présence de turbulence

1.2.3.1 Effets optiques de la turbulences atmosphérique

Lorsqu'on observe un astre depuis la Terre, la lumière issue de cet astre traverse l'atmosphère avant de parvenir jusqu'à nous. L'atmosphère est un milieu dynamique : de nombreux courants d'air chaud et d'air froid se croisent dans l'ensemble de l'atmosphère. Ces croisements sont à l'origine de tourbillons qui constituent alors un milieu de température différente. Or, la température influe sur l'indice optique, dont les variations sont à l'origine du phénomène de réfraction décrit par la loi de Descartes. Les tourbillons, qui ont une géométrie sphérique, agissent alors comme des lentilles. La taille et la position des tourbillons étant aléatoires et évoluant rapidement, la trajectoire des rayons est fortement perturbée. Cela a une incidence directe sur le front d'onde de la lumière issue de l'astre. Ce front d'onde, initialement plan à son entrée dans l'atmosphère, est déformé lorsqu'il arrive au sol.

1.2.3.2 Courte pose et longue pose en présence d'aberrations

On considère désormais un télescope idéal associé à une propagation à travers la turbulence. L'atmosphère peut être vue comme une série de couches minces. La propagation à travers la turbulence se résume alors à une fluctuation de la phase du front d'onde liée à la sommation des fluctuations d'indice rencontrées à la traversée de chacune des couches turbulentes.

Dans ces conditions, la fonction d'étalement de point s'exprime de la façon suivante :

$$FEP(\alpha, t) = |TF [P(\mathbf{r}) \psi_\alpha(\mathbf{r}, t)]_\alpha|^2 \quad (1.8)$$

où ψ_α décrit le champ complexe. Comme nous ne considérons que des fluctuations de phase, ceci se résume à :

$$FEP(\alpha, t) = |TF [P(\mathbf{r}) \exp \{i\phi_{turb}(\mathbf{r}, t)\}]_\alpha|^2 \quad (1.9)$$

où P définit la fonction pupille, et ϕ_{turb} la phase turbulente. La turbulence étant un processus évolutif, la phase ainsi que la FEP dépendent du temps t .

On peut considérer la FEP "instantanée" ou courte pose, associée à une phase turbulente unique. On voit apparaître alors des tavelures (ou *speckles*) liés à l'interférence des divers rayons lumineux dans le plan image (voir figure 1.5). Chaque grain de speckle a pour taille caractéristique la limite de diffraction. La FEP courte pose s'étale sur un domaine de taille caractéristique λ/r_0 , r_0 étant le paramètre de Fried qui caractérise les effets de la turbulence [Rodier, 1999].

Dans la pratique, l'imagerie d'objets de faible intensité incite à augmenter le temps de pose. On doit donc considérer une FEP "longue pose", somme temporelle de FEP courtes poses.

La FEP longue pose n'est rien d'autre que l'intégration temporelle de FEP courte pose. Elle s'exprime donc simplement de la façon suivante :

$$FEP_{lp}(\alpha) = \langle FEP(\alpha, t) \rangle_t = \frac{1}{T} \int_0^T |TF [P(\mathbf{r}) \exp \{i\phi_{turb}(\mathbf{r}, t)\}]_{\alpha}|^2 dt \quad (1.10)$$

Le moyennage temporel conduit à la superposition de ces FEP, donc des tavelures, et finalement à la formation d'une tache de profil typiquement gaussien de largeur à mi-hauteur λ/r_0 (Fig. 1.5).

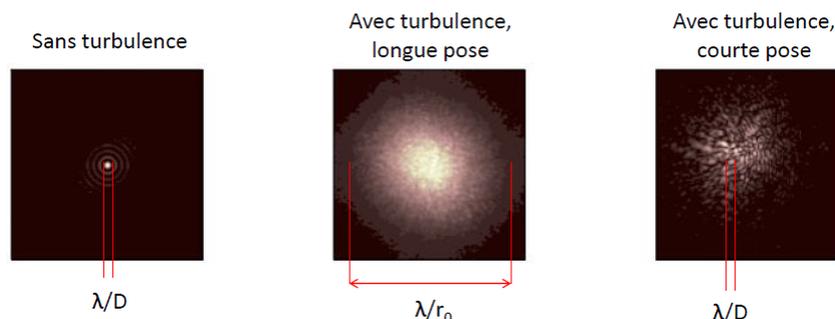


FIGURE 1.5 – Effet de la turbulence sur l'image. L'image sans turbulence correspond à la tache d'Airy. Des tavelures de la taille du pic de la tache d'Airy (λ/D) sont discernables dans l'image courte pose. Dans l'image longue pose, ces tavelures se moyennent et l'on obtient une tache de taille λ/r_0 .

L'image obtenue en présence de turbulence a donc une résolution beaucoup moins élevée que la résolution théorique du télescope. Afin de remédier à cela, on peut compenser les effets de la turbulence sur le front d'onde grâce à l'optique adaptative.

1.3 Optique adaptative

Nous expliquons ici le principe de l'optique adaptative et les différents éléments qui permettent de caractériser la qualité de correction du front d'onde. Nous nous intéressons ensuite aux différentes sources d'erreur qui mènent à une correction imparfaite. Nous décrivons enfin les systèmes d'optique adaptative dits "assistés par laser", qui vont nous intéresser dans la thèse.

1.3.1 Principe

L'Optique Adaptative (OA) est un système opto-mécanique qui corrige en temps réel le front d'onde turbulent intercepté par le télescope. Pour cela, un système d'optique adaptative se compose de trois éléments clés :

- un Analyseur de Surface d'Onde (ASO) qui mesure les défauts de *phase du front d'onde* incident
- un système informatique qui traite les mesures de l'ASO en temps réel (on parle souvent de Real Time Computer : RTC) et qui contrôle le troisième élément clé :

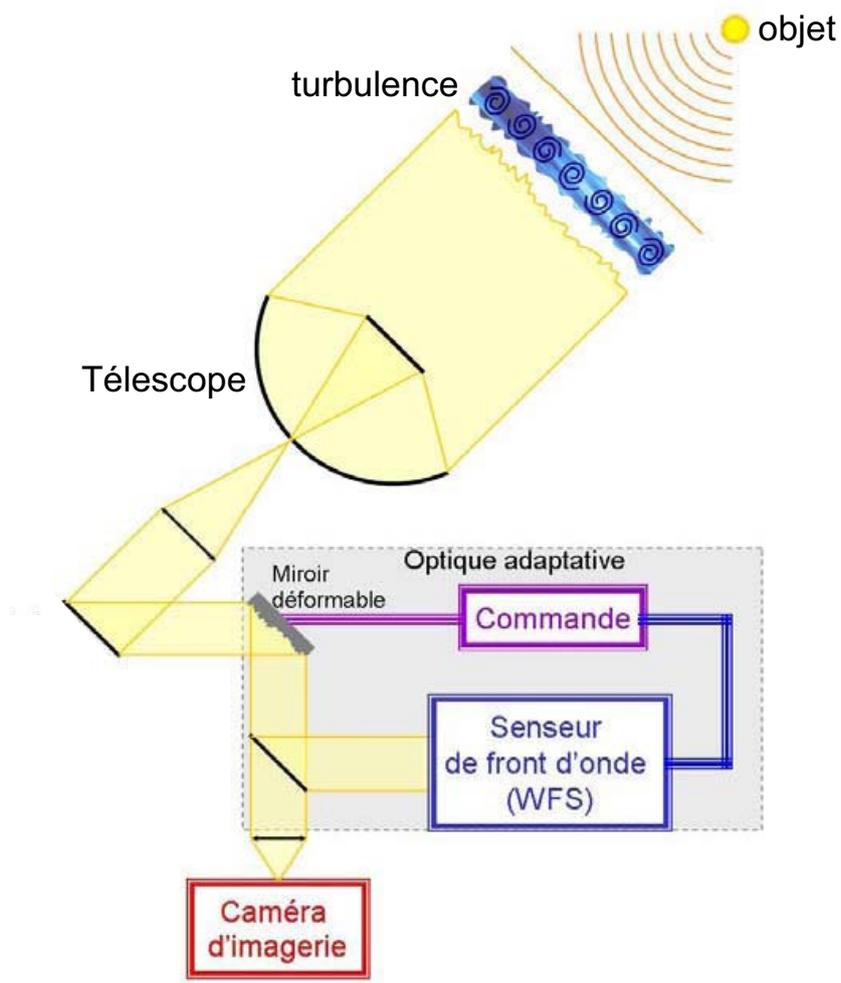


FIGURE 1.6 – Schéma de principe d'une optique adaptative.

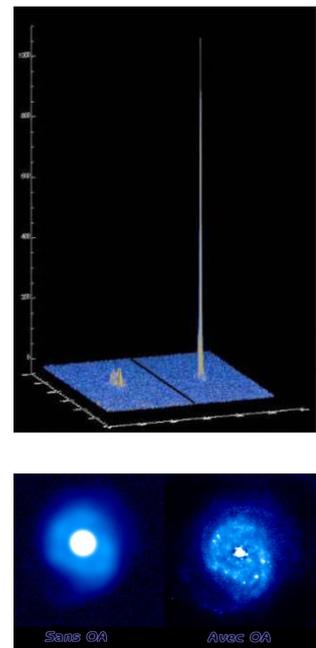


FIGURE 1.7 – Exemple de résultats d'OA. [en haut] cas d'une étoile simple ; [en bas] cas d'une galaxie

- un Miroir Déformable (DM en anglais) qui assure la correction en imprimant au front d'onde incident une déformation inverse de celle qui a été mesurée.

Le but de tout système d'OA est d'obtenir un front d'onde corrigé aussi plan que possible et permettre ainsi d'observer (sur une voie d'imagerie) une image corrigée proche de la limite théorique de diffraction du télescope. Comme on peut le remarquer sur la figure 1.6, le système de correction (miroir déformable) est situé en amont du système de mesure (analyseur de surface d'onde). Ce type de configuration est dite "boucle fermée" et reste la plus utilisée en OA. Le temps d'évolution de la turbulence étant très rapide (de l'ordre de 10 ms), la correction doit se faire à une cadence de quelques centaines de Hertz.

1.3.2 Qualité de correction

En pratique, de nombreuses limitations apparaissent et la correction d'une OA n'est jamais parfaite. Nous décrivons ici les différents critères pour évaluer la performance d'une OA.

1.3.2.1 Variances résiduelles

La variance de la phase après correction par optique adaptative (ou phase résiduelle) est le critère le plus couramment utilisé pour quantifier les performances d'une OA car il permet de travailler directement sur la grandeur sur laquelle le système agit. On définit tout d'abord la variance spatiale "instantanée" de la phase :

$$\sigma_{\phi_{res}}^2(t) = \frac{1}{S} \int (\phi_{res}(\mathbf{r}, t))^2 d\mathbf{r} - \left(\frac{1}{S} \int \phi_{res}(\mathbf{r}, t) d\mathbf{r} \right)^2 \quad (1.11)$$

qui caractérise l'énergie présente dans la phase résiduelle à chaque instant (c'est-à-dire après chaque correction par le miroir déformable). Si la correction est parfaite, la phase résiduelle sera nulle et sa variance aussi. L'intérêt de l'OA étant de corriger en temps réel et d'obtenir une longue pose (c'est-à-dire une image intégrée sur une durée significative par rapport à l'évolution typique du phénomène turbulent), on s'intéresse souvent à des quantités moyennées temporellement pour caractériser l'OA et ses performances. Ainsi, on définit une variance "moyenne" comme suit :

$$\sigma_{\phi_{res}, tot}^2 = \frac{1}{T} \int_0^T \sigma_{\phi_{res}}^2(t) dt \quad (1.12)$$

Notons que la variance de la phase est chromatique

$$\sigma_{\phi_{res}}^2(t, \lambda_2) = \sigma_{\phi_{res}}^2(t, \lambda_1) \left(\frac{\lambda_1}{\lambda_2} \right)^2 \quad (1.13)$$

Pour une OA donnée et donc un niveau de correction donné, la performance de l'optique.

1.3.2.2 Rapport de Strehl

Le rapport de Strehl (ou Strehl Ratio [SR] en anglais) est un critère particulièrement bien adapté à la caractérisation de systèmes d'OA à haute performance. Il est défini comme suit :

$$SR = \frac{FEP(\boldsymbol{\alpha} = 0)}{Airy(\boldsymbol{\alpha} = 0)} = \frac{\int FTO(\boldsymbol{\kappa}) d\boldsymbol{\kappa}}{\int FTO_{Airy}(\boldsymbol{\kappa}) d\boldsymbol{\kappa}} \quad (1.14)$$

avec $FTO_{Airy}(\kappa)$ la fonction de transfert de la tache d'Airy (c'est-à-dire de la réponse parfaite de l'instrument). Le rapport de Strehl donne donc une valeur intégrée quantifiant l'écart entre une FEP (ou FTO) corrigée par OA et la FEP (ou FTO) théorique d'un télescope parfait. Bien que le rapport de Strehl soit un critère lié à la FEP (resp. FTO) de l'instrument, il est possible de le relier très simplement à la variance de la phase résiduelle : $\sigma_{\phi_{res}}^2(t)$ (pour le cas d'une imagerie courte pose) ou $\sigma_{\phi_{res,tot}}^2$ (pour le cas d'une imagerie longue pose), en appliquant la formule de Maréchal

$$SR(t) \simeq Ec(t) = e^{-\sigma_{\phi_{res}}^2} \quad (1.15)$$

où Ec est l'énergie cohérente à savoir l'énergie présente dans le pic cohérent de la FEP corrigée par OA. Le rapport de Strehl décroît très vite lorsque la performance du système d'OA est dégradée (c'est-à-dire lorsque la variance résiduelle augmente) ce qui en fait un critère de performance pertinent pour les applications d'imagerie à très haute résolution.

1.3.2.3 Choix d'une base pour décrire la phase

Le choix d'une base de modes est essentiel dès que l'on veut étudier finement une optique adaptative, son comportement et ses performances. La variance de phase décrite plus haut donne une information spatiale intégrée. Il est nécessaire d'avoir un niveau de définition plus élevé en décomposant cette variance (et donc la phase elle même) sur une base judicieusement choisie. En pratique on utilisera principalement les bases décrites ci-après.

Base des pixels c'est le support le plus naturel pour décrire la phase. Chaque pixel définit une étendue spatiale (carrée) à laquelle on associe une valeur du signal continu (en général sa moyenne spatiale sur le support considéré). Plus le pixel sera petit plus la résolution spatiale sera importante.

Base des polynômes de Zernike Les polynômes de Zernike constituent une base sur un support circulaire. Proposés par Zernike en 1934, ces polynômes s'expriment en fonction d'ordres radiaux n , de fréquences azimutales m et de la parité de leur numéro i :

$$\begin{aligned} Z_{i \text{ pair}} &= \sqrt{n+1} R_n^m(\mathbf{r}) \sqrt{2} \cos(m\theta) & m \neq 0 \\ Z_{i \text{ impair}} &= \sqrt{n+1} R_n^m(\mathbf{r}) \sqrt{2} \sin(m\theta) & m \neq 0 \\ Z_i &= \sqrt{n+1} R_n^0(\mathbf{r}) & m = 0 \end{aligned}$$

avec

$$R_n^m(\mathbf{r}) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! [(n+m)/2 - s]! [(n-m)/2 - s]!} \mathbf{r}^{n-2s} \quad (1.16)$$

A noter que la base des polynômes de Zernike est une base orthonormée sur une pupille circulaire de surface S :

$$\frac{1}{S} \int_S Z_i(\mathbf{r}) Z_j(\mathbf{r}) d\mathbf{r} = \delta_{ij} \quad (1.17)$$

avec δ_{ij} valant 1 si $i = j$ et 0 autrement. On peut aussi remarquer, comme le montre la figure 1.8 que les premiers polynômes représentent les aberrations optiques classiques (comas, défocalisation, astigmatismes ...).

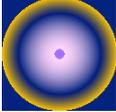
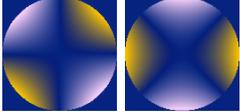
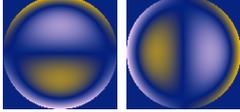
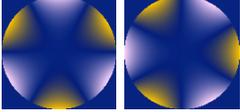
$\mathcal{Z}_p(\rho, \theta)$	m=0	m=1	m=2	m=3
n=0	$\mathcal{Z}_1 = 1$  piston			
n=1		$\begin{Bmatrix} \mathcal{Z}_2 \\ \mathcal{Z}_3 \end{Bmatrix} = 2\rho \begin{Bmatrix} \cos \theta \\ \sin \theta \end{Bmatrix}$  tip, tilt		
n=2	$\mathcal{Z}_4 = \sqrt{3}(2\rho^2 - 1)$  défocalisation		$\begin{Bmatrix} \mathcal{Z}_5 \\ \mathcal{Z}_6 \end{Bmatrix} = \sqrt{6}\rho^2 \begin{Bmatrix} \sin 2\theta \\ \cos 2\theta \end{Bmatrix}$  astigmatisme	
n=3		$\begin{Bmatrix} \mathcal{Z}_7 \\ \mathcal{Z}_8 \end{Bmatrix} = \sqrt{8}(3\rho^3 - 2\rho) \begin{Bmatrix} \sin \theta \\ \cos \theta \end{Bmatrix}$  Coma		$\begin{Bmatrix} \mathcal{Z}_9 \\ \mathcal{Z}_{10} \end{Bmatrix} = \sqrt{8}\rho^3 \begin{Bmatrix} \sin 3\theta \\ \cos 3\theta \end{Bmatrix}$  Trifoil

FIGURE 1.8 – Polynômes de Zernike de 1 à 21 rangés par ordres radiaux et fréquences azimutales. A noter la correspondance des premiers ordres avec les aberrations optiques classiques.

En utilisant cette base de polynômes, la phase turbulente s'exprime de la façon suivante :

$$\varphi(\mathbf{r}) = \sum_{i=1}^{\infty} a_i Z_i(\mathbf{r}) \quad (1.18)$$

où $\varphi(\mathbf{r})$ représente la phase turbulente et $Z_i(\mathbf{r})$ le $i^{\text{ème}}$ polynôme de Zernike.

Inversement, un coefficient de Zernike a_i de la décomposition modale de $\varphi(\mathbf{r})$ sera obtenu par :

$$a_i = \frac{1}{S} \int_S \varphi(\mathbf{r}) Z_i(\mathbf{r}) d\mathbf{r} \quad (1.19)$$

Les polynômes de Zernike sont classés par ordres radiaux et azimutaux croissants. Ce classement a de nombreux avantages, le premier étant qu'il correspond au classement en fréquences spatiales, c'est-à-dire les polynômes d'ordres radiaux élevés correspondent aux hautes fréquences spatiales. Cette analogie entre ordre radial et fréquence spatiale, ainsi que les expressions analytiques des polynômes de Zernike permettent de représenter, d'étudier et de caractériser le comportement statistique de la turbulence et l'effet sur l'optique adaptative à l'aide de formules analytiques simples. Le deuxième avantage de ce classement est qu'il correspond à l'énergie turbulente présente dans les modes. Plus les ordres sont élevés, plus l'énergie turbulente est faible.

Base des polynômes de Karhunen-Loève Les polynômes de Karhunen-Loève constituent une base orthogonale (orthonormée ici) que l'on peut calculer pour décomposer n'importe quel processus aléatoire, tel que les déformations du front d'onde par la turbulence, à partir de la statistique de ce processus. Ainsi, on peut décomposer la phase de la manière suivante :

$$\varphi(\mathbf{r}) = \sum_{i=1}^{\infty} b_i K_i(\mathbf{r}) \quad (1.20)$$

où K_i désigne le i -ème polynôme de Karhunen-Loève et b_i est le coefficient correspondant. On définit, de même que pour les polynômes de Zernike, l'orthonormalité par :

$$\frac{1}{S} \int_S K_i(\mathbf{r}) K_j(\mathbf{r}) d\mathbf{r} = \delta_{ij} \quad (1.21)$$

Les coefficients b_i sont statistiquement décorrélés, on a alors la covariance de b_i et b_j qui vaut :

$$\langle b_i b_j \rangle = \delta_{ij} \sigma_i^2 \quad (1.22)$$

où $\langle . \rangle$ désigne une moyenne temporelle et σ_i^2 est la variance du coefficient b_i .

D'autre part, la base des polynômes de Karhunen-Loève maximise l'énergie dans les bas ordres, ce qui permet d'avoir une bonne représentation de la phase en la décomposant sur peu de modes. Cette base possède donc les bonnes propriétés pour estimer finement un front d'onde turbulent [Lane et Tallon, 1992].

Il n'existe pas d'expression analytique des polynômes de Karhunen-Loève pour la décomposition d'un front d'onde turbulent. Cependant, ceux-ci peuvent être calculés en diagonalisant la matrice de covariance des coefficients de Zernike [Rodier, 1990]. Cannon a aussi proposé une méthode moins coûteuse en temps de calcul [Cannon, 1996], que j'utiliserai dans cette thèse.

1.3.2.4 Densité Spectrale de Puissance

La densité spectrale de puissance de la phase permet d'avoir accès à une répartition de la variance (énergie) d'une grandeur par fréquence spatiale :

$$DSP_{\phi_{res}}(\mathbf{f}) = \langle |TF[\phi_{res}(\mathbf{r})]|^2 \rangle \quad (1.23)$$

où $\langle . \rangle$ représente une moyenne d'ensemble. En pratique, l'équation 1.23 se calcule en remplaçant la moyenne d'ensemble par une moyenne temporelle.

$$DSP_{\phi_{res}}(\mathbf{f}) = \frac{1}{T} \int_0^T \left(|TF[\phi_{res,t}(\mathbf{r}, t)]_{f_x, f_y}|^2 \right) dt \quad (1.24)$$

Cette équation tend vers le cas théorique quand T tend vers l'infini. En pratique, en supposant que le temps de corrélation de la phase résiduelle vaut τ , la convergence de 1.24 vers 1.23 évolue en T/τ . Enfin on rappelle que :

$$\int \int_{-\infty}^{+\infty} DSP_{\phi_{res}(\mathbf{r})}(\mathbf{f}) d\mathbf{f} = \sigma_{\phi_{res}(\mathbf{r})}^2 \quad (1.25)$$

Cette grandeur est utile pour caractériser les erreurs à l'origine d'une correction imparfaite en optique adaptative. Dans la suite, nous détaillons les différentes sources d'erreur.

1.3.3 Budget d'erreur

Comme schématisé en figure 1.9, l'OA peut être représentée comme un système linéaire (ou autrement dit un système de filtrage spatial) transformant une phase turbulente en phase résiduelle (pour laquelle les basses fréquences spatiales ont été fortement atténuées). L'objectif étant d'obtenir le filtrage le plus efficace possible et de tendre idéalement vers une phase nulle. En pratique, ce but n'est jamais atteint, le processus de filtrage étant entaché de différentes erreurs provenant de diverses sources qu'on peut regrouper en 4 grandes catégories :

- Erreurs dites "fondamentales" qui représentent les limitations ultimes de l'OA et qui ne sont dues qu'à l'objet observé et la répartition volumique de la turbulence au dessus du télescope :
 - son extension angulaire qui, combiné à la répartition volumique de la turbulence, va conduire au phénomène d'anisoplanétisme (voir 1.3.3.1)
 - son flux limité qui va conduire à des erreurs de mesure de la surface d'onde ...
- Erreurs dites "technologiques", c'est-à-dire liées aux composants de la boucle (nombre d'actionneurs du miroir, nombre de points de mesure de l'analyseur, type de détecteurs...)
- Erreurs dites "systèmes", à savoir les erreurs de calibration (de la boucle d'OA) et les erreurs de "couplage" entre le système d'OA, le télescope et l'instrument ;
- Erreur dites "chromatiques" liées eux effets de propagation et de diffraction à travers la turbulence. Ces erreurs sont le plus souvent négligeables et nous ne les traiterons pas ici.

On va, en première approximation, considérer ces différentes erreurs comme décorréliées les unes par rapport aux autres. Ainsi on peut définir simplement, un budget d'erreur fréquence

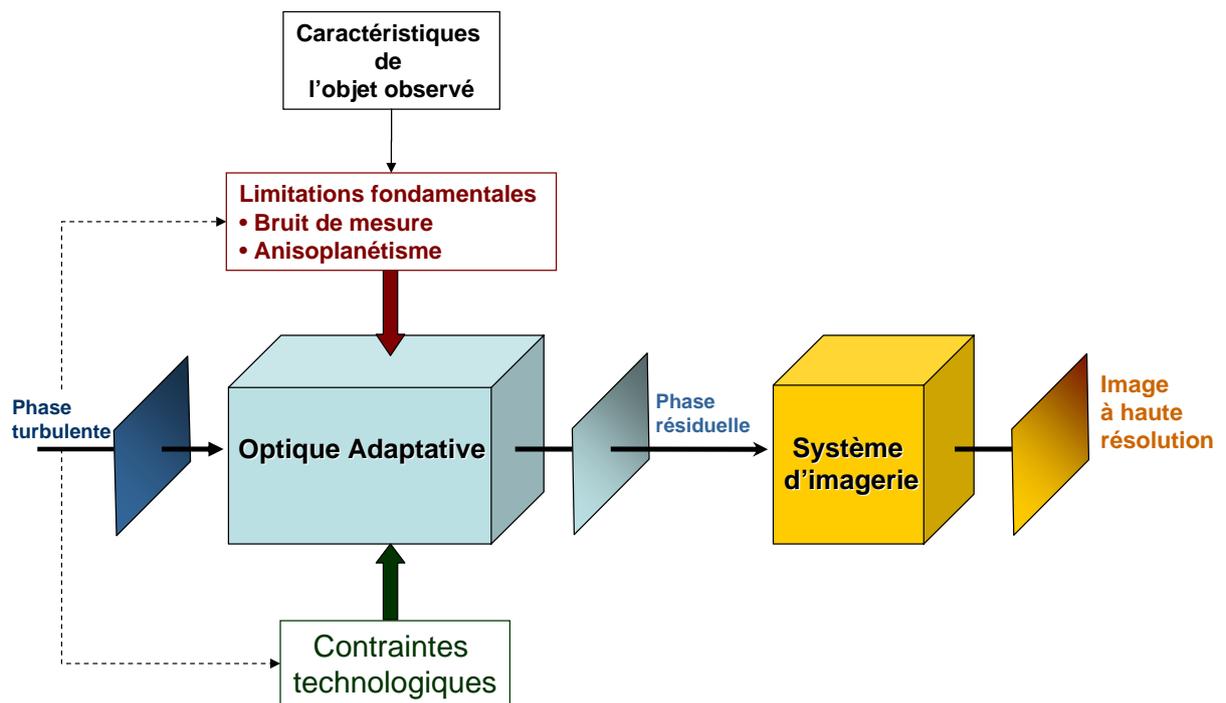


FIGURE 1.9 – Principe de la chaîne d'acquisition incluant un système d'OA.

spatiale par fréquence spatiale qui s'écrit comme une somme de densités spectrales de puissance associées à chaque type d'erreur, soit :

$$\begin{aligned}
 DSP_{res,tot} = & \underbrace{DSP_{anisoplanétisme} + DSP_{bruit\ de\ photons}}_{\text{Limitations fondamentales}} \quad (1.26) \\
 & + \underbrace{DSP_{ajustement} + DSP_{repliement} + DSP_{retard\ de\ boucle} + DSP_{bruit\ de\ détection}}_{\text{Limitations "technologiques"}} \\
 & + \underbrace{DSP_{étalonnage} + DSP_{aberrations\ non\ communes} + DSP_{exogènes}}_{\text{Limitations "systèmes"}}
 \end{aligned}$$

Le but principal de toute définition et dimensionnement d'un système d'OA sera de minimiser ce budget d'erreur dans un espace de paramètres qui sera défini par :

- les besoins scientifiques : types d'objets à observer ;
- les contraintes atmosphériques : conditions de turbulences à considérer ;
- les contraintes technologiques : types de composants disponibles ;
- les contraintes instrumentales : liens avec les instruments scientifiques et le télescope.

Ces différents termes d'erreurs sont détaillés ci-après.

1.3.3.1 Anisoplanétisme

Le phénomène d'anisoplanétisme limite la performance du système d'optique adaptative dès que la direction d'imagerie diffère de la direction d'analyse. C'est le cas lors de l'imagerie de scènes étendues ou bien lorsque l'objet d'intérêt scientifique est trop peu lumineux pour asservir la boucle dessus et que l'on a été contraint de choisir une autre source guide.

Comme le montre le schéma de la Fig. 1.10, le volume de turbulence traversé par l'onde lumineuse utilisée pour l'analyse de front d'onde est alors différent du volume turbulent affectant la voie d'imagerie. Plus précisément, c'est la turbulence d'altitude qui limite le domaine isoplanétique de l'instrument. La correction apportée n'est alors pas adaptée à la voie imagerie.

On caractérise cette erreur en évaluant l'angle isoplanétique du site d'observation. On retiendra ici la définition de Fried, d'après laquelle l'angle isoplanétique est la distance angulaire par rapport à l'axe optique pour laquelle l'erreur d'estimation de la phase turbulente reste inférieure à 1 rad² en variance. Roddier a exprimé cet angle en fonction du paramètre de Fried de la turbulence et de son profil [Roddier, 1999] :

$$\theta_0 = 0.314 \frac{r_0}{\bar{h}} \quad \text{avec} \quad \bar{h} = \left(\frac{\int_0^\infty h^{5/3} C_n^2(h) dh}{\int_0^\infty C_n^2(h) dh} \right)^{5/3} \quad (1.27)$$

où $C_n^2(h)$ désigne la constante de structure de l'indice de réfraction, qui définit la force locale des variations d'indice à une altitude donnée. L'angle isoplanétique étant dépendant de r_0 , il est chromatique. A titre d'exemple, il est de l'ordre de 1.8 arcsec à 0.5 μ m pour le site du VLT sur le mont Paranal au Chili.

Ce problème d'anisoplanétisme est très limitant pour certains champs applicatifs de l'optique adaptative. Il se traduit en premier lieu par une limitation drastique de la portion de ciel accessible aux instruments équipés d'une OA. On verra plus loin que la solution consiste alors à multiplier les directions d'analyse pour accéder à la connaissance de la répartition volumique de la turbulence.

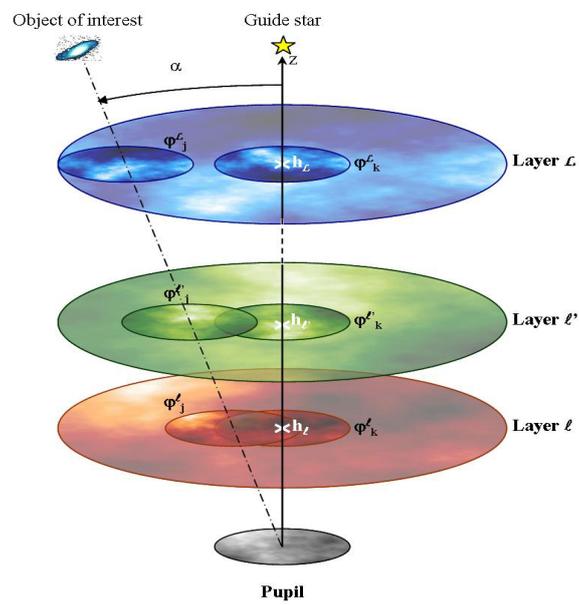


FIGURE 1.10 – Représentation schématique du phénomène d'anisoplanétisme.

1.3.3.2 Erreur d'ajustement

L'erreur d'ajustement est l'erreur qu'on commet en ne corrigeant qu'un nombre fini de modes (ou fréquences spatiales) de la phase turbulente. Cette erreur est intrinsèquement liée à la structure de l'élément correcteur (miroir déformable) qui va agir comme un filtre passe haut sur les fréquences spatiales de la turbulence. La coupure de ce filtre f_c^{DM} est donnée par

$$f_c^{DM} = 1/2d_{DM} \quad (1.28)$$

où d_{DM} est le pas inter-actionneur reporté dans la pupille, souvent appelé pitch DM. Si l'on suppose ici que la correction du DM est parfaite à toutes les fréquences plus petites que f_c^{DM} : aucune atténuation de la phase estimée ne sera considérée.

$$\tilde{\phi}_{corr} = N\tilde{\phi}_{turb} \quad (1.29)$$

avec :

$$N = \begin{cases} 1 & \text{pour } f_x \text{ et } f_y \leq f_c^{DM} \\ 0 & \text{sinon} \end{cases} \quad (1.30)$$

Ainsi la densité de puissance de l'erreur d'ajustement s'écrit

$$DSP_{ajustement}(f_x, f_y) = \begin{cases} \frac{0.023}{r_0^{5/3}} (f_x^2 + f_y^2)^{-11/6} & \text{pour } f_x \text{ et } f_y > f_c^{DM} \\ 0 & \text{sinon} \end{cases} \quad (1.31)$$

et la variance associée est égale à

$$\sigma_{ajustement}^2 = \int \int DSP_{ajustement}(f_x, f_y) df_x df_y = 0,232 \left(\frac{d_{DM}}{r_0} \right)^{5/3} \quad (1.32)$$

Notons qu'en fonction de la technologie utilisée pour réaliser le miroir déformable, nous nous approcherons plus ou moins du cas parfait. Le coefficient 0,232 sera donc à ajuster en fonction du type de miroir et de la disposition des points de correction dans la pupille.

1.3.3.3 Erreur de mesure

La structure de l'ASO introduit des erreurs de mesures, les deux principales étant le repliement et le bruit de mesure propagé lors de la reconstruction du front d'onde.

Effet de repliement Prenons l'exemple d'un analyseur de type plan pupille (cf §1.4 p.46). L'ASO réalise une mesure échantillonnée spatialement de la phase en divisant la pupille en sous-pupilles. Cet échantillonnage de la mesure provoque une périodisation de son spectre dans le plan de Fourier avec une fréquence $f_{ech} = 1/d$. Ainsi, lors de la mesure d'une fréquence f on est non seulement sensible à la valeur du spectre de la phase pour cette fréquence mais aussi à la valeur du spectre d'une haute fréquence ($f_{ech} - f$) introduit par le phénomène de périodisation. Ainsi, la mesure des basses fréquences de la phase (fréquences les plus importantes car contenant le plus d'énergie turbulente) va être entachée d'une erreur plus moins importante selon l'allure du spectre et l'échantillonnage de la mesure (taille de la sous-pupille). Cette erreur dépend donc des caractéristiques de l'ASO et du spectre haute fréquence de la phase turbulente.

Afin de minimiser les effets de repliement des hautes fréquences spatiales de la phase, un filtrage par un trou dans le plan focal peut être effectué [Poyneer et Macintosh, 2004].

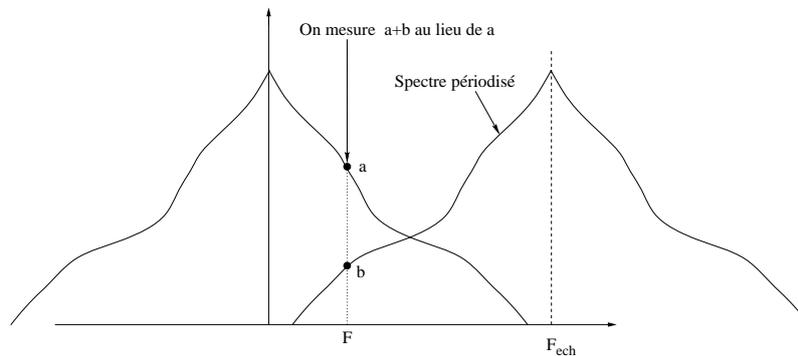


FIGURE 1.11 – Spectre spatial d’un signal échantillonné et mise en évidence du phénomène de repliement.

Propagation de bruit La propagation du bruit est un élément déterminant du budget d’erreur d’une OA. Elle est liée à la mesure de la phase turbulente et au processus de reconstruction. Quel que soit l’analyseur envisagé (cf. chapitre 1.4), l’information sur la phase turbulente est codée en terme d’intensité lumineuse lue sur un détecteur. A ce moment, la mesure est dégradée par du bruit de photon et du bruit de détecteur. Ensuite, la propagation du bruit lors du processus de reconstruction dépend de l’algorithme de reconstruction utilisé.

Cette erreur est souvent liée à la magnitude limite, qui désigne le flux minimal requis pour que l’erreur de reconstruction de l’ASO soit inférieure à un objectif fixé. La magnitude est une unité logarithmique utilisée en astronomie pour quantifier la luminosité d’un astre. Les magnitudes faibles correspondent à des flux élevés, et on passe d’une magnitude à la magnitude supérieure en divisant le flux par 2,5. La référence est l’étoile Véga, de magnitude 0.

1.3.3.4 Erreur temporelle

La boucle d’optique adaptative contient deux éléments introduisant des retards de la commande sur l’évolution de la turbulence : le détecteur de l’analyseur de front d’onde et la commande. Des deux, c’est généralement le premier qui génère le plus grand retard. Ainsi, la boucle d’OA est rythmée temporellement par l’acquisition des mesures de phase. A l’instar de ce qui se passe pour l’anisoplanétisme, le retard de la boucle sur la turbulence génère une erreur, qu’il est possible de compenser en partie en utilisant un modèle d’évolution temporelle de la turbulence et une commande de type prédictive, que l’on ne détaillera pas ici.

1.3.3.5 Erreurs d’étalonnage de la boucle d’OA

L’optique adaptative intègre un analyseur de front d’onde, un calculateur temps réel et un élément correcteur. Ces différents composants doivent être étalonnés (individuellement et les

uns par rapport aux autres) afin de pouvoir trouver les bonnes déformations à appliquer en fonction des mesures obtenues. La validité de ces étalonnages évolue au cours du temps, en particulier à cause aux instabilités opto-mécaniques du système (dilatations/contractions dues aux fluctuations de températures, contraintes dues aux changements d'orientations par rapport au vecteur gravité ...). Le temps dédié à ces étalonnages est souvent réduit sur un système opérationnel, ce qui conduit à des erreurs sur la mesure du front d'onde et sur l'application de la correction qui se traduisent par une perte en performance du système complet. Il est donc essentiel lors de la conception même d'un système d'optique adaptative de s'assurer de la stabilité de ces composants mais aussi de prévoir et d'optimiser (en terme de vitesse et d'efficacité) les procédures d'étalonnage.

1.3.3.6 Aberrations non communes

Par construction, un système d'OA va tendre à fournir un front d'onde plan au niveau de l'analyseur de surface d'onde (c'est à dire là où l'on mesure effectivement la perturbation que l'on va vouloir corriger avec le système). A contrario, l'acquisition des données scientifiques se fait sur la voie d'imagerie. Toute aberration différentielle entre ces 2 voies conduit à une perte de performance. Ces aberrations différentielles sont souvent désignées par le terme aberrations non communes (non-common path aberrations, NCPA en anglais). Elles se séparent en 2 catégories :

- les aberrations situées en aval de la lame séparatrice *sur la voie d'analyse*. Elles sont vues par l'analyseur de surface d'onde et donc mesurées et réinjectées sur le miroir déformable. Elles vont donc être corrigées, mais comme elles n'affectent **que** la voie d'analyse, elles vont se rajouter sur la voie d'imagerie et finalement dégrader l'image scientifique d'intérêt.
- les aberrations situées en aval de la lame séparatrice *sur la voie d'imagerie*. Elles ne sont pas vues par l'analyseur de surface d'onde et donc non-corrigées par l'OA.

La solution est de mesurer ces deux jeux d'aberrations et de modifier la phase de correction au niveau du miroir déformable (ou, et c'est souvent le cas, de modifier les références de l'analyseur de surface d'onde) pour ignorer les aberrations de la voie d'analyse et pre-compenser les aberrations de la voie d'imagerie. On assure ainsi que l'OA permettra de tendre vers un front d'onde plan non plus au niveau de la voie d'analyse mais au niveau de la voie d'imagerie, c'est-à-dire là où l'on va acquérir les données scientifiques.

1.3.3.7 Erreurs exogènes

Ce terme regroupe les sources d'erreurs issues de l'environnement du système d'OA et qui perturbent son fonctionnement. Il peut contenir les erreurs dues aux vibrations mécaniques se propageant dans le système et affectant la mesure du front d'onde ou l'acquisition d'images.

Dans la suite, nous revenons sur les limitations induites par l'erreur d'anisoplanétisme, et montrons les solutions envisagées pour dépasser ces limites.

1.3.4 Optique adaptative assistée par laser

L'anisoplanétisme limite le nombre d'objets pour lesquels on peut avoir une bonne qualité de correction. Afin d'augmenter ce nombre, de nouveaux systèmes, qui utilisent des étoiles

artificielles créées par laser, ont vu le jour. Nous expliquons ici l'effet de l'anisoplanétisme sur ce qu'on appelle la couverture de ciel, ainsi que le principe des étoiles laser et des différents systèmes les utilisant.

1.3.4.1 Couverture de ciel

La mesure des aberrations peut se faire soit directement sur l'objet d'intérêt si celui-ci est suffisamment brillant, soit sur une source proche de celui-ci. Dans le cas d'observation de galaxies lointaines, la faible brillance de surface de ces objets rend nécessaire l'utilisation d'une autre référence. Du fait de la décorrélation angulaire de la turbulence et de l'erreur d'anisoplanétisme qui en découle, la source de référence doit se situer dans un rayon relativement proche de l'objet d'intérêt, de l'ordre de θ_0 . Au delà, la décorrélation angulaire du front d'onde introduira une perte sensible de performance et donc de résolution angulaire. Cette contrainte limite énormément le nombre d'objets pour lesquels une bonne correction de la turbulence est réalisable. On parle de couverture du ciel, c'est-à-dire de zones de ciel où l'on peut espérer une correction meilleure qu'un certain seuil (seuil à fixer en fonction du programme scientifique).

En considérant l'ensemble de la sphère céleste, la surface totale du ciel est de $2 \times 2\pi r^2 = 4\pi(360/2\pi)^2 = 41253$ degrés carrés soit 5.34×10^{11} secondes d'arc carrées. Dans l'infrarouge, le domaine isoplanétique est de l'ordre de quelques dizaines de secondes d'arc. On peut donc estimer qu'il y a quelques centaines de millions de domaine isoplanétique sur l'ensemble du ciel. Or, le nombre d'étoiles suffisamment brillantes pour permettre une bonne mesure, et donc une bonne correction, en OA est lui de l'ordre de quelques millions à quelques dizaines de millions. Il y a donc statistiquement quelques pourcents (%) de chance d'avoir une étoile suffisamment brillante dans un domaine isoplanétique. Ce calcul simpliste, qui ne tient pas compte de l'inhomogénéité de distribution des étoiles dans le ciel, ni de la magnitude limite de l'ASO considéré, donne néanmoins un ordre d'idée assez proche de la réalité. Les étoiles étant plus nombreuses dans le plan galactique qu'aux pôles, la couverture du ciel déduite de ce calcul est ainsi sous-estimée au plan galactique et surestimée aux pôles galactiques. Notons qu'il dépend aussi de la longueur d'onde. La couverture de ciel augmente aux grandes longueurs d'onde et devient quasi nulle dès que l'on s'intéresse à l'optique adaptative dans le visible.

1.3.4.2 Étoiles laser

Pour accroître la couverture de ciel, plusieurs auteurs ont étudié l'utilisation d'étoiles artificielles (étoile laser, principe proposé en 1985 par R. Foy et A. Labeyrie [Foy et Labeyrie, 1985]). Le principe d'une étoile laser est basé sur la ré-émission de la lumière par des particules excitées. Il existe deux grandes classes d'étoiles laser :

- Les étoiles laser de type Rayleigh utilisent la diffusion Rayleigh (choc élastique) des atomes ou molécules dans l'atmosphère. L'émission induite par un laser induit un moment de dipôles qui est ré-émis à la même fréquence que la fréquence d'excitation. La diffusion Rayleigh étant proportionnelle à λ^{-4} , il est préférable d'utiliser de courtes longueurs d'ondes (UV, bleu ou vert) afin d'augmenter son efficacité. Ce type d'étoile laser est utilisé par exemple sur le démonstrateur CANARY [Morris *et al.*, 2014].
- Les étoiles laser de type Sodium qui excitent une couche de molécules de Sodium présente dans la haute atmosphère à environ 90km d'altitude (mésosphère terrestre). La

couche de Sodium est d'une épaisseur caractéristique d'environ 10km [Pfrommer et Hickson, 2010]. La lumière observée provient de l'absorption de l'énergie émise par le laser qui change l'état d'excitation de la transition atomique D2 des molécules de Sodium. Par désexcitation spontanée, l'énergie est ré-émise à la même fréquence que la fréquence d'émission ($\lambda = 589\text{nm}$, couleur orange). Ce type d'étoile laser est utilisé par exemple au télescope Gemini Sud [Neichel *et al.*, 2010].

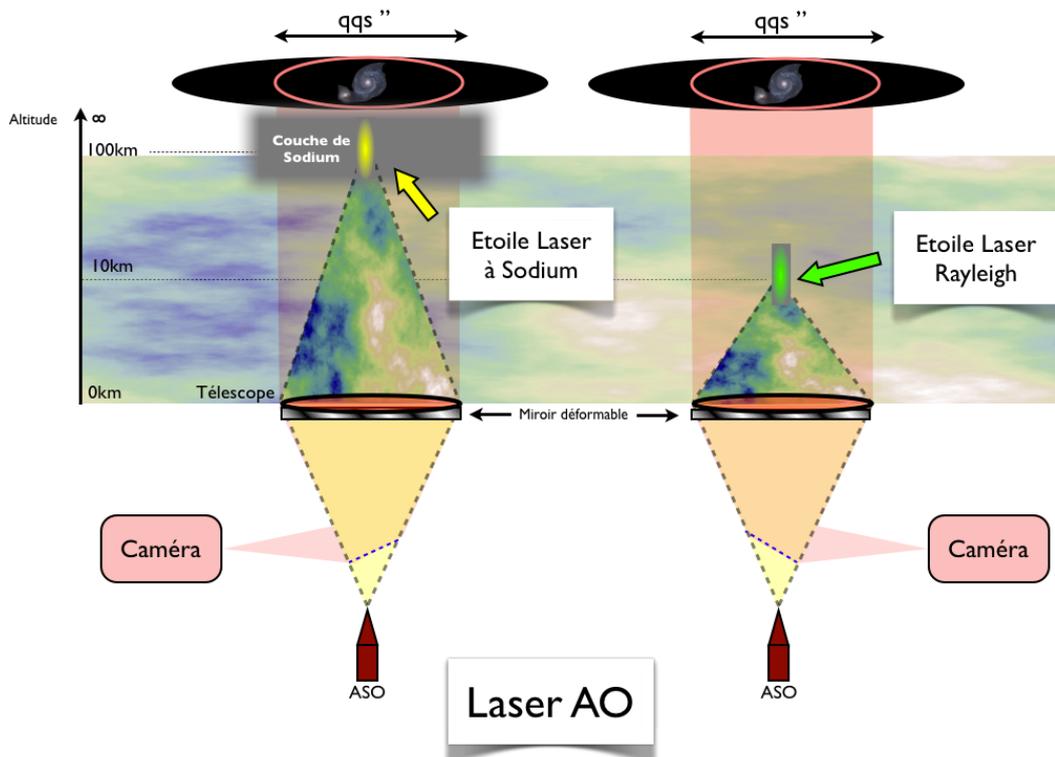


FIGURE 1.12 – Schéma représentatif d'une OA fonctionnant avec une étoile laser Sodium et une étoile laser Rayleigh.

En créant une source brillante n'importe où dans le ciel, l'étoile laser est donc la solution aux problèmes de couverture de ciel. Malheureusement, l'étoile laser souffre de plusieurs limitations intrinsèques qui rendent nécessaire l'utilisation d'une étoile naturelle pour un fonctionnement optimal. La couverture de ciel apportée par les étoiles lasers n'est donc pas nécessairement de 100%, mais elle reste quand même très supérieure au cas de l'OA fonctionnant uniquement sur étoiles naturelles.

Parmi ces limitations on notera essentiellement :

- l'effet de cône : l'EG laser émet sa lumière à une altitude finie, contrairement au cas d'un objet stellaire. L'onde reçue n'est donc plus plane mais sphérique. Ainsi les perturbations de phase vues par l'ASO ne sont pas exactement les mêmes que celles obtenues sur une EG naturelle. Plus le diamètre du télescope est grand, et plus l'étoile laser est focalisée près du sol, plus l'effet de cône devient important [Tallon et Foy, 1990] (voir figure 1.13).

Pour limiter son impact, il est nécessaire d'utiliser plusieurs EG lasers afin de synthétiser le cylindre d'une étoile naturelle avec plusieurs faisceaux coniques (cf §1.3.4 p.39).

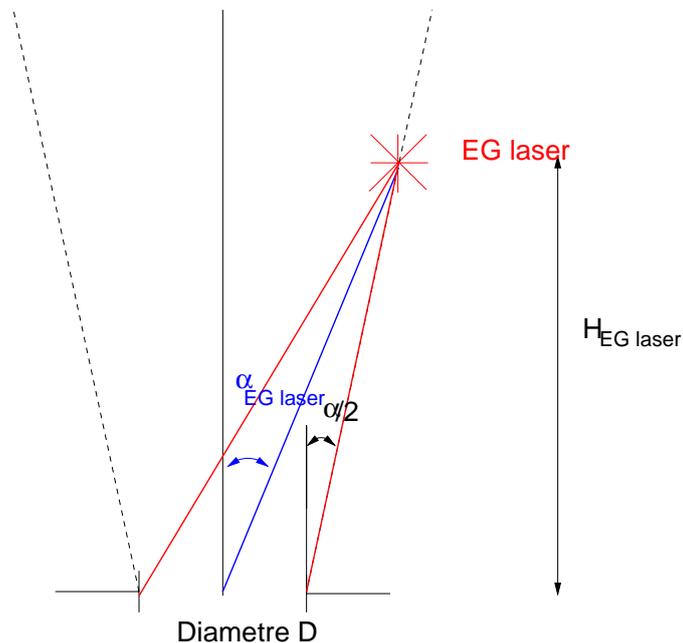


FIGURE 1.13 – Représentation schématique du problème d'effet de cône avec étoile laser.

- Indétermination du tip-tilt : La position de l'EG laser sur le ciel est toujours inconnue du fait du principe de retour inverse de la lumière. Le faisceau laser parcourt toujours le même trajet et n'est pas affecté par les modes tip-tilts dus à l'atmosphère. La mesure de ces modes n'est donc pas possible sur une EG laser [Rigaut et Gendron, 1992, Ellerbroek et Rigaut, 2001]. Pour pallier ce défaut, la plupart des systèmes à EG laser prévoit l'utilisation d'au moins une étoile naturelle pour faire la mesure des tip-tilt. Néanmoins la contrainte sur cette étoile naturelle est moins forte que dans le cas de l'OA classique. Cette étoile ne servant qu'à mesurer le mode tip-tilt (qui est le mode de plus basse fréquence spatiale), peut donc être moins brillante (utilisation de la pleine pupille, ce qui augmente le flux disponible). Bien évidemment, le tilt est lui aussi soumis au phénomène d'anisoplanétisme.
- Indétermination du focus : La couche de sodium a une altitude moyenne qui varie au cours du temps, ce qui provoque une erreur dans l'estimation du focus sur une étoile laser de type sodium [Neichel *et al.*, 2013, Pfrommer et Hickson, 2010]. En effet, la variation d'altitude s'ajoute au focus turbulent, et il est impossible de différencier l'un de l'autre dans la mesure. De la même manière que pour le tip/tilt, une étoile naturelle est donc nécessaire pour mesurer le focus lorsqu'on utilise des étoiles sodium.

1.3.4.3 Systèmes d'optique adaptative assistée par laser

Les systèmes d'optique adaptative qui sont décrits dans ce paragraphe reposent sur la connaissance du volume de turbulence et sont dits "tomographiques". Selon les objectifs, une optique adaptative doit permettre d'optimiser la correction dans une direction choisie (LTAO)

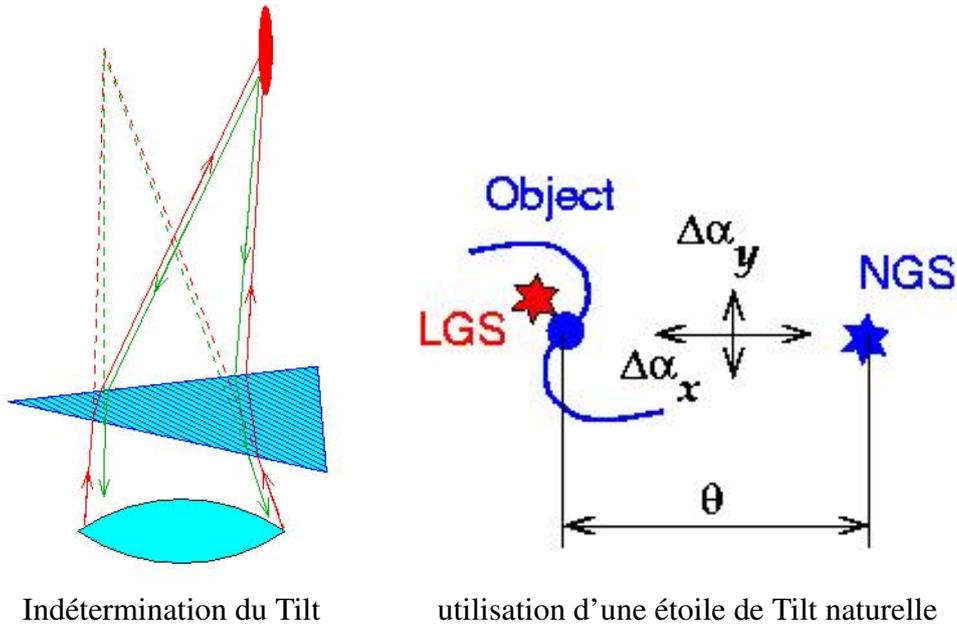


FIGURE 1.14 – [A gauche] Principe d'indétermination du tilt avec une étoile laser. [A droite] mesure de tilt sur une étoile naturelle.

ou d'obtenir la meilleure correction possible dans un champ de quelques arcmin (MCAO).

Optique adaptative tomographique multi-laser : LTAO L'utilisation d'une étoile laser seule ne peut pas donner une performance équivalente à celle délivrée par une étoile naturelle, en particulier à cause de l'effet de cône. Cet effet peut néanmoins être contourné en utilisant plusieurs étoiles laser qui, pointées autour de l'objet d'intérêt, permettent de sonder tout le volume de turbulence au dessus du télescope [Tallon et Foy, 1990, Le Louarn *et al.*, 2000, Baranec *et al.*, 2006]. Le principe est illustré en figure 1.15. Ce schéma impose alors une analyse multi-directionnelle de la turbulence. L'idée est d'utiliser autant d'ASO qu'il y a d'étoiles laser. Chacun d'entre eux observe la lumière provenant d'une étoile laser placée dans le champ. Chaque ASO va alors mesurer l'intégrale de la perturbation atmosphérique dans chaque direction d'analyse hors de l'axe de correction. La diversité angulaire des mesures permet de reconstruire le volume 3D de turbulence à partir des informations 2D intégrées dans chaque direction. Une fois connu le volume de turbulence dans chaque direction d'analyse, il faut alors reconstruire le volume de turbulence à corriger dans l'axe d'intérêt et appliquer la correction au miroir déformable. Ce type d'optique adaptative, qui a pour but de fournir une correction équivalente à celle que pourrait fournir une OA classique dans l'axe avec étoile naturelle, est appelée Laser Tomography Adaptive Optics ou LTAO. Grâce à l'utilisation d'étoiles laser, elle a pour but de fournir une couverture du ciel (proche) de 100%.

Correction dans tout le champ : MCAO La création d'étoiles laser permet donc d'augmenter considérablement la couverture du ciel, passant de quelques pourcents pour l'optique adaptative classique avec étoile naturelle à, potentiellement, 100%. En revanche, les systèmes restent toujours limités par la taille du champ de correction. Toutefois, en supposant que nous

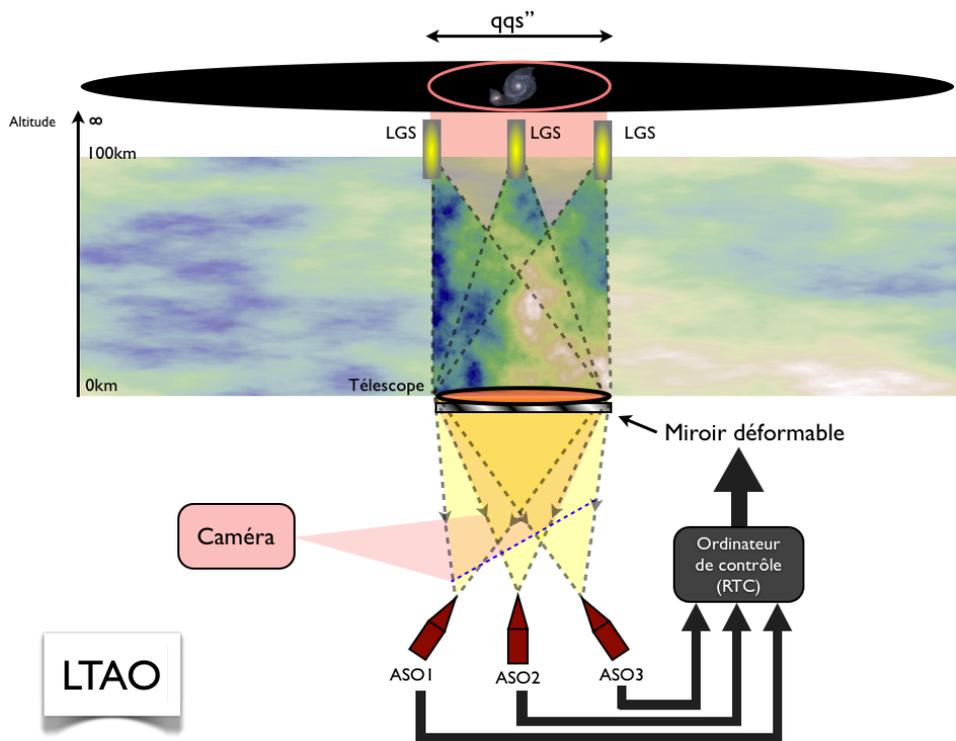


FIGURE 1.15 – Schéma de principe de la LTAO

maîtrisons parfaitement cette opération tomographique et que nous soyons donc capables de bien connaître le volume de turbulence au dessus du télescope, il est possible de corriger le volume de turbulence en conjuguant optiquement plusieurs miroirs déformables à plusieurs distances équivalentes dans le volume turbulent [Tallon *et al.*, 1992, Johnston et Welsh, 1994, Fusco *et al.*, 2001] : c'est le concept d'optique adaptative multiconjuguée (ou Multiconjugate adaptive optics [MCAO] en anglais). Un tel système (voir figure 1.16 utilise typiquement 2 à 3 miroirs

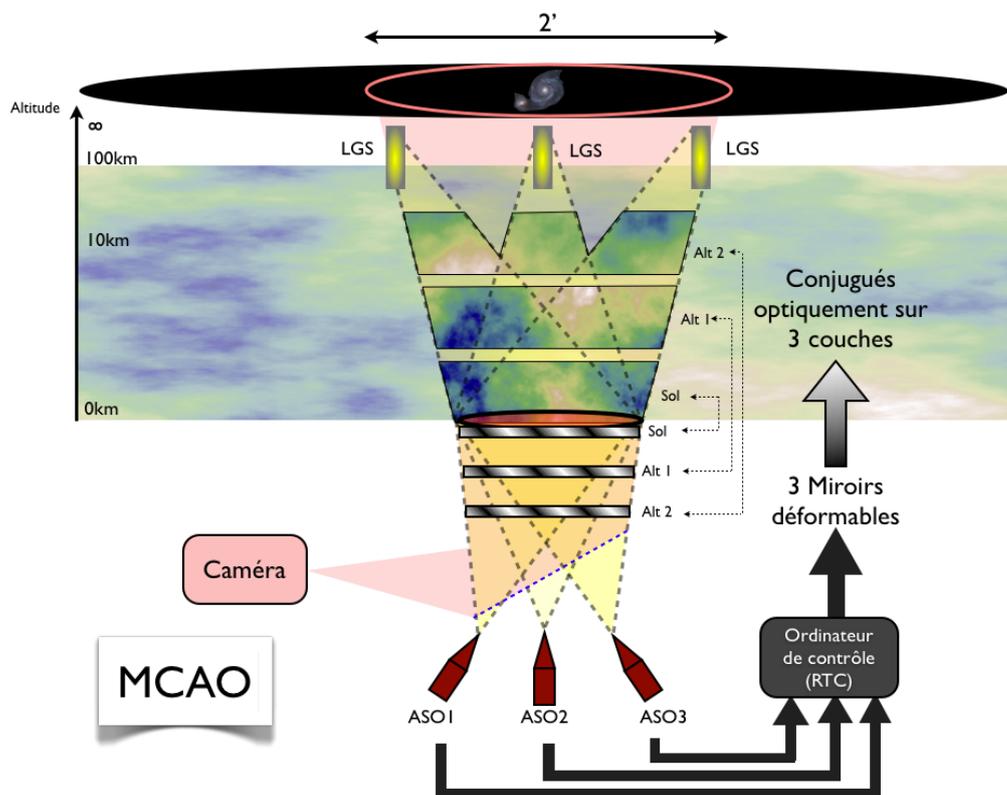


FIGURE 1.16 – Schéma de principe de la MCAO

déformables (cela est suffisant, d'après [Fusco *et al.*, 1999]) : 1 conjugué optiquement à la couche au sol (0m) et 2 autres conjugués à des altitudes correspondant aux couches les plus turbulentes de l'atmosphère sur le site d'observation (ex : 2000m et 10000m). Comme dans le cas de la LTAO, l'utilisation d'étoiles laser permet de s'affranchir du problème de la couverture du ciel. L'objectif de ce type de système est d'obtenir une correction proche de la limite de diffraction pour des champs de l'ordre de 1' à 2' (soit typiquement 5 à 10 fois plus grands qu'en OA "simple").

Ces systèmes, LTAO et MCAO améliorent donc grandement la couverture de ciel et permettent de s'affranchir de l'erreur d'anisoplanétisme. Cependant, les autres termes d'erreur (cf §1.3.3 p.33), notamment les erreurs liées à l'analyseur (ou aux analyseurs) de front d'onde subsistent. D'autre part, la couverture de ciel reste limitée, car une étoile naturelle est requise pour l'estimation des bas ordres. Le choix des analyseurs d'un système est donc déterminant à la fois pour la qualité de la correction et pour la couverture de ciel.

Dans la suite, nous nous intéressons aux analyseurs qui peuvent être envisagés sur un système d'optique adaptative, ainsi qu'à une méthode pour les comparer.

1.4 Analyseurs de front d'onde et leur comparaison

L'analyseur de front d'onde est le sous-système de l'optique adaptative chargé de mesurer la déformation du front d'onde incident. Dans les applications de l'analyse de front d'onde pour l'astronomie, on demande essentiellement à ces senseurs une bonne robustesse au bruit et un temps de mesure de la phase compatible avec le caractère temps réel du système (de l'ordre de quelques millisecondes). Plusieurs types de senseurs de front d'onde ont été proposés pour l'optique adaptative, ils sont le plus souvent déduits de concepts existants qui sont optimisés pour répondre aux besoins spécifiques de l'OA (faible flux, temps réel ...). Les plus utilisés en optique adaptative sont présentés ci-dessous. On pourra distinguer ici 2 grandes catégories :

- les analyseurs "en plan focal" ;
- les analyseurs "en plan pupille".

Les premiers s'appuient sur la répartition de l'intensité au plan focal pour reconstruire le front d'onde tandis que les seconds permettent d'estimer localement la phase en échantillonnant la pupille.

Nous nous intéresserons dans la thèse à deux configurations pour chaque analyseur :

- analyse des bas ordres (tip/tilt et focus) sur étoile naturelle afin de compléter l'analyse des hauts ordres sur étoile laser de type sodium.
- analyse des hauts ordres sur étoile naturelle.

Dans le premier cas, on estime peu de modes. On peut alors optimiser l'analyseur pour minimiser la sensibilité au bruit, ce qui permet d'analyser le front d'onde sur des étoiles plus faibles.

Dans le deuxième cas, on privilégie le nombre de modes estimés plutôt que la sensibilité au bruit. Nous nous limitons au cas de l'étoile naturelle car les étoiles laser forment un objet étendu qui varie dans le champ, ce qui complique les modèles. On peut cependant supposer que tous les analyseurs sont impactés de manière similaire par cet effet. De plus, les études qui seront présentées resteront valides pour les systèmes d'optique adaptative n'utilisant pas d'étoiles laser.

1.4.1 Analyseurs classiques

Nous présentons ici le principe des principaux analyseurs existants ou envisagés sur une optique adaptative. Nous décrivons dans chaque cas le principe, ainsi que la méthode de simulation et le modèle de bruit utilisés dans la thèse. Nous discutons aussi de leurs limitations respectives pour l'analyse des bas ordres sur une étoile naturelle faible. Les limitations pour l'analyse des hauts ordres seront évoquées dans le chapitre 5 p.149.

1.4.1.1 Shack-Hartmann

Le Shack-Hartmann est un analyseur de front d'onde plan-pupille. C'est l'analyseur le plus communément utilisé en optique adaptative.

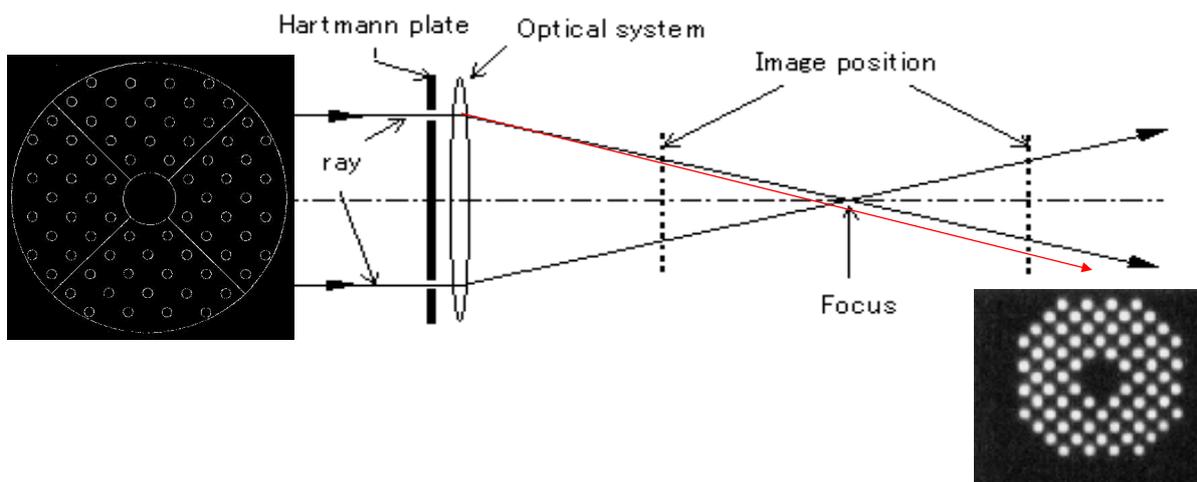


FIGURE 1.17 – Schéma de principe de fonctionnement d'un test de Hartmann.

Principe Il est basé sur le principe du test de Hartmann (voir figure 1.17). Un réseau de trous placé dans la pupille matérialise des rayons. La déviation angulaire d'un rayon est proportionnelle à la pente locale du front d'onde. La déviation angulaire est proportionnelle au déplacement du rayon dans un plan extra-focal. L'avantage de ce type de test est sa simplicité de mise en oeuvre (tests de grandes optiques). Ses deux inconvénients majeurs sont une perte de flux et un mauvais échantillonnage spatial du front d'onde.

Le Shack-Hartmann reprend le principe du test de Hartmann en le modifiant pour s'affranchir de ses limitations. Les trous sont remplacés par des micro-lentilles qui permettent d'exploiter l'ensemble des photons incidents sur la pupille du télescope [Shack et Platt, 1971]. Chaque micro-lentille définit une sous-pupille. Les images formées au plan focal de chaque sous-pupille vont nous donner une information sur la pente moyenne du front d'onde par l'intermédiaire d'une mesure de centre de gravité (voir figure 1.18). La pente moyenne locale de

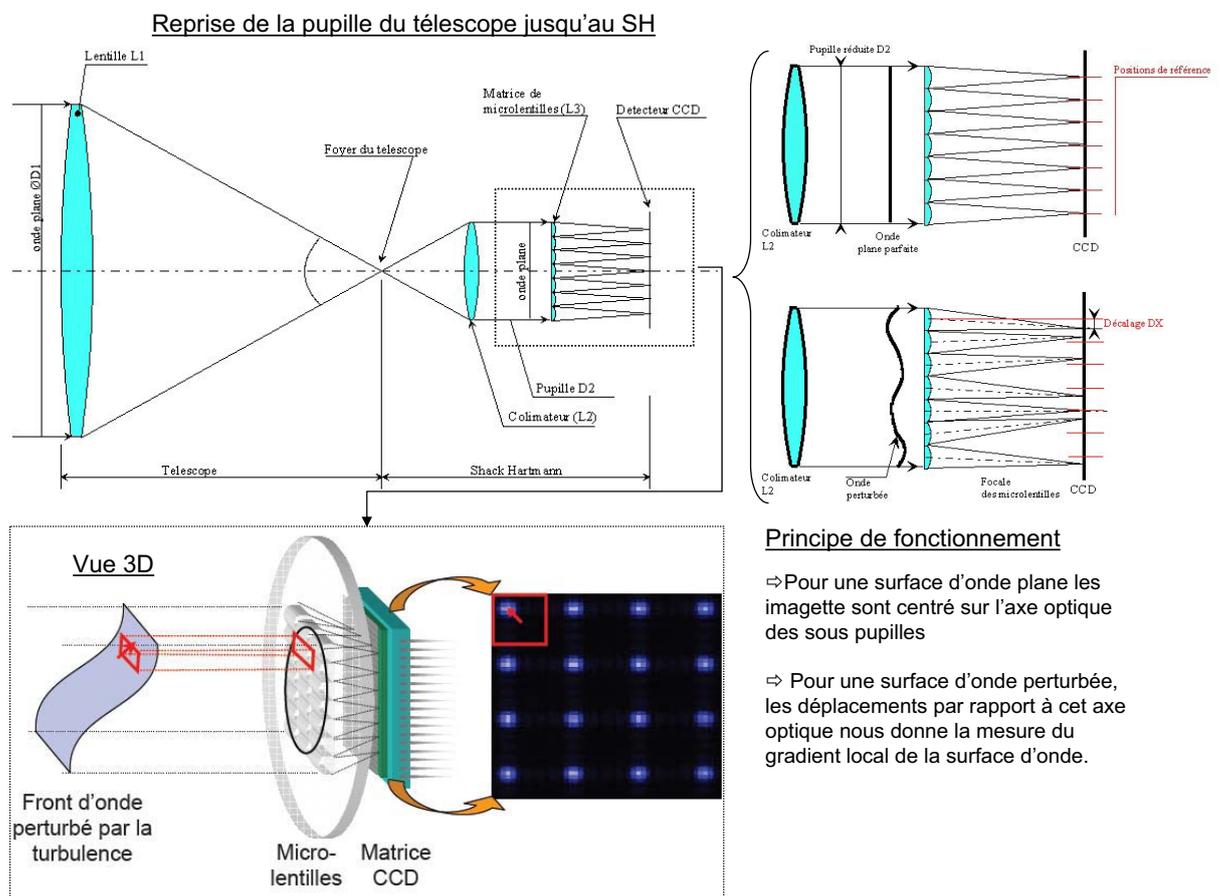


FIGURE 1.18 – Schéma de principe de fonctionnement d'un SH.

la phase sur la sous-pupille (ou angle d'arrivée) est liée au déplacement du centre de gravité

(CDG) de la tache focale au foyer d'une sous-pupille par la relation suivante :

$$\alpha_x = \frac{x_c}{f_{ml}} = \frac{\lambda}{2\pi\mathcal{S}} \iint_{sspup} \frac{\delta\varphi}{\delta x} dx dy$$

$$\alpha_y = \frac{y_c}{f_{ml}} = \frac{\lambda}{2\pi\mathcal{S}} \iint_{sspup} \frac{\delta\varphi}{\delta y} dx dy$$

avec \mathcal{S} la surface de la sous pupille, f_{ml} la focale des micro-lentilles ¹ et (x_c, y_c) le déplacement du centre de gravité par rapport à la position de référence de la sous pupille.

Méthode de simulation Les images de chaque sous-pupille sont calculées une par une, puis assemblées pour former l'image globale donnée par la matrice de micro-lentilles. Les interférences normalement produites par les micro-lentilles sont ici négligées.

Nous calculons les pentes de trois manières différentes : un centre de gravité classique, un centre de gravité pondéré corrigé pour avoir un coefficient de linéarité unitaire [Nicolle, 2006, Thomas *et al.*, 2006]. Le centre de gravité est efficace en régime de bruit de photons et le centre de gravité pondéré est efficace en bruit de détecteur.

Il est aussi possible d'estimer les pentes avec un quad-cell, que nous n'utiliserons pas dans les simulations. C'est un ensemble de 4 pixels plus grands que la tache image. En regardant la répartition d'intensité dans les 4 pixels, on en déduit facilement le centre de gravité de la tache. Cette méthode est précise mais elle a un domaine de linéarité très faible.

Modèle de bruit Les variances d'erreur sur la mesure de la pente peut être retrouvée grâce aux formules théoriques décrites par [Nicolle *et al.*, 2004] (centre de gravité) et [Nicolle, 2006, Thomas *et al.*, 2006] (centre de gravité pondéré corrigé) :

$$\text{Centre de gravité : } \sigma_{CdG_{ph}}^2 = \frac{\pi^2}{2 \ln(2) n_{ph}} \left(\frac{N_t}{N_d} \right)^2 \quad (1.33)$$

$$\sigma_{CdG_{det}}^2 = \frac{\pi^2}{3} \frac{\sigma_e^2}{n_{ph}^2} \left(\frac{N_s^2}{N_d} \right)^2 \quad (1.34)$$

$$\text{Centre de gravité pondéré : } \sigma_{CdGP_{ph}}^2 = \frac{\pi^2}{2 \ln(2) n_{ph}} \left(\frac{N_t}{N_d} \right)^2 \frac{(N_t^2 + N_w^2)^4}{(2N_t^2 + N_w^2)^2 N_w^4} \quad (1.35)$$

$$\sigma_{CdGP_{det}}^2 = \frac{\pi^3}{32(\ln 2)^2} \frac{\sigma_e^2}{n_{ph}^2} \frac{(N_t^2 + N_w^2)^4}{N_d^2 N_w^4} \quad (1.36)$$

$$(1.37)$$

avec n_{ph} le flux, σ_e^2 la variance du bruit de détecteur, N_t la largeur à mi-hauteur de la tache sur laquelle on fait la mesure, N_d la largeur à mi-hauteur de la tache théorique donnée par la diffraction, N_s^2 le nombre de pixels pris en compte pour la mesure du centre de gravité et N_w la largeur à mi-hauteur de la pondération utilisée pour le centre de gravité pondéré. Ces variances sont données pour la mesure d'un écart de phase bord à bord d'une sous-pupille et s'expriment en radians².

1. On règle la sensibilité d'un Shack-Hartmann en jouant sur la focale f_{ml} des micro-lentilles

Limitations pour la mesure des bas ordres sur étoile naturelle Pour l'estimation du tip/tilt et du focus, le Shack-Hartmann nécessite au moins 2×2 sous-pupilles. Par rapport à une tache pleine pupille, on divise donc le flux en 4 et on a une taille de tache dans chaque sous-pupille deux fois plus élevée. Le rapport signal-à-bruit par pixel est alors fortement diminué (cf §1.4.8 p.61). En particulier, le Shack-Hartmann perd surtout en précision dans les bas ordres [Rigaut et Gendron, 1992]. Cette solution n'est donc pas bien adaptée à l'analyse du tip/tilt et du focus sur étoile naturelle.

1.4.1.2 Pyramide

L'analyseur à pyramide est un analyseur plan pupille.

Principe Cet analyseur s'inspire de l'expérience du couteau de Foucault. Un prisme pyramidal coupe le plan focal en 4. Les faisceaux en sortie sont ensuite détectés dans un plan pupille (Fig. 1.19). Les 4 images de pupille résultantes sont assimilables à ce que l'on obtiendrait en utilisant un couteau de Foucault dans deux directions. La pyramide est généralement utilisée avec une modulation : on fait tourner la tache autour du sommet de la pyramide. On peut alors, en combinant les 4 images de pupille, remonter aux pentes locales du front d'onde, comme dans un Shack-Hartmann [Ragazzoni, 1996].

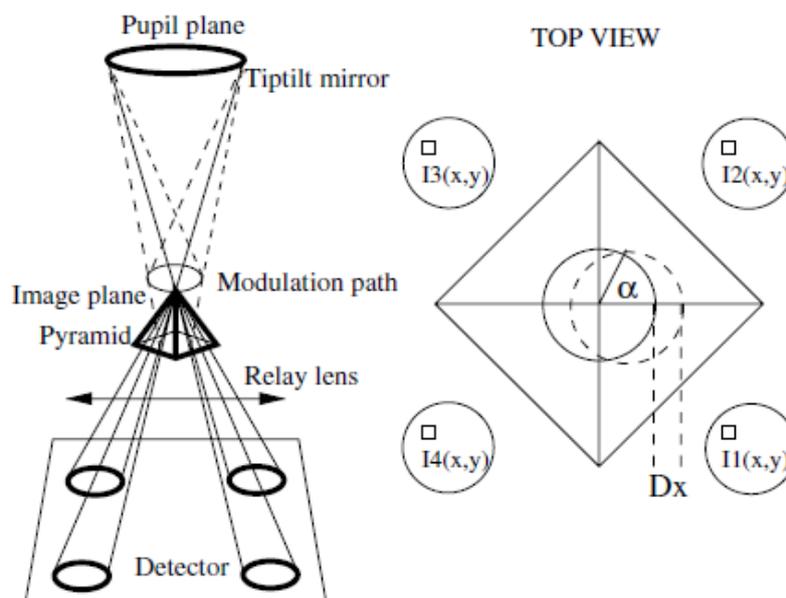


FIGURE 1.19 – Schéma de principe de l'analyseur à pyramide. Le sommet de la pyramide est placé au foyer, coupant le plan focal en 4. Les 4 faisceaux en sortie forment autant d'images de la pupille.

Source : [Vérinaud, 2004]

Méthode de simulation Nous calculons tout d'abord le champ complexe au plan focal. On lui applique alors le masque de phase correspondant à la pyramide, puis on calcule l'intensité

résultante dans le plan pupille. L'image ainsi obtenue inclut les interférences entre les 4 images de pupille, qui sont, dans notre cas, espacées entre elles de deux diamètres pupillaires. La pyramide peut être simulée sans modulation ou avec une modulation circulaire. Dans ce dernier cas, le cercle de modulation est échantillonné en plusieurs points d'intégration et une image est calculée pour chaque point. La somme de toutes les images obtenues forme l'image finale. Pour que la modulation soit bien échantillonnée, le nombre de points d'intégration est choisi de manière à ce qu'il y ait au maximum un diamètre de tache d'Airy entre chaque point.

Le signal issu de la pyramide est calculé de la manière suivante pour un pixel k :

$$\begin{aligned} S_x[k] &= \frac{(P_1[k] + P_3[k]) - (P_2[k] + P_4[k])}{N} \\ S_y[k] &= \frac{(P_1[k] + P_2[k]) - (P_3[k] + P_4[k])}{N} \\ \text{avec } N &= \frac{1}{n_{sspup}} \sum_k P_1[k] + P_2[k] + P_3[k] + P_4[k] \end{aligned} \quad (1.38)$$

avec P_i l'image de la pupille issue de la face i et n_{sspup} le nombre de sous-pupilles. S_x est le signal lié aux pentes en x et S_y le signal lié aux pentes en y . N est le flux par sous-pupille moyenné sur l'ensemble des sous-pupilles.

Modèle de bruit On considère que le nombre de sous-pupilles est suffisamment élevé pour que la variance de bruit sur N soit négligeable.

En bruit de photons, la variance de bruit sur le pixel $P_i[k]$ est égale à son flux moyen. La variance d'erreur du numérateur est alors la somme des flux moyens de $P_1[k]$, $P_2[k]$, $P_3[k]$ et $P_4[k]$, qui est égal au flux moyen de $P_1[k] + P_2[k] + P_3[k] + P_4[k]$.

En bruit de détecteur, la variance de bruit sur le pixel $P_i[k]$ est égale à σ_e^2 , avec σ_e^2 la variance de bruit de détecteur. La variance d'erreur du numérateur est alors $4\sigma_e^2$.

Finalement, la variance d'erreur sur $S_x[k]$ et $S_y[k]$ est $(P_1[k] + P_2[k] + P_3[k] + P_4[k])/N^2$ en bruit de photons et $4\sigma_e^2/N^2$ en bruit de détecteur.

Limitations pour la mesure des bas ordres sur étoile naturelle L'avantage de la pyramide par rapport au Shack-Hartmann est que le signal obtenu dépend du champ complexe au foyer de la pleine pupille. On bénéficie alors en partie de ce qu'on appelle le "gain pleine pupille", dont on reparlera dans la partie 1.4.8 p.61, et on obtient une meilleure performance qu'avec le Shack-Hartmann [Ragazzoni et Farinato, 1999, Vérinaud, 2004, Guyon, 2005]. La pyramide est donc a priori adaptée à la mesure des bas ordres à faible flux. Cependant, c'est un analyseur complexe à utiliser, puisqu'il requiert des composants optiques spécifiques et une parfaite maîtrise des alignements optiques.

1.4.1.3 Analyseur de courbure

L'analyseur de courbure est un analyseur plan pupille.

Principe Il repose sur le fait que les variations de phase se transforment en variations d'intensité lors de la propagation d'un champ. Roddier a proposé un analyseur dont le schéma est

présenté en Fig. 1.20 [Roddiier, 1988]. Une première lentille focalise le champ incident, puis une deuxième, d'une longueur focale deux fois moindre, est placée dans le plan focal de la première lentille. Cette deuxième lentille permet d'assurer la conjugaison entre deux plans symétriques par rapport au foyer. Roddiier a démontré qu'en prenant deux images à une distance l de part et d'autre du foyer et en les combinant, on peut obtenir un signal proportionnel à la courbure de la phase. Il faut cependant noter que sur les bords de la pupille, ce signal n'est plus proportionnel à la courbure mais à la dérivée première de la phase. Il faut aussi noter que cette linéarité n'est conservée que si la relation $\lambda(f - l)/d_0 \ll d_0 l/f$, avec f la focale de la première lentille et d_0 le rayon de cohérence de la phase, est vérifiée. $\lambda(f - l)/d_0$ correspond à la taille du flou résultant de la diffraction par un élément de taille d_0 , et $d_0 l/f$ est la taille des fluctuations que l'on veut observer.

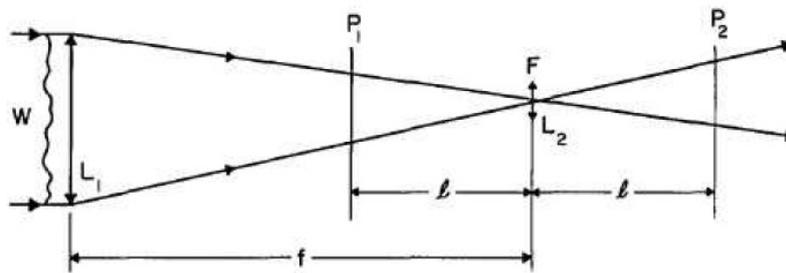


FIGURE 1.20 – Schéma de principe de l'analyseur de courbure. La focale de la lentille L_2 vaut $f/2$. En combinant les deux images obtenues dans les plans P_1 et P_2 , on obtient un signal proportionnel à la courbure de la phase. En pratique, la lentille L_2 est remplacée par une membrane vibrante, permettant ainsi de faire les deux images successivement sur un même détecteur.

Source : [Roddiier, 1988]

En pratique, il est difficile de placer un détecteur dans le plan P_1 . La lentille L_2 est donc remplacée par une membrane réfléchissante vibrante, qui permet d'acquérir tour à tour les deux images sur un même détecteur à haute cadence. Les détecteurs utilisés actuellement pour l'analyseur de courbure sont des matrices de photodiodes à avalanche disposées en couronnes. Chaque photodiode correspond à une sous-pupille de l'analyseur.

Méthode de simulation La propagation d'une onde sphérique sur une certaine distance est équivalente à la propagation d'une onde plane sur une distance plus grande. Ainsi, les intensités dans les deux plans extra-pupillaires, qui sont situés dans un faisceau convergent et un faisceau divergent, peuvent être calculées en propageant l'onde incidente dans l'espace libre sur une distance équivalente. Pour un plan situé à une distance l du foyer, la distance équivalente pour la propagation dans l'espace libre vaut $\pm f^2/l$, f étant la longueur focale.

En sachant cela, on est alors capable de calculer les intensités dans les plans extra-pupillaires en appliquant une propagation de Fresnel dans l'espace libre. On construit ensuite un masque qui permet de représenter la matrice de photodiodes à avalanches, qui est arrangée en couronnes (Fig. 1.21). L'intensité sur chaque détecteur est alors intégrée pour former les données.

Un signal peut être calculé à partir des données brutes pour déduire la courbure du front d'onde [Roddiier, 1988], mais nous ne nous y intéresserons pas dans la thèse.

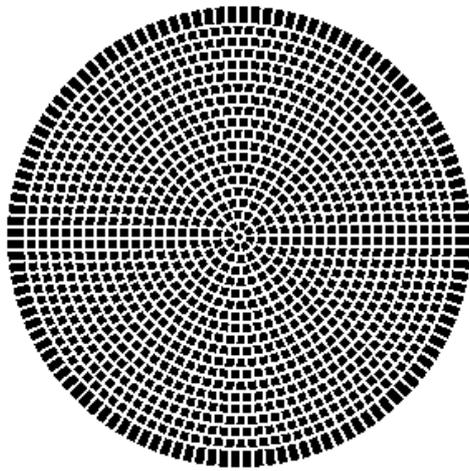


FIGURE 1.21 – Masque représentant la matrice de photodiodes à avalanches qui constitue le détecteur de l'analyseur de courbure. Le masque correspondant à une photodiode peut être isolé afin d'intégrer le flux sur cette photodiode.

Modèle de bruit Les données étant directement le flux sur les photodiodes, le modèle de bruit est celui d'une image classique.

Limitations pour la mesure des bas ordres sur étoile naturelle L'analyseur de courbure a l'avantage d'utiliser des photodiodes à avalanche, qui ne présentent pas de bruit de détecteur. Néanmoins, en bruit de photons, il est moins performant que le Shack-Hartmann pour estimer des bas ordres [Guyon, 2005]. Une variante permettant de mesurer les pentes et les courbures locales simultanément a été proposée par Paterson [Paterson et Dainty, 2000], mais elle ne dépasse pas la performance d'un Shack-Hartmann à faible flux [Barwick, 2009].

1.4.1.4 Masque de Zernike

Le masque de Zernike est un analyseur plan pupille.

Principe Le masque de Zernike est une pastille de taille λ/D , avec D le diamètre de la pupille, qui introduit un déphasage de $\pi/2$ et se place au foyer [Zernike, 1934]. Lorsque l'on fait ensuite une image en plan pupille, on obtient un signal proportionnel à la phase en entrée, à une constante près.

Cet effet est valable lorsque la phase est faible. Le champ incident s'exprime alors :

$$\Psi_P = \mathbf{P} \exp(i\varphi) \simeq \mathbf{P}(1 + i\varphi) \quad (1.39)$$

Le champ au plan focal s'écrit :

$$\Psi_F = \tilde{\mathbf{P}} + i\tilde{\mathbf{P}} * \tilde{\varphi} \quad (1.40)$$

où $\tilde{}$ désigne la transformée de Fourier et $*$ est le produit de convolution. Le premier terme correspond à la tache de diffraction, de largeur à mi-hauteur λ/D , tandis que le deuxième terme correspond au champ induit par la phase aberrante. On considère que ce dernier a une amplitude nulle dans un rayon inférieur à λ/D . En effet, une période spatiale l dans la phase se traduit par des speckles éloignés du centre de la tache d'une distance λ/l . Les plus grandes variations spatiales de φ sont limitées au diamètre D , le deuxième terme ne présente donc aucun signal dans un rayon inférieur à λ/D .

Le masque ajoute alors une phase $\pi/2$ uniquement sur le premier terme. Le champ en plan focal devient :

$$\Psi'_F = \tilde{\mathbf{P}} \exp(i\pi/2) + i\tilde{\mathbf{P}} * \tilde{\varphi} = i(\tilde{\mathbf{P}} + \tilde{\mathbf{P}} * \tilde{\varphi}) \quad (1.41)$$

En revenant dans le plan pupille, on a :

$$\widetilde{\Psi}'_F = i(\mathbf{P} + \mathbf{P}\varphi) \quad (1.42)$$

Et l'image s'écrit :

$$\mathbf{I}(\varphi) = |\widetilde{\Psi}'_F|^2 = \mathbf{P}^2(1 + \varphi)^2 \quad (1.43)$$

Comme on est en approximation faible phase, on néglige le terme en φ^2 pour finalement obtenir :

$$\mathbf{I}(\varphi) = \mathbf{P}^2(1 + 2\varphi) \quad (1.44)$$

Une autre version de cet analyseur, le masque de Zernike différentiel, fait une image avec un déphasage $\pi/2$ et une autre avec un déphasage $-\pi/2$, qui va avoir une intensité en $\mathbf{P}^2(1 - 2\varphi)$. La différence des deux images est alors directement proportionnelle à la phase.

Méthode de simulation Pour modéliser le masque, on crée un masque de phase circulaire dont le diamètre est égal à la largeur à mi-hauteur de la tache d’Airy. La phase vaut $\pi/2$ à l’intérieur du cercle et 0 à l’extérieur. En appliquant ce masque au champ complexe en plan focal, on peut alors calculer l’image résultante en plan pupille.

Modèle de bruit Le modèle de bruit est celui d’une image classique.

Limitations pour la mesure des bas ordres sur étoile naturelle Cet analyseur est en théorie idéal puisque le signal est directement la phase à estimer [Paterson, 2008, Guyon, 2005]. Cependant, il ne permet d’estimer que des phases très faibles, et il ne peut pas être utilisé en large bande spectrale, à moins d’utiliser des moyens spécifiques tel qu’un correcteur de Wynne [Wynne, 1979].

1.4.1.5 Diversité de phase

La diversité de phase est une méthode d’analyse plan focal.

Principe Dans un système parfait, la FEP d’un point source est une tache d’Airy. Cependant, les aberrations du front d’onde ont une influence sur la forme et la taille de la FEP. En effet la FEP est liée au front d’onde par la formule 1.8. Cependant, une estimation directe de la phase est difficile car elle peut présenter des ambiguïtés. Les modes pairs donnent en effet la même image quel que soit le signe de leur amplitude. La diversité de phase consiste à prendre une autre FEP avec un déphasage connu ϕ_d , appelé phase de diversité [Gonsalves, 1982, Paxman *et al.*, 1992]. On obtient alors la combinaison de formules :

$$FEP_1(\phi) = |TF\{P \cdot e^{i\phi}\}|^2 \quad (1.45)$$

$$FEP_2(\phi) = |TF\{P \cdot e^{i(\phi+\phi_d)}\}|^2 \quad (1.46)$$

Avec ces deux images, il est possible de retrouver la phase sans ambiguïté. La défocalisation est la phase de diversité habituellement utilisée. C’est en effet une aberration facilement applicable, et qui permet de lever l’ambiguïté sur les modes pairs.

Méthode de simulation Les données sont très faciles à simuler, puisqu’il suffit de créer une image en plan focal et une image dans un plan extra-focal.

Modèle de bruit La diversité de phase s’appuie directement sur des images classiques. Le modèle de bruit utilisé est donc celui décrit au paragraphe 1.2.1.1 p.22.

Limitations pour la mesure des bas ordres sur étoile naturelle On considère souvent que les inconvénients de cette méthode sont qu’elle requiert un traitement complexe, associé à un coup de calcul peu compatible avec des utilisations en temps réel, et qu’elle ne peut pas être utilisée facilement en large bande spectrale [Rousset, 1999]. De plus, elle nécessite une division du flux par deux pour acquérir les deux images. Cependant cette technique est en pleine mutation et des algorithmes sont actuellement en développement, notamment à l’ONERA, pour améliorer et accélérer le traitement en vue de l’appliquer en bande large, en temps réel [Mocœur *et al.*, 2009]. Un traitement sur une image unique a aussi été envisagé [Tokovinin et Heathcote, 2006].

1.4.2 LIFT

La plupart des analyseurs cités ci-dessus ne présentent pas les bonnes propriétés pour analyser les bas ordres (tip/tilt et focus) sur une étoile naturelle faible dans un système d'optique adaptative assisté par laser. LIFT (LInearized Focal-plane Technique) est un analyseur qui a été développé dans ce but.

1.4.2.1 Historique

L'idée de LIFT a été inspirée par plusieurs concepts. On peut tout d'abord citer le *Modified Quad-Cell*, proposé par Dekany pour le télescope Keck afin d'analyser le tip/tilt et focus sur étoile naturelle. Cet analyseur consiste à ajouter un astigmatisme sur la phase pour rendre le quad-cell sensible à une défocalisation. En effet, sans astigmatisme, la tache est identique si l'on défocalise dans un sens ou dans l'autre avec la même amplitude absolue, et on mesure un signal nul sur le quad-cell. Si l'on ajoute l'astigmatisme, la tache défocalisée s'allonge dans une direction ou dans la direction perpendiculaire selon le signe du focus (Fig. 1.22). Le modified quad-cell permet alors, avec la pondération adaptée, de mesurer un focus. Cependant, cet analyseur a un domaine de linéarité très limité et il est très sensible au repliement.

Par ailleurs, Tokovinin a proposé une méthode de diversité de phase qui n'utilise que l'image défocalisée [Tokovinin et Heathcote, 2006]. Il montre notamment qu'un focus ou un astigmatisme peuvent facilement et rapidement être estimés à partir de l'élongation de la tache image en x et y . Néanmoins, la défocalisation utilisée est importante, et la tache résultante est très étalée sur les pixels. Le rapport signal-à-bruit n'est donc sans doute pas optimisé pour l'estimation de ces modes.

LIFT est un analyseur plan focal qui reprend l'idée de l'astigmatisme pour estimer le focus, ainsi que l'utilisation d'une image unique, présentant une phase de diversité.

1.4.2.2 Principe

LIFT utilise une phase de diversité ϕ_d égale à 0,5 radians en astigmatisme (coefficient de Zernike $a_5=0,5$). Cette phase résout le problème d'indétermination pour les modes pairs, dont le focus fait partie (Fig. 1.22). L'estimation, de type maximum de vraisemblance, est faite sur une image pleine pupille unique.

En considérant \mathbf{r} le vecteur des coordonnées spatiales dans le plan focal, ϕ la phase à mesurer et n le bruit, l'image voit son intensité exprimée d'une manière semblable à l'équation 1.8 :

$$I(\mathbf{r}, \phi) = |TF\{\underbrace{P e^{i\phi_d}}_{P_d} \cdot e^{i\phi}\}_{\mathbf{r}}|^2 + n(\mathbf{r}) \quad (1.47)$$

On peut décomposer ϕ sur la base des polynômes de Zernike Z_i tel que $\phi = \sum_i a_i Z_i$. On note \mathbf{A} le vecteur des coefficients a_i , et I désigne la fonction de \mathbf{A} donnant l'intensité dans le plan focal. $\mathbf{I}(\mathbf{A})$ est un vecteur dont l'élément $\mathbf{I}(\mathbf{A})[p]$ correspond au p -ième pixel de l'image (selon l'ordre lexicographique). Le développement de Taylor au premier ordre de ce vecteur donne :

$$\mathbf{I}(\mathbf{A}) - \mathbf{I}(\mathbf{0}) \simeq \sum_k a_k \mathbf{I}'_k(\mathbf{0}) + \mathbf{n} \quad (1.48)$$

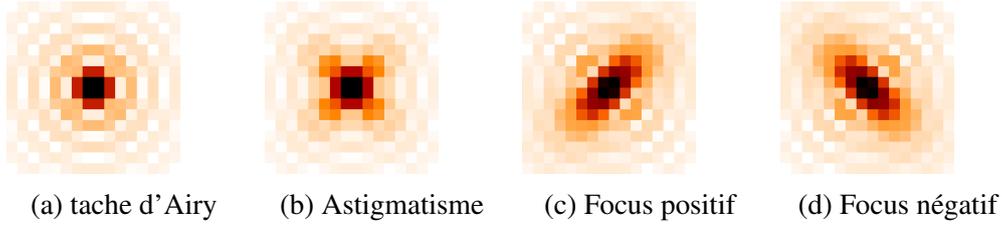


FIGURE 1.22 – Levée de l'indétermination du focus grâce à l'astigmatisme. La tache s'étire dans une direction ou dans la direction perpendiculaire selon le signe du focus. Au plan focal, l'énergie est assez peu étalée par rapport à la tache d'Airy.

avec $\mathbf{I}'_k(\mathbf{0}) = \left. \frac{\partial \mathbf{I}(\mathbf{A})}{\partial a_k} \right|_{\mathbf{A}=\mathbf{0}}$ et \mathbf{n} le vecteur de bruit dans l'image. On peut montrer que $\mathbf{I}'_k(\mathbf{0}) = 2\Re\{\widetilde{\mathbf{P}}_d^* \cdot iTF\{\mathbf{P}_d \cdot \mathbf{Z}_k\}\}$ (cf annexe A.1 p.189), où l'opérateur \cdot est le produit terme à terme et $\widetilde{\mathbf{P}}_d^*$ le conjugué de la transformée de Fourier de \mathbf{P}_d . Les vecteurs \mathbf{P}_d et \mathbf{Z}_k représentent respectivement l'amplitude complexe de diversité et le polynôme de Zernike d'ordre k dans la pupille. \mathbf{n} , \mathbf{P}_d et \mathbf{Z}_k ont leurs éléments définis de la même manière que $\mathbf{I}(\mathbf{A})$.

En notant le vecteur $\Delta\mathbf{I}$ tel que $\Delta\mathbf{I}[p] \triangleq \mathbf{I}(\mathbf{A})[p] - \mathbf{I}(\mathbf{0})[p]$, et la matrice \mathbf{H} telle que l'élément à la p -ième ligne et la k -ième colonne s'exprime $\mathbf{H}[p, k] \triangleq \mathbf{I}'_k[p]$, on a alors un modèle linéaire :

$$\Delta\mathbf{I} = \mathbf{H}(\mathbf{0})\mathbf{A} + \mathbf{n} \quad (1.49)$$

où l'on cherche \mathbf{A} . Pour pouvoir l'estimer, il va falloir prendre en compte le bruit \mathbf{n} . On suppose ici qu'il s'agit d'un bruit gaussien de moyenne nulle (cf §1.2.1.1 p.22). On appellera sa matrice de covariance $\mathbf{R}_n = \langle \mathbf{n}\mathbf{n}^t \rangle$ où \mathbf{n}^t est la transposée de \mathbf{n} et $\langle \cdot \rangle$ est une moyenne sur les occurrences de bruit. L'estimateur donné par le maximum de vraisemblance est le suivant (cf annexe A.2 p.189) :

$$\hat{\mathbf{A}}_{MV} = P_{MV}(\mathbf{0})\Delta\mathbf{I}; \quad P_{MV}(\mathbf{0}) = (\mathbf{H}(\mathbf{0})^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{0}))^{-1} \mathbf{H}(\mathbf{0})^t \mathbf{R}_n^{-1} \quad (1.50)$$

L'estimation revient finalement à appliquer une pondération $P_{MV}(\mathbf{0})$ à la différence d'images $\Delta\mathbf{I}$. $P_{MV}(\mathbf{0})$ contient une carte de pondération par mode. La pondération s'adapte aux conditions de bruit. A fort flux, le bruit de détecteur a une influence négligeable, donc la pondération prend tous les pixels utiles en compte. A faible flux, le bruit de détecteur devient dominant, la pondération évite alors de prendre en compte les pixels en périphérie de la tache, qui sont noyés dans le bruit (Fig. 1.23). L'équation 1.48 n'utilise le développement de Taylor qu'au premier ordre, ce qui limite le domaine de linéarité de LIFT autour de 0. Pour augmenter le domaine de linéarité, on itère plusieurs fois à partir de l'estimée donnée par l'équation 1.50 (voir §2.2 p.66).

Comme on estime peu de modes, on peut être affecté par une erreur induite par le repliement des hauts ordres. L'impact des hauts ordres sur la FEP se manifeste assez loin de l'axe optique, on peut donc limiter le repliement en appliquant l'algorithme sur une zone restreinte de l'image. Ce fenêtrage numérique est analogue au concept de Shack-Hartmann filtré par un diaphragme de champ dans un plan focal intermédiaire, introduit par Lisa Poyneer [Poyneer et Macintosh, 2004]. Cet aspect sera détaillé dans le chapitre 3 p.95.

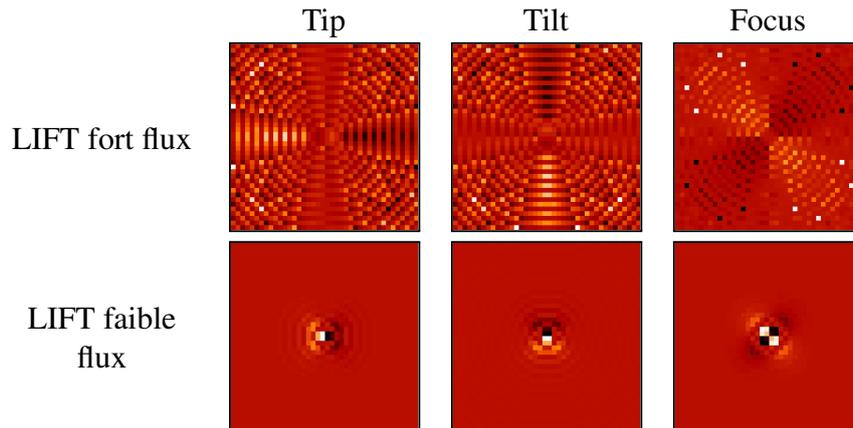


FIGURE 1.23 – Cartes de pondération pour LIFT en fort flux et faible flux.

Dans la suite de la thèse, nous serons amenés à comparer LIFT à d’autres analyseurs. Nous présentons dans la partie suivante les notions nécessaires à la comparaison des analyseurs de front d’onde.

1.4.3 Comparaison des analyseurs de front d’onde

La comparaison des analyseurs a une place importante dans la littérature, puisque la performance d’une optique adaptative est en partie limitée par la performance de l’analyseur de front d’onde. Les analyseurs peuvent être notamment comparés en termes de dynamique ou de propagation du bruit. La dynamique correspond à l’amplitude maximale mesurable pour un mode donné, et la propagation du bruit décrit l’impact du bruit de mesure sur l’estimation. Pour une application donnée, la dynamique, ainsi que le nombre de modes à estimer, sont fixés. L’approche classique est alors de comparer les analyseurs en terme de propagation de bruit, en évaluant la matrice de covariance d’erreur d’estimation, qui détermine l’erreur sur la reconstruction du front d’onde [Rousset, 1999, Rigaut et Gendron, 1992, Guyon, 2005, Vérinaud, 2004, Ragazzoni et Farinato, 1999, Ellerbroek *et al.*, 1997, Welsh *et al.*, 1995, Meynadier *et al.*, 1999]². Cependant, la matrice de covariance d’erreur dépend de la méthode d’estimation pour un analyseur donné.

Afin d’avoir une juste comparaison des analyseurs, nous nous sommes appuyé dans cette thèse sur l’information qui se trouve dans les données, qui peut être quantifiée grâce à la *matrice d’information de Fisher* [Van Trees, 2004]. Dans la suite, nous allons montrer que cette approche est en fait équivalente à la précédente dans le cas particulier d’une estimation de type maximum de vraisemblance, avec l’hypothèse d’un bruit additif gaussien. Ce type d’estimation est couramment utilisé, par exemple pour reconstruire le front d’onde à partir des pentes d’un Shack-Hartmann. Nous privilégions tout de même l’approche de l’information de Fisher, qui est plus générale.

Dans un premier temps, nous décrivons le calcul de la propagation du bruit pour une estimation du front d’onde de type maximum de vraisemblance (§1.4.4 p.59). Nous définissons la covariance d’erreur d’estimation, qui caractérise la précision de l’estimation, à partir de la-

2. Ces références feront l’objet d’une étude bibliographique dans le chapitre 5 p.149.

quelle nous déduisons des coefficients pour quantifier facilement la propagation du bruit d'un analyseur (§1.4.5 p.60).

Nous définissons ensuite la matrice d'information de Fisher (§1.4.6 p.60), puis nous montrons le lien entre la propagation de bruit donnée par l'information de Fisher et celle donnée par une estimation de maximum de vraisemblance (§1.4.7 p.61).

Enfin, nous nous intéressons à la notion de gain pleine pupille, qui permet de mieux appréhender la physique qui se cache derrière la propagation du bruit (§1.4.8 p.61).

1.4.4 Propagation de bruit dans un estimateur Maximum de Vraisemblance

Un analyseur de front d'onde, dans son domaine de linéarité, peut être représenté par un modèle linéaire. Nous prenons l'exemple d'une estimation de coefficients de Zernike³ avec un Shack-Hartmann. On décrit le modèle linéaire par l'équation suivante :

$$\mathbf{y} = \mathbf{D}\mathbf{A} + \mathbf{b} \quad (1.51)$$

avec \mathbf{y} le vecteur des pentes, \mathbf{D} la matrice d'interaction, \mathbf{A} le vecteur des coefficients de Zernike et \mathbf{b} le bruit sur les pentes.

La solution donnée par l'estimateur maximum de vraisemblance est :

$$\hat{\mathbf{A}} = (\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{y} \quad (1.52)$$

avec $\mathbf{C}_b = \langle \mathbf{b}\mathbf{b}^t \rangle$ la matrice de covariance du bruit. Cette solution existe si la densité de probabilité de vraisemblance $p(\mathbf{y}|\mathbf{A})$ (probabilité d'avoir \mathbf{y} sachant \mathbf{A}) admet un maximum. En outre, pour appliquer cette solution analytique, la matrice $(\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1}$ doit être inversible. Cette condition ne peut être remplie que si le nombre de données est supérieur ou égal au nombre de coefficients à estimer (taille de $\mathbf{y} \geq$ taille de \mathbf{A}).

L'erreur d'estimation vaut alors :

$$\mathbf{E} = \hat{\mathbf{A}} - \mathbf{A} = (\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{b} \quad (1.53)$$

On peut ainsi calculer la matrice de covariance d'erreur d'estimation \mathbf{C}_E :

$$\begin{aligned} \mathbf{C}_E &= \langle \mathbf{E}\mathbf{E}^t \rangle = \langle (\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{b}\mathbf{b}^t \mathbf{C}_b^{-1} \mathbf{D} (\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1} \rangle \\ \mathbf{C}_E &= (\mathbf{D}^t \mathbf{C}_b^{-1} \mathbf{D})^{-1} \end{aligned} \quad (1.54)$$

La variance d'erreur sur chaque mode est égale aux coefficients diagonaux de cette matrice. La trace (somme des coefficients diagonaux) donne alors la variance d'erreur totale sur le front d'onde. En tenant compte du fait que le bruit est homogène et décorréolé sur toutes les pentes, on retrouve les coefficients de Rigaut et Gendron, définis par la diagonale de $(\mathbf{D}^t \mathbf{D})^{-1}$ [Rigaut et Gendron, 1992]. On a en effet dans ce cas $\mathbf{C}_b = \sigma_{pente}^2 \times Id$, donc la variance d'erreur sur chaque mode vaut $\sigma_i^2 = [(\mathbf{D}^t \mathbf{D})^{-1}]_{i,i} \times \sigma_{pente}^2$.

3. Les calculs restent bien sûr valides pour d'autres modes estimés.

1.4.5 Coefficients de propagation du bruit

La variance d'erreur d'estimation d'un analyseur peut s'exprimer comme la somme d'un terme représentant la propagation du bruit de photons et d'un terme représentant la propagation du bruit de détecteur :

$$\sum_i \sigma^2(\hat{a}_i - a_i) = \overbrace{\left(\sum_i \alpha_i\right)}^{\alpha} \frac{1}{n_{ph}} + \overbrace{\left(\sum_i \beta_i\right)}^{\beta} \left(\frac{\sigma_e}{n_{ph}}\right)^2 \quad (1.55)$$

avec n_{ph} le flux et σ_e l'écart-type du bruit de détecteur. α_i et β_i sont respectivement les coefficients de propagation du bruit de photons et du bruit de détecteur pour le i -ème mode. α et β sont les coefficients de propagation sur l'estimation totale. Plus ces coefficients sont petits, moins la variance d'erreur est élevée. On veut donc qu'ils soient minimaux.

Il est possible de retrouver ces coefficients théoriquement en calculant la matrice de covariance d'erreur d'estimation \mathbf{C}_E :

- en annulant le bruit de détecteur. On a alors $\alpha_i = \mathbf{C}_{Ei,i} \times n_{ph}$.
- en annulant le bruit de photons et en fixant l'écart-type de bruit de détecteur à 1 électron. On a alors $\beta_i = \mathbf{C}_{Ei,i} \times n_{ph}^2$.

Dans la thèse, nous nous appuyerons sur ces coefficients pour comparer les analyseurs.

1.4.6 Information de Fisher

La matrice d'information de Fisher est un outil classiquement utilisé en traitement du signal pour évaluer la quantité d'information présente dans les données (voir par exemple [Retout *et al.*, 2001, Vallisneri, 2008, Machado *et al.*, 2009, Barrett *et al.*, 1995]). A partir de cette matrice, on peut calculer une limite inférieure de la variance d'erreur d'un estimateur, appelée borne de Cramér-Rao [Cramér, 1999]. Je donne ici la définition de la matrice de Fisher et je fais le lien avec la borne de Cramér-Rao.

Soit $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ l'ensemble des données (pentes, pixels...), qui dépendent d'un ensemble de coefficients de modes (des polynômes de Zernike par exemple) $\mathbf{A} = \{a_1, a_2, \dots, a_m\}$ et d'un bruit $\mathbf{b} = \{b_1, b_2, \dots, b_n\}$. Les éléments de la matrice de Fisher sont définis de la manière suivante :

$$F_{i,j}(\mathbf{A}_p) = E\left\{\left[\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p} \left[\frac{\partial}{\partial a_j} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p}\right\} \quad (1.56)$$

avec $p(\mathbf{y}|\mathbf{A})$ la densité de probabilité de vraisemblance et \mathbf{A}_p le point de fonctionnement. La probabilité de vraisemblance est la probabilité, lorsque les coefficients \mathbf{A} sont connus, d'obtenir les données qui correspondent à ces coefficients⁴. On peut aussi écrire :

$$F_{i,j}(\mathbf{A}_p) = -E\left\{\frac{\partial^2}{\partial a_i \partial a_j} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p}\right\} \quad (1.57)$$

La matrice de Fisher est liée à la borne de Cramér-Rao, ou *Cramer-Rao Lower Bound* en anglais (CRLB) de la manière suivante :

$$CRLB(\hat{\mathbf{A}}) = F(\mathbf{A}_p)^{-1} \quad (1.58)$$

4. Notons que la statistique des coefficients de Zernike n'intervient pas dans cette probabilité, puisque les coefficients sont considérés comme étant connus.

Cette limite inférieure est valide pour tous les estimateurs non biaisés, c'est-à-dire dont la moyenne sur beaucoup d'occurrences de bruit tend vers la valeur vraie.

Il faut noter que, puisque \mathbf{A} est un vecteur de coefficients, la borne de Cramér-Rao est ici une matrice. Pour une matrice de covariance d'erreur d'estimation \mathbf{C}_E , on a donc la relation

$$\mathbf{C}_E \geq CRLB(\hat{\mathbf{A}}) \quad (1.59)$$

qui est équivalente à dire que la matrice $\mathbf{C}_E - CRLB(\hat{\mathbf{A}})$ est semi-définie positive.

On peut ainsi calculer des coefficients de propagation du bruit ne dépendant que de l'information contenue dans les données en remplaçant \mathbf{C}_E par $F(\mathbf{A}_p)^{-1}$ dans le paragraphe 1.4.5 p.60. Cette méthode sera utilisée dans la thèse pour comparer les analyseurs.

Il existe une version de la matrice de Fisher, dite "matrice d'information Bayésienne", qui prend en compte la statistique des modes à estimer et qui permet de déterminer la borne de Cramér-Rao pour des estimateurs biaisés [Dauwels et Korl, 2006, Bay *et al.*, 2008]. Nous avons choisi de ne pas prendre en compte la statistique des modes à estimer ici, afin d'avoir une comparaison indépendante du contexte.

Dans le paragraphe suivant, nous montrons le lien entre la matrice d'information de Fisher et la matrice de covariance de l'estimateur de vraisemblance définie au paragraphe 1.4.4 p.59.

1.4.7 Lien entre l'information de Fisher et l'estimateur de maximum de vraisemblance

Dans l'hypothèse d'un bruit additif gaussien, l'estimateur de maximum de vraisemblance atteint la borne de Cramér-Rao [Barrett *et al.*, 1995]. Ce résultat est redémontré dans l'annexe D.2 p.196 dans le cadre de l'estimation d'un front d'onde. Ainsi, l'approche de la matrice d'information de Fisher et celle de la matrice de covariance d'erreur d'estimation pour calculer la propagation du bruit sont bien équivalentes.

Dans le cas du Shack-Hartmann, cela implique que la borne de Cramér-Rao sur les pentes locales peut être directement déduite des coefficients de Rigaut et Gendron [Rigaut et Gendron, 1992], habituellement utilisés pour calculer la propagation de bruit dans un Shack-Hartmann.

Ces considérations sont très mathématiques et il peut être intéressant de revenir parfois à des concepts plus physiques. La notion de gain pleine pupille permet d'avoir un ordre de grandeur des différences de propagation de bruit entre deux types d'analyseurs. C'est l'objet du dernier paragraphe de ce chapitre.

1.4.8 Gain pleine pupille

On appelle "gain pleine pupille" le gain obtenu sur la propagation du bruit dans une estimation en utilisant la pleine pupille plutôt que de la diviser en plusieurs sous-pupilles. D'une part, le flux n'est pas divisé et d'autre part, la tache image donnée par la pleine pupille est plus étroite que la tache image donnée par une sous-pupille. On a alors un flux plus concentré et un meilleur rapport signal-à-bruit. Ce gain se perd lorsqu'on s'éloigne de la limite de diffraction. En effet, plus les aberrations sont fortes, plus la tache donnée par la pleine pupille est éclatée.

Nous illustrons ici ce gain en prenant l'exemple de l'estimation d'un centre de gravité pour retrouver la pente d'un front d'onde α (Fig. 1.24). Nous considérons que nous sommes à la limite de diffraction, et que l'on fait une estimation de la pente sur la pleine pupille. La mesure

du centre de gravité est directement proportionnelle à la pente du front d'onde. Cependant, on exprime généralement l'erreur sur la mesure de la pente en différence de phase entre les bords de la pupille. Cette différence de phase $\Delta\varphi$ est liée à α via la différence bord à bord d'écart aberrant du front d'onde $\Delta\delta$. On a $\Delta\delta = \Delta\varphi \times \lambda/2\pi$ et $\alpha = \Delta\delta/D$ avec D le diamètre de la pupille.

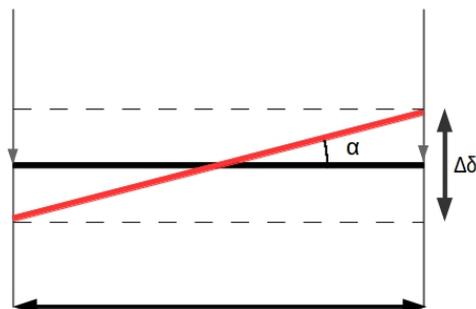


FIGURE 1.24 – Front d'onde tilté arrivant sur une lentille.

La variance d'erreur sur la différence de phase bord à bord de la pupille est égale à, en bruit de photons [Rousset, 1999] :

$$\sigma_{\Delta\varphi,ph}^2 = \frac{\pi^2}{2} \frac{1}{n_{ph}} \quad (1.60)$$

et en bruit de détecteur :

$$\sigma_{\Delta\varphi,det}^2 = \frac{\pi^2}{3} \frac{\sigma_e^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \quad (1.61)$$

avec n_{ph} le flux et σ_e^2 la variance de bruit de détecteur. La variance d'erreur en bruit de détecteur $\sigma_{\Delta\varphi,det}^2$ dépend de N_S^2 , le nombre de pixels utilisés pour l'estimation et de N_D , la largeur à mi-hauteur de la tache de diffraction. Sachant que $\alpha = \Delta\varphi \times \lambda/2\pi D$, on a $\sigma_\alpha^2 = (\lambda/2\pi D)^2 \sigma_{\Delta\varphi}^2$, donc :

$$\sigma_{\alpha,ph}^2 = \left(\frac{\lambda}{2\pi D} \right)^2 \frac{\pi^2}{2} \frac{1}{n_{ph}} = \frac{1}{8} \frac{1}{n_{ph}} \left(\frac{\lambda}{D} \right)^2 \quad (1.62)$$

et

$$\sigma_{\alpha,det}^2 = \left(\frac{\lambda}{2\pi D} \right)^2 \frac{\pi^2}{3} \frac{\sigma_e^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 = \frac{1}{12} \frac{\sigma_e^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{D} \right)^2 \quad (1.63)$$

Dans la suite, nous allons considérer que le flux est divisé de manière égale sur N_V voies d'imagerie. On estime un centre de gravité dans chaque voie, puis on moyenne ces estimations. La variance d'erreur sera alors divisée par le nombre d'estimations, soit N_V . Nous allons considérer deux cas de figure : une division simple du flux et une division du flux associée à une division en sous-pupilles.

1.4.8.1 Division simple du flux

La division simple du flux consisterait par exemple à mettre plusieurs lames séparatrices pour faire une image pleine pupille sur plusieurs détecteurs. Un tel cas n'aurait évidemment aucun intérêt en pratique, il s'agit seulement d'une illustration. La variance d'erreur d'estimation de la pente sur la voie i est, en bruit de photons :

$$\sigma_{\alpha,ph,i}^2 = \frac{1}{8} \frac{N_V}{n_{ph}} \left(\frac{\lambda}{D} \right)^2 \quad (1.64)$$

et en bruit de détecteur :

$$\sigma_{\alpha,det,i}^2 = \frac{1}{12} \frac{\sigma_e^2 N_V^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{D} \right)^2 \quad (1.65)$$

En moyennant les résultats, on divise la variance d'erreur par N_V , et on obtient :

$$\sigma_{\alpha_{moy},ph}^2 = \frac{1}{N_V} \sigma_{\alpha,ph,i}^2 = \frac{1}{8} \frac{1}{n_{ph}} \left(\frac{\lambda}{D} \right)^2 = \sigma_{\alpha,ph}^2 \quad (1.66)$$

et

$$\sigma_{\alpha_{moy},det}^2 = \frac{1}{N_V} \sigma_{\alpha,det,i}^2 = \frac{1}{12} \frac{\sigma_e^2 N_V}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{D} \right)^2 = N_V \sigma_{\alpha,det}^2 \quad (1.67)$$

La variance d'erreur reste la même en bruit de photons, mais on l'a multipliée par N_V en bruit de détecteur.

1.4.8.2 Division en sous-pupilles

Considérons maintenant que la pupille est divisée à l'aide d'une matrice de $N_{sspup} \times N_{sspup}$ micro-lentilles de taille $d = D/N_{sspup}$. Pour une pupille circulaire, le flux est divisé sur $N_V = \pi/4 \times N_{sspup}^2$ sous-pupilles. Dans chaque sous-pupille, la variance d'erreur d'estimation de la pente s'écrit alors, en bruit de photons :

$$\sigma_{\alpha,ph,i}^2 = \frac{1}{8} \frac{N_V}{n_{ph}} \left(\frac{\lambda}{d} \right)^2 = \frac{1}{8} \frac{N_V N_{sspup}^2}{n_{ph}} \left(\frac{\lambda}{D} \right)^2 \quad (1.68)$$

et en bruit de détecteur :

$$\sigma_{\alpha,det,i}^2 = \frac{1}{12} \frac{\sigma_e^2 N_V^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{d} \right)^2 = \frac{1}{12} \frac{\sigma_e^2 N_V^2 N_{sspup}^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{D} \right)^2 \quad (1.69)$$

En moyennant les résultats, on divise la variance d'erreur par N_V , et on obtient :

$$\sigma_{\alpha_{moy},ph}^2 = \frac{1}{N_V} \sigma_{\alpha,ph,i}^2 = \frac{1}{8} \frac{N_{sspup}^2}{n_{ph}} \left(\frac{\lambda}{D} \right)^2 = N_{sspup}^2 \sigma_{\alpha,ph}^2 \quad (1.70)$$

et

$$\sigma_{\alpha_{moy},det}^2 = \frac{1}{N_V} \sigma_{\alpha,det,i}^2 = \frac{1}{12} \frac{\sigma_e^2 N_V N_{sspup}^2}{n_{ph}^2} \left(\frac{N_S^2}{N_D} \right)^2 \left(\frac{\lambda}{D} \right)^2 = N_V N_{sspup}^2 \sigma_{\alpha,det}^2 \quad (1.71)$$

La variance d'erreur est finalement multipliée par N_{sspup}^2 en bruit de photons et par $N_V N_{sspup}^2 \sim N_{sspup}^4$ en bruit de détecteur. Par rapport à une division simple du flux, le facteur est de N_{sspup}^2 sur les deux bruits.

En conclusion, la sensibilité au bruit de photons n'est pas influencée par la division du flux, mais plutôt par l'utilisation de pupilles plus petites pour échantillonner la pupille. En termes de bruit de détecteur, la variance d'erreur est démultipliée dans tous les cas. Un gain important en précision peut donc être atteint en faisant tendre N_V vers 1.

1.4.8.3 Application au Shack-Hartmann et à la pyramide

Considérons maintenant un Shack-Hartmann et une pyramide modulée avec $N_{sspup} \times N_{sspup}$ sous-pupilles estimant un tip/tilt sur la pupille globale. Ces deux analyseurs estiment des pentes locales et reçoivent le même flux par sous-pupille. Cependant, les sous-pupilles de la pyramide, contrairement au Shack-Hartmann, ne sont pas des lentilles, mais sont représentées par les pixels sur le détecteur. Il n'y a ainsi pas de perte de résolution par rapport à la pleine pupille. Finalement, le rapport des propagations de bruit entre le Shack-Hartmann et la pyramide est égal au rapport entre l'erreur d'estimation des pentes pour une division de la pupille sur des micro-lentilles et l'erreur d'estimation des pentes pour une division simple du flux.

On a donc un gain en précision de la pyramide par rapport au Shack-Hartmann de l'ordre de N_{sspup}^2 . Ce gain est à modérer avec la modulation de la pyramide. En effet, la modulation entraîne un agrandissement artificiel de la tache focale, donc une perte de résolution. Si le diamètre de la modulation est de l'ordre de la taille d'une tache de sous-pupille du Shack-Hartmann, le gain pleine pupille est totalement perdu [Vérinaud, 2004]. D'autre part, si l'on estime des modes au-delà du tip/tilt, il faut s'intéresser à des parties de la tache qui sont de plus en plus loin du centre au fur et à mesure qu'on augmente l'ordre radial du mode à estimer. Cela est équivalent à une baisse de résolution, proportionnelle à l'ordre radial [Ragazzoni et Farinato, 1999]. Pour les modes les plus élevés, la pyramide finit donc par perdre totalement le gain pleine pupille.

Chapitre 2

Optimisation de LIFT et validation en laboratoire

2.1 Introduction

Dans le premier chapitre, nous avons évoqué la nécessité d'analyser les bas ordres sur une étoile naturelle dans un système d'optique adaptative assisté par laser (§1.3.4 p.39). Plusieurs analyseurs, dont LIFT, peuvent être envisagés pour cette analyse. Leurs principes et leurs limitations ont été présentés dans la partie 1.4 p.46. LIFT ressort comme une solution qui semble avantageuse, mais cet analyseur manquait de maturité et n'avait jamais été testé expérimentalement.

Le premier objectif de ce chapitre est donc l'optimisation de LIFT : je cherche un réglage optimal pour que LIFT soit aussi compétitif que possible. Le second objectif est de valider le concept expérimentalement.

Je reprends tout d'abord la description de LIFT de la partie 1.4 p.46 en l'approfondissant. J'ajoute notamment l'aspect polychromatique de la formation des données et je détaille l'algorithme.

J'aborde ensuite la première étape de mon travail de thèse, à savoir l'optimisation de LIFT et sa validation en laboratoire. Dans un premier temps, j'ai cherché une phase de diversité optimale. Je me suis alors intéressé aux propriétés de convergence de l'algorithme, puis j'ai cherché à minimiser les coefficients de propagation du bruit.

Une fois LIFT optimisé, nous voulions vérifier qu'il était robuste à des incertitudes expérimentales. En particulier, une erreur sur l'échantillonnage ou l'amplitude de la phase de référence peut facilement conduire à une sur- ou sous-estimation systématique du focus. Une première étude en simulation a donc été faite, afin de s'assurer de la robustesse de LIFT à ces deux paramètres. Ensuite, j'ai mené des tests sur un banc dédié pour confirmer expérimentalement la linéarité de l'estimation, en lumière monochromatique ou polychromatique, ainsi que la propagation du bruit.

Enfin, je me suis appuyé sur la matrice de Fisher pour comparer LIFT à d'autres analyseurs dans le cadre de l'estimation du tip/tilt et du focus sur étoile naturelle.

La partie 2.2 p.66 décrit le principe de LIFT. L'optimisation et la validation en laboratoire sont présentées dans un article publié au cours de la thèse, et reproduit dans la partie 2.3 p.70. Cette partie inclut les résultats de l'étude de robustesse faite en simulation (§2.3.4 p.88). Bien que cette étude ait été faite avant la validation en laboratoire, elle n'apparaît dans l'article et est

donc présentée dans le paragraphe qui suit l'article. Nous verrons enfin, dans la partie 2.4 p.89, la comparaison de LIFT avec d'autres analyseurs à l'aide de la matrice de Fisher.

2.2 LIFT : un analyseur plan focal

Nous détaillons ici l'explication de la partie 1.4.2 p.56. Nous reprenons la formation d'images en ajoutant l'aspect polychromatique, puis nous nous intéressons à l'algorithme.

2.2.1 Formation d'images

Soit r le vecteur de coordonnées dans le plan focal, ρ le vecteur de coordonnées dans le plan pupille, et λ la longueur d'onde. La répartition d'intensité I sur le détecteur s'écrit :

$$\begin{aligned} I(r, \varphi + \varphi_d) &= \int_{\lambda} F(\lambda) |\text{FT}\{\underbrace{P(\rho) \exp[i\varphi_d(\rho, \lambda)]}_{P_d} \times \exp[i\varphi(\rho, \lambda)]\}_r|^2 d\lambda \quad (2.1) \\ &= \int_{\lambda} F(\lambda) I_{\lambda}(r, \varphi + \varphi_d) d\lambda \end{aligned}$$

avec $P(\rho)$ la transmission de la pupille, $F(\lambda)$ le flux, φ la phase à estimer et φ_d la phase de diversité. $\text{FT}\{f(\rho)\}_r$ est la transformée de Fourier du plan pupille vers le plan focal induite par la diffraction de Fraunhofer. $I_{\lambda}(r, \varphi + \varphi_d)$ est la répartition d'intensité à un λ donné. L'intégrale peut être vue comme une somme discrète sur les longueurs d'onde :

$$I(r, \varphi + \varphi_d) = \sum_{\lambda} F_{\lambda} I_{\lambda}(r, \varphi + \varphi_d) \quad (2.2)$$

φ peut être décomposé sur la base des modes de Zernike Z_i tel que $\varphi = \sum_i a_i Z_i$. Soit \mathbf{A} le vecteur de coefficients a_i . On suppose que les aberrations, en différence de chemin optique ($\mathbf{A}_{\text{OPD}} = \mathbf{A} \times \lambda/2\pi$), ne dépendent pas de la longueur d'onde.

Pour les besoins de la simulation numérique, on considère maintenant la répartition d'intensité comme un vecteur comprenant tous les pixels du détecteur. On peut alors l'exprimer uniquement en fonction de \mathbf{A}_{OPD} . L'équation (2.2) devient :

$$\mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)[p] = \sum_{\lambda} F_{\lambda} \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)[p] + \mathbf{n}[p] \quad (2.3)$$

avec $\mathbf{I}(\mathbf{A}_d)[p]$ l'intégration sur le p -ième pixel de $I(r, \varphi_d)$ et $\mathbf{n}[p]$ le bruit sur le pixel p . On fait un développement de Taylor au premier ordre pour linéariser l'équation (2.3), comme décrit par Gonsalves pour l'estimation de faible phase [Gonsalves, 2001] :

$$\sum_{\lambda} F_{\lambda} \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d) \simeq \sum_{\lambda} F_{\lambda} [\mathbf{I}_{\lambda}(\mathbf{A}_d) + \sum_k a_k \mathbf{I}'_{k\lambda}(\mathbf{A}_d)] + \mathbf{n} \quad (2.4)$$

ce qui est équivalent à :

$$\sum_{\lambda} F_{\lambda} [\mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d) - \mathbf{I}_{\lambda}(\mathbf{A}_d)] \simeq \sum_{\lambda} \sum_k a_k F_{\lambda} \mathbf{I}'_{k\lambda}(\mathbf{A}_d) + \mathbf{n} \quad (2.5)$$

avec $\mathbf{I}'_{k\lambda}(\mathbf{A}_d) = \left. \frac{\partial \mathbf{I}_\lambda(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)}{\partial a_k} \right|_{\mathbf{A}_{\text{OPD}}=0}$.

On définit $\Delta \mathbf{I}_\lambda$ par :

$$\Delta \mathbf{I}_\lambda[p] = \mathbf{I}_\lambda(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)[p] - \mathbf{I}_\lambda(\mathbf{A}_d)[p] \quad (2.6)$$

et $\mathbf{H}_\lambda(\mathbf{A}_d)$ la matrice dont l'élément à la p-ième ligne et k-ième colonne est défini par :

$$\mathbf{H}_\lambda(\mathbf{A}_d)[p, k] = \mathbf{I}'_{k\lambda}(\mathbf{A}_d)[p] \quad (2.7)$$

L'équation (2.5) devient alors :

$$\sum_{\lambda} F_{\lambda} \Delta \mathbf{I}_{\lambda} = \left(\sum_{\lambda} F_{\lambda} \mathbf{H}_{\lambda}(\mathbf{A}_d) \right) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (2.8)$$

Par analogie avec l'équation (2.3), on pose $\Delta \mathbf{I} = \sum_{\lambda} F_{\lambda} \Delta \mathbf{I}_{\lambda}$ et $\mathbf{H}(\mathbf{A}_d) = \sum_{\lambda} F_{\lambda} \mathbf{H}_{\lambda}(\mathbf{A}_d)$. La relation précédente s'écrit :

$$\Delta \mathbf{I} = \mathbf{H}(\mathbf{A}_d) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (2.9)$$

L'estimation de \mathbf{A}_{OPD} est faite par maximum de vraisemblance. Si la bande spectrale est suffisamment étroite, on peut ne prendre en compte que la longueur d'onde centrale :

$$\Delta \mathbf{I}_{\bar{\lambda}} = \mathbf{H}_{\bar{\lambda}}(\mathbf{A}_d) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (2.10)$$

avec $\bar{\lambda} = (\sum_{\lambda} F_{\lambda} \lambda) / (\sum_{\lambda} F_{\lambda})$. En pratique, on utilisera toujours un modèle monochromatique à la longueur d'onde centrale, mais le modèle polychromatique pourrait facilement être mis en oeuvre. L'efficacité de cette approximation en lumière polychromatique est évaluée dans le paragraphe 5.2.2 de l'article présenté dans la partie 2.3 p.70.

On fait l'approximation classique que le bruit de photon sur le k-ième pixel est un bruit additif gaussien de moyenne nulle, avec une variance égale au flux moyen \bar{I}_k [Mugnier *et al.*, 2004] (cf §1.2.1.1 p.22). Le bruit de détecteur suit une loi gaussienne de moyenne nulle avec un écart-type σ_e considéré uniforme sur le détecteur. La variance de bruit totale pour le pixel k est donc :

$$\sigma_{n_k}^2 = \bar{I}_k + \sigma_e^2 \quad (2.11)$$

Cette expression définit les éléments de la matrice de covariance $\mathbf{R}_n = \langle \mathbf{nn}^t \rangle$, supposée diagonale. \mathbf{R}_n est la notation utilisée dans l'article reproduit au paragraphe 2.3 p.70, elle correspond à C_b dans le chapitre 1 p.21.

2.2.2 Algorithme

On cherche \mathbf{A}_{OPD} tel que la vraisemblance $p(\mathbf{I} | \mathbf{A}_{\text{OPD}}, \mathbf{H}(\mathbf{A}_d), \mathbf{I}(\mathbf{A}_d))$ est maximale. A partir des hypothèses sur la statistique du bruit, la log-vraisemblance prend la forme :

$$-\ln [p(\mathbf{I} | \mathbf{A}_{\text{OPD}}, \mathbf{H}(\mathbf{A}_d), \mathbf{I}(\mathbf{A}_d))] = \frac{1}{2} (\Delta \mathbf{I} - \mathbf{H}(\mathbf{A}_d) \mathbf{A}_{\text{OPD}})^t \mathbf{R}_n^{-1} (\Delta \mathbf{I} - \mathbf{H}(\mathbf{A}_d) \mathbf{A}_{\text{OPD}}) + \text{cst} \quad (2.12)$$

La solution du maximum de vraisemblance est alors :

$$\hat{\mathbf{A}}_{\text{OPD}} = \mathbf{P}_{\text{MV}}(\mathbf{A}_d) \Delta \mathbf{I} \quad \text{avec} \quad \mathbf{P}_{\text{MV}}(\mathbf{A}_d) = (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_d))^{-1} \mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \quad (2.13)$$

Le développement de Taylor vu dans l'équation (2.4) a un domaine de validité limité, qui restreint le domaine de linéarité de l'analyseur. Cependant, ce domaine peut être agrandi en itérant l'estimation : une première estimation \hat{A}_{OPD}^0 est faite en 0, puis on définit cette estimation comme le nouveau point de fonctionnement ($\mathbf{H}(\mathbf{A}_d)$ devient $\mathbf{H}(\mathbf{A}_d + \hat{\mathbf{A}}_{\text{OPD}}^0)$), à partir duquel on fait une nouvelle estimation, et ainsi de suite (Fig. 2.1). On a finalement l'algorithme suivant :

- Itération 0 : $\Delta \mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_d)$; $\hat{\mathbf{A}}_{\text{OPD}}^0 = \mathbf{P}_{\text{MV}}(\mathbf{A}_d)\Delta \mathbf{I}$
 - Itération 1 : $\Delta \mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^0 + \mathbf{A}_d)$; $\hat{\mathbf{A}}_{\text{OPD}}^1 = \mathbf{P}_{\text{MV}}(\hat{\mathbf{A}}_{\text{OPD}}^0 + \mathbf{A}_d)\Delta \mathbf{I} + \hat{\mathbf{A}}_{\text{OPD}}^0$
 - Itération 2 : $\Delta \mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^1 + \mathbf{A}_d)$; $\hat{\mathbf{A}}_{\text{OPD}}^2 = \mathbf{P}_{\text{MV}}(\hat{\mathbf{A}}_{\text{OPD}}^1 + \mathbf{A}_d)\Delta \mathbf{I} + \hat{\mathbf{A}}_{\text{OPD}}^1$
- ⋮

avec \mathbf{I}_{data} l'image acquise sur le détecteur et $\mathbf{I}(\mathbf{X})$ une image simulée à partir de notre modèle.

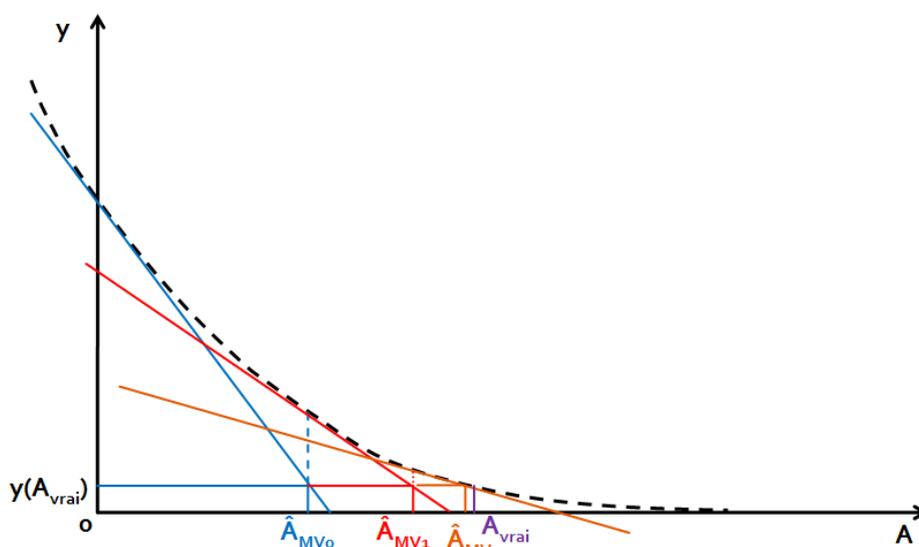


FIGURE 2.1 – Compensation de l'approximation faite par la linéarisation grâce aux itérations. On considère une mesure y , qui est une fonction de A . On cherche à estimer A_{vrai} , qui a donné la mesure $y(A_{\text{vrai}})$. L'algorithme part de $A = 0$, où la fonction est linéarisée, ce qui revient à prendre sa tangente en ce point. On regarde ensuite pour quel A cette tangente donne la mesure $y(A_{\text{vrai}})$. On obtient ainsi la première estimation $\hat{A}_{\text{MV}0}$. On linéarise la fonction à nouveau en ce point, puis on réitère le processus jusqu'à converger vers A_{vrai} .

L'extension du domaine de linéarité en fonction du nombre d'itérations est montré sur la figure 2.2 pour le tip/tilt et le focus. Quelques itérations sont suffisantes pour améliorer la linéarité. La tache n'étant pas forcément située au centre du détecteur, une estimation du centre de gravité de la tache à la première itération permet d'initialiser le point de fonctionnement. Le domaine de linéarité du tip/tilt n'est alors limité que par les bords du détecteur.

LIFT estimera des bas ordres (tip/tilt/focus) dans un contexte où les hauts ordres sont déjà corrigés grâce aux étoiles laser. Les images à analyser présenteront donc peu de hauts ordres. De plus, afin de limiter les effets de repliement et par analogie avec la méthode du trou filtrant dans le Shack-Hartmann [Poyneer et Macintosh, 2004], seule une zone carrée limitée (~ 5

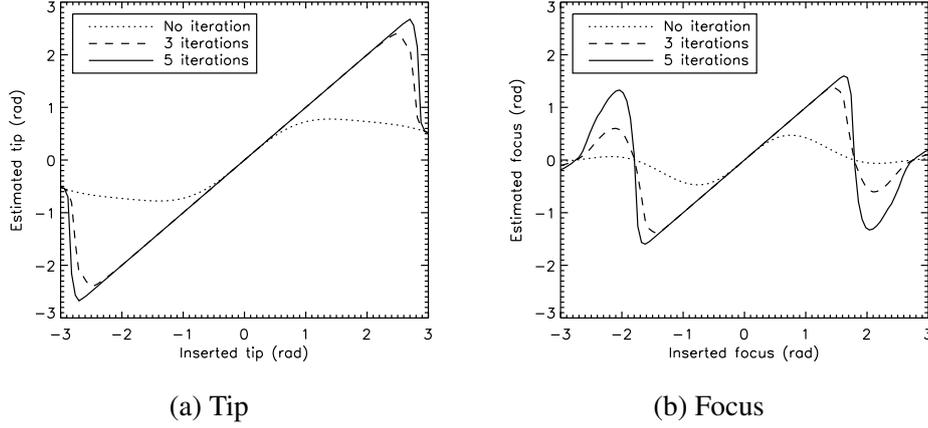


FIGURE 2.2 – Validation par simulation de l’augmentation du domaine de linéarité avec différents nombres d’itérations pour le tip et le focus. Phase de référence : 0.4 rad d’astigmatisme. Lumière monochromatique.

diamètres de tache d’Airy) centré sur la tache image est prise en compte. Cela diminue aussi le temps de calcul. D’autre part, LIFT offre la possibilité d’estimer un peu plus de modes, ce qui peut aider à limiter les effets de repliement [Meimon *et al.*, 2010a]. En première approximation, ces effets sont considérés comme négligeables, par conséquent les hauts ordres sont nuls dans toute la suite de ce chapitre. L’étude du comportement de LIFT en présence de hauts ordres est abordée au chapitre 3 p.95.

D’autre part, une estimation du flux est nécessaire pour évaluer le bruit et créer l’image modèle. Le flux est calculé en faisant la somme des pixels de la zone utile. Le but est de ne pas prendre trop de pixels en compte, au cas où le bruit de détecteur est dominant ou le fond de la caméra est mal soustrait. Comme une grande majorité des photons se trouvent au centre de la tache, le flux n’est que légèrement sous-estimé, et cela n’induit pas d’erreur sur l’estimation des modes.

Discutons du critère utilisé pour arrêter les itérations. Bien qu’il y ait une linéarisation dans le modèle, qui a permis de trouver une solution analytique, notre méthode s’inspire d’une estimation MV classique, qui utilise un modèle non-linéaire. Sans l’approximation faite par la linéarisation, l’estimateur MV $\hat{\mathbf{A}}_{\text{OPD}}^{\text{MV}}$ vérifie :

$$\hat{\mathbf{A}}_{\text{OPD}}^{\text{MV}} = \arg \min_{\mathbf{A}_{\text{OPD}}} ([\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)]^t \mathbf{R}_n^{-1} [\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_d)]) \quad (2.14)$$

On a donc choisi d’évaluer, à chaque itération k , le critère :

$$C_k = [\mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^{k-1} + \mathbf{A}_d)]^t \mathbf{R}_n^{-1} [\mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^{k-1} + \mathbf{A}_d)] \quad (2.15)$$

L’algorithme itère jusqu’à ce que la différence relative $|C_k - C_{k-1}|/C_k$ soit inférieure à 10^{-6} ou jusqu’à ce que la moyenne de la différence entre deux estimations successives $\langle \|\hat{\mathbf{A}}_{\text{OPD}}^k - \hat{\mathbf{A}}_{\text{OPD}}^{k-1}\| \rangle$ soit inférieure à 10^{-3} rad.

L’utilisation de LIFT en pratique est détaillée dans l’annexe B p.191.

2.3 Optimisation et validation en laboratoire

2.3.1 Guide de lecture

Cette étude a fait l'objet d'un article reproduit dans la suite. La première partie (§2) est une description de LIFT semblable à celle du paragraphe 2.2 p.66. Les notations sont les mêmes, sauf pour la phase de diversité φ_d qui devient une *offset phase* φ_{off} . Le lecteur pourra passer rapidement sur cette partie. Ensuite, nous discutons du choix des paramètres (§3). Notamment, pour le choix de la phase de diversité, nous nous appuyons sur la forme du critère à minimiser (§3.2, §3.3.1) ainsi que la propagation du bruit sur l'estimation (§3.3.2). Nous profitons de cette étude pour comparer LIFT au Shack-Hartmann 2x2 et à la pyramide en termes de sensibilité au bruit (§4). Nous présentons ensuite les résultats de la validation expérimentale en laboratoire (§5). Ces tests ont été fait sur des phases statiques ne présentant que du tip/tilt et du focus. Nous montrons d'abord la linéarité de l'estimation du tip/tilt en monochromatique (§5.2.1), puis du focus en monochromatique et polychromatique (§5.2.2). Enfin, nous confirmons notre modèle de propagation du bruit (§5.3).

2.3.2 Article

L'article est reproduit dans les pages suivantes.

Experimental validation of LIFT for estimation of low-order modes in low-flux wavefront sensing

C. Plantet,^{1*} S.Meimon,¹ J.-M. Conan¹ and T. Fusco^{1,2}

¹ONERA, The French Aerospace Lab
92322 Chatillon, France

²Aix Marseille Universite, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR
7326

13388 Marseille, France

*cedric.plantet@onera.fr

Abstract: Laser Tomographic and Multi-Conjugate Adaptive Optics systems rely on natural guide stars to sense low order aberrations (tip/tilt and focus). LIFT is a novel focal plane wavefront sensor (WFS), performing a maximum likelihood phase retrieval on a single image, with better sensitivity than a 2×2 Hartmann-Shack WFS. Its performance for the estimation of tip/tilt and focus is similar to a pyramid WFS without modulation, but with a simpler set-up. We present here the LIFT concept and associated data processing, as well as experimental results. We validate the estimation of tip/tilt and focus, with monochromatic and large bandwidth light, and verify the low noise sensitivity predicted by theory.

© 2013 Optical Society of America

OCIS codes: (010.1080) Active or adaptive optics; (010.7350) Wave-front sensing.

References and links

1. T. Fusco, S. Meimon, Y. Clenet, M. Cohen, H. Schnetler, J. Paufigue, V. Michau, J.-P. Amans, D. Gratadour, C. Petit, C. Robert, P. Jagourel, E. Gendron, G. Rousset, J.-M. Conan and N. Hubin, "ATLAS: the E-ELT laser tomographic adaptive optics system," *Proc. SPIE* **7736** 77360D (2010).
2. E. Diolaiti, J.-M. Conan, I. Foppiani, M. Lombini, E. Marchetti, C. Petit, C. Robert, L. Schreiber, A. Baruffolo, M. Bellazzini, G. Bregoli, P. Ciliegi, G. Cosentino, V. Biliotti, S. D'Odorico, T. Fusco, N. Hubin, S. Meimon and J.-F. Sauvage, "MAORY E-ELT multi-conjugate adaptive optics module," *Proc. AO4ELT 02007* (2011).
3. G. Herriot, P. Hickson, B. L. Ellerbroek, D. A. Andersen, T. Davidge, D. A. Erickson, I. P. Powell, R. Clare, L. Gilles, C. Boyer, M. Smith, L. Saddlemyer and J.-P. Veran, "NFIRAOS: TMT narrow field, near-infrared facility adaptive optics," *Proc. SPIE* **6272** 62720Q (2006).
4. B. Neichel, F. Rigaut, M. Bec, M. Boccas, F. Daruich, C. D'Orgeville, V. Fesquet, R. Galvez, A. Garcia-Rissmann, G. Gausachs, M. Lombini, G. Perez, G. Trancho, V. Upadhy and T. Vucina, "The Gemini MCAO System GeMS: nearing the end of a lab-story," *Proc. SPIE* **7736** 773606-1 (2010).
5. G. Rousset, "Wave-front sensors," in *Adaptive Optics in Astronomy*, F. Roddier, (Cambridge University, 1999) 91–130.
6. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," *Appl. Opt.* **27**(7), 1223–1225 (1988).
7. O. Guyon, "Limits of adaptive optics for high contrast imaging," *Astrophys. J.* **629**, 592–614 (2005).
8. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," *J. Mod. Opt.* **43**(2), 289–293 (1996).
9. R. Ragazzoni and J. Farinato, "Sensitivity of a pyramidal wave front sensor in closed loop adaptive optics," *Astron. Astrophys.* **350**, L23–L26 (1999).
10. C. Verinaud, "On the nature of the measurements provided by a pyramid wave-front sensor," *Opt. Commun.* **233**, 27–38 (2004).
11. R. Dekany, Caltech Optical Observatories, Physics, Math and Astronomy Division, California Institute of Technology, Pasadena, Calif. 91125, USA (personal communication).

12. C. Paterson and J. C. Dainty, "Hybrid curvature and gradient wave-front sensor," *Opt. Lett.* **25**(23), 1687–1689 (2000).
13. S. Barwick, "Performance comparison between Shack-Hartmann and astigmatic hybrid wavefront sensors," *Appl. Opt.* **48**(36), 6967–6972 (2009).
14. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.* **21**, 829–832 (1982).
15. I. Mocoer, L. M. Mugnier and F. Cassaing, "Analytical solution to the phase-diversity problem for real-time wavefront sensing," *Opt. Lett.* **34**(22), 3487–3489 (2009).
16. S. Meimon, T. Fusco and L. M. Mugnier, "LIFT: a focal-plane wavefront sensor for real-time low-order sensing on faint sources," *Opt. Lett.* **35**(18), 3036–3038 (2010).
17. C. Plantet, S. Meimon, T. Fusco, and J.-M. Conan, "Experimental validation of the linearized focal-plane technique (LIFT)," *Proc. AO4ELT2* (2011).
18. R. A. Gonsalves, "Small-phase solution to the phase-retrieval problem," *Opt. Lett.* **26**(10), 684–685 (2001).
19. L.A. Poyneer and B. Macintosh, "Spatially filtered wave-front sensor for high-order adaptive optics," *J. Opt. Soc. Am.* **21**, 810–819 (2004).
20. S. Meimon, T. Fusco and F. Cassaing, "A focal plane sensor for low-order sensing on laser tomographic systems: LIFT," *Proc. SPIE* **7736** 773611 (2010).
21. D. K. Cohen, W. H. Gee, M. Ludeke and J. Lewkowicz, "Automatic focus control: the astigmatic lens approach," *J. Opt. Soc. Am.* **23**, 565–570 (1984).
22. M. Nicolle, T. Fusco, G. Rousset, and V. Michau, "Improvement of Shack-Hartmann wave-front sensor measurement for extreme adaptive optics," *Opt. Lett.* **29**(18), 2743–2745 (2004).
23. S. Thomas, T. Fusco, A. Tokovinin, M. Nicolle, V. Michau and G. Rousset, "Comparison of centroid computation algorithms in a Shack-Hartmann sensor," *Mon. Not. R. Astron. Soc.* **371**, 323–336 (2006).
24. P. Feautrier, J.-L. Gach, M. Downing, P. Jorden, J. Kolb, J. Rothman, T. Fusco, P. Balard, E. Stadler, C. Guillaume, D. Boutolleau, G. Destefanis, N. Lhermet, O. Pacaud, M. Vuillemermet, A. Kerlain, N. Hubin, J. Reyes, M. Kasper, O. Ivert, W. Suske, A. Walker, M. Skegg, S. Derelle, J. Deschamps, C. Robert, N. Vedrenne, F. Chazalet, J. Tanchon, T. Trolier, A. Ravex, G. Zins, P. Kern, T. Moulin and O. Preis, "Advances in detector technologies for visible and infrared wavefront sensing," *Proc. SPIE* **8447**, 84470Q (2012).
25. C. Plantet, B. Neichel, S. Meimon, T. Fusco, and J.-M. Conan, "LIFT, a noise-effective low order focal-plane sensor: from theory to full experimental validation," *Proc. SPIE* (2012).

1. Introduction

Tomographic adaptive optics uses laser guide stars to build a 3D model of the turbulence. Nevertheless, natural guide stars are still needed to sense low order modes [1–4]. To get a maximum sky coverage, we need to sense stars as faint as possible. Hence, wavefront sensors must have a very low sensitivity to noise. Several wavefront sensors are already used or could be considered for the analysis of natural stars. We can cite the Shack-Hartmann (SH), curvature, pyramid, modified quad-cell sensors and phase diversity.

The Shack-Hartmann has been used and studied a lot, making it a reliable and widespread sensor. It is a pupil-plane sensor using microlenslets dividing the pupil into subapertures [5]. The local wavefront slope is measured in each subaperture to retrieve the wavefront surface. The signal to noise ratio is inversely proportional to the number of subapertures, since the total flux is separated among them. As it needs at least 4 subapertures to estimate the focus, it may not be effective enough for the faintest stars.

The curvature sensor, introduced by Roddier in 1988 [6], measures second derivatives of the wavefront instead of slopes. It has been shown that it is more sensitive to noise in low spatial frequencies than a Shack-Hartmann [7]. It is thus not adapted to the estimation of low order modes on natural guide stars.

Ragazzoni proposed in 1996 a new pupil-plane sensor using a pyramidal prism [8]. This sensor has been studied in terms of photon noise, and has better performance than a Shack-Hartmann [7, 9, 10]. This sensor could be a competitive one in the search for the best possible sky coverage. Nevertheless, it requires a complex setup and the pyramidal prism comes as an additional component.

On the other side of the WFS spectrum, we find focal-plane sensors. A simple and attractive solution, the modified quad-cell, has been proposed by Dekany for the Keck telescope [11]. It

consists in adding an astigmatism offset before focalization on a 4 pixel detector. This method allows the estimation of a focus in addition of the classical tip/tilt. Nevertheless, this sensor has poor linearity and is strongly sensitive to aliasing. Another concept, based on the modified quad-cell and the curvature sensor has been developed by Paterson and Dainty [12]. Unfortunately, this sensor seems to lose efficiency at low flux and does not outperform a Shack-Hartmann [13].

Phase diversity is another focal-plane sensor relying on two images [14]. One is taken at the focal plane and one with a focus offset. An algorithm deduces the aberrations from these images. Mocoer *et al.* [15] proposed a method to make this sensor fitted for real-time analysis, and showed good results for small phase estimations of low-order aberrations. Though, the flux is divided between the two images, which makes the set-up a bit complex and potentially reduces the detection efficiency.

We recently proposed a wavefront sensor called LIFT (Linearized Focal-plane Technique), dedicated to faint natural guide stars [16], and requiring a very simple hardware. The estimation by LIFT is based upon a single image taken at the focal plane of the telescope. The relation between the aberration and the intensity pattern is linearized to make computations easier and faster. Besides, we add a known phase offset to avoid phase indeterminations. First experiments on ONERA's test bench have already helped validating this sensor [17]. The present article gives a detailed description of LIFT concept, as well as a theoretical analysis of its linearity and noise properties. The outcome of this analysis is compared with experimental results. The noise propagation of the Shack-Hartmann and pyramid sensors is also studied for comparison.

In section 2, we focus on the concept of LIFT and detail its algorithm. We then discuss in section 3 the optimization of the phase offset by means of analytical calculations on the convergence of estimation and noise propagation. This last analysis is used to compare LIFT with the Shack-Hartmann and pyramid sensors in section 4. We finally present in section 5 the full in-lab experimental validation of LIFT for the estimation of low order modes: tip, tilt and focus.

2. LIFT : a phase-retrieval wavefront sensor

LIFT is a focal plane WFS, comprising a focusing optical system introducing a known amount of astigmatism, a focal plane detection, and an original phase retrieval algorithm. LIFT, as a natural guide star low order WFS, is meant to be used in infrared (IR) light with a high order correction on laser guide stars. After a description of data formation, we go into the details of this algorithm.

2.1. Data formation

Let r be the vector of coordinates in the focal plane, ρ the vector of coordinates in the pupil plane, and λ the wavelength. The intensity pattern I on the imaging sensor is described by:

$$\begin{aligned} I(r, \phi + \phi_{off}) &= \int_{\lambda} F(\lambda) |\text{FT}\{\underbrace{P(\rho) \exp[i\phi_{off}(\rho, \lambda)]}_{P_{off}} \times \exp[i\phi(\rho, \lambda)]\}_r|^2 d\lambda \\ &= \int_{\lambda} F(\lambda) I_{\lambda}(r, \phi + \phi_{off}) d\lambda \end{aligned} \quad (1)$$

with $P(\rho)$ the pupil transmission, $F(\lambda)$ the incoming flux, ϕ the aberrated phase to be estimated and ϕ_{off} the known phase offset. $\text{FT}\{f(\rho)\}_r$ is the Fourier Transform from the pupil plane to the focal plane as described by Fraunhofer diffraction. $I_{\lambda}(r, \phi + \phi_{off})$ is the intensity pattern at a given λ . The integral can be considered as a discrete sum over the wavelengths:

$$I(r, \phi + \phi_{off}) = \sum_{\lambda} F_{\lambda} I_{\lambda}(r, \phi + \phi_{off}) \quad (2)$$

ϕ can be decomposed on Zernike modes Z_i so that $\phi = \sum_i a_i Z_i$. Let \mathbf{A} be the vector of coefficients a_i . We assume that the aberrations, in optical path difference ($\mathbf{A}_{\text{OPD}} = \mathbf{A} \times \lambda / 2\pi$), do not depend on the wavelength.

For the needs of numerical computations, we consider now the intensity pattern as a vector comprising all the pixels of the imaging detector. It can then be expressed as a function of \mathbf{A}_{OPD} only. Equation (2) becomes:

$$\mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})[p] = \sum_{\lambda} F_{\lambda} \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})[p] + \mathbf{n}[p] \quad (3)$$

with $\mathbf{I}(\mathbf{A}_{\text{off}})[p]$ corresponding to the integration on the p-th pixel of $I(r, \phi_{\text{off}})$. \mathbf{n} is the noise on pixels. We use a first order Taylor expansion to linearize Eq. (3), as described by Gonsalves for small phase estimation [18]:

$$\sum_{\lambda} F_{\lambda} \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}}) \simeq \sum_{\lambda} F_{\lambda} [\mathbf{I}_{\lambda}(\mathbf{A}_{\text{off}}) + \sum_k a_k \mathbf{I}'_{k\lambda}(\mathbf{A}_{\text{off}})] + \mathbf{n} \quad (4)$$

which is equivalent to:

$$\sum_{\lambda} F_{\lambda} [\mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}}) - \mathbf{I}_{\lambda}(\mathbf{A}_{\text{off}})] \simeq \sum_{\lambda} \sum_k a_k F_{\lambda} \mathbf{I}'_{k\lambda}(\mathbf{A}_{\text{off}}) + \mathbf{n} \quad (5)$$

with $\mathbf{I}'_{k\lambda}(\mathbf{A}_{\text{off}}) = \left. \frac{\partial \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})}{\partial a_k} \right|_{\mathbf{A}_{\text{OPD}} = \mathbf{0}}$.

Let $\Delta \mathbf{I}_{\lambda}$ be the vector so that $\Delta \mathbf{I}_{\lambda}[p] = \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})[p] - \mathbf{I}_{\lambda}(\mathbf{A}_{\text{off}})[p]$, and $\mathbf{H}_{\lambda}(\mathbf{A}_{\text{off}})$ be the matrix of which the element at the p-th row and k-th column is defined by $\mathbf{H}_{\lambda}(\mathbf{A}_{\text{off}})[p, k] = \mathbf{I}'_{k\lambda}(\mathbf{A}_{\text{off}})[p]$. Then Eq. (5) becomes:

$$\sum_{\lambda} F_{\lambda} \Delta \mathbf{I}_{\lambda} = \left(\sum_{\lambda} F_{\lambda} \mathbf{H}_{\lambda}(\mathbf{A}_{\text{off}}) \right) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (6)$$

By analogy with Eq. (3), we define $\Delta \mathbf{I} = \sum_{\lambda} F_{\lambda} \Delta \mathbf{I}_{\lambda}$ and $\mathbf{H}(\mathbf{A}_{\text{off}}) = \sum_{\lambda} F_{\lambda} \mathbf{H}_{\lambda}(\mathbf{A}_{\text{off}})$. The previous relation writes:

$$\Delta \mathbf{I} = \mathbf{H}(\mathbf{A}_{\text{off}}) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (7)$$

The estimation of \mathbf{A}_{OPD} is made by maximum likelihood (ML). If the optical bandwidth is narrow enough, we can consider the central wavelength only:

$$\Delta \mathbf{I}_{\bar{\lambda}} = \mathbf{H}_{\bar{\lambda}}(\mathbf{A}_{\text{off}}) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (8)$$

with $\bar{\lambda} = (\sum_{\lambda} F_{\lambda} \lambda) / (\sum_{\lambda} F_{\lambda})$. In practice and in this article, we always use a monochromatic model at the central wavelength, but the polychromatic model could be easily implemented. The accuracy of estimation with this approximation in polychromatic light is evaluated in section 5.2.2.

LIFT is meant to sense low orders (tip/tilt/focus) in a context where higher orders are corrected based on laser guide star measurements. The images to be analyzed should therefore present a low amount of high orders. Besides, in order to limit aliasing and by analogy with the pinhole filtering method in a Shack-Hartmann WFS [19], only a limited square zone (~ 5 Airy diameters) centered on the image spot is taken into account. This lowers computation time as well. Also, LIFT offers the possibility to estimate a few modes beyond tip/tilt/focus, which can help reduce aliasing [20]. As a first approximation, aliasing effects are assumed negligible, and consequently higher modes are set to zero in the present paper. A detailed study of LIFT's behavior in the presence of high orders will come in a future publication.

2.2. Algorithm

We approximate the k -th pixel photon noise by an additive zero-mean Gaussian noise, with a variance equal to the mean flux \bar{I}_k . Read-out noise follows a zero-mean Gaussian distribution with a standard deviation σ_e considered uniform on the detector, hence the global noise variance for the pixel k :

$$\sigma_{n_k}^2 = \bar{I}_k + \sigma_e^2 \quad (9)$$

This expression defines the elements of the covariance matrix $\mathbf{R}_n = \langle \mathbf{nn}^t \rangle$, which is assumed to be diagonal. The Gaussian approximation for photon noise is accurate with at least 10 photons per pixel. Under 10 photons, the read-out noise is dominant in IR imaging.

We search \mathbf{A}_{OPD} so that the likelihood $p(\mathbf{I}|\mathbf{A}_{\text{OPD}}, \mathbf{H}(\mathbf{A}_{\text{off}}), \mathbf{I}(\mathbf{A}_{\text{off}}))$ is maximum. Based on the assumed noise statistics, the log likelihood takes the form:

$$-\ln [p(\mathbf{I}|\mathbf{A}_{\text{OPD}}, \mathbf{H}(\mathbf{A}_{\text{off}}), \mathbf{I}(\mathbf{A}_{\text{off}}))] = \frac{1}{2} (\Delta\mathbf{I} - \mathbf{H}(\mathbf{A}_{\text{off}})\mathbf{A}_{\text{OPD}})^t \mathbf{R}_n^{-1} (\Delta\mathbf{I} - \mathbf{H}(\mathbf{A}_{\text{off}})\mathbf{A}_{\text{OPD}}) + \text{cst} \quad (10)$$

The maximum likelihood solution is thus:

$$\hat{\mathbf{A}}_{\text{OPD}} = \mathbf{P}_{\text{ML}}(\mathbf{A}_{\text{off}})\Delta\mathbf{I} \quad \text{with} \quad \mathbf{P}_{\text{ML}}(\mathbf{A}_{\text{off}}) = (\mathbf{H}(\mathbf{A}_{\text{off}})^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_{\text{off}}))^{-1} \mathbf{H}(\mathbf{A}_{\text{off}})^t \mathbf{R}_n^{-1} \quad (11)$$

and the variance of the estimation error $\mathbf{E} = \mathbf{A}_{\text{OPD}} - \hat{\mathbf{A}}_{\text{OPD}}$ is given by:

$$\sigma_{\text{error}}^2 = \text{Tr}\{\langle \mathbf{E}\mathbf{E}^t \rangle\} = \text{Tr}\{(\mathbf{H}(\mathbf{A}_{\text{off}})^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_{\text{off}}))^{-1}\} \quad (12)$$

Each diagonal element of $\langle \mathbf{E}\mathbf{E}^t \rangle$ gives the variance of the estimation error for one mode.

Of course, first order Taylor expansion as seen in Eq. (4) has a limited validity range, which defines the sensor linearity domain. However, this domain can be extended by iterating the estimation. We use the following algorithm:

- Iteration 0: $\Delta\mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{off}})$; $\hat{\mathbf{A}}_{\text{OPD}}^0 = \mathbf{P}_{\text{ML}}(\mathbf{A}_{\text{off}})\Delta\mathbf{I}$
- Iteration 1: $\Delta\mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^0 + \mathbf{A}_{\text{off}})$; $\hat{\mathbf{A}}_{\text{OPD}}^1 = \mathbf{P}_{\text{ML}}(\hat{\mathbf{A}}_{\text{OPD}}^0 + \mathbf{A}_{\text{off}})\Delta\mathbf{I} + \hat{\mathbf{A}}_{\text{OPD}}^0$
- Iteration 2: $\Delta\mathbf{I} = \mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^1 + \mathbf{A}_{\text{off}})$; $\hat{\mathbf{A}}_{\text{OPD}}^2 = \mathbf{P}_{\text{ML}}(\hat{\mathbf{A}}_{\text{OPD}}^1 + \mathbf{A}_{\text{off}})\Delta\mathbf{I} + \hat{\mathbf{A}}_{\text{OPD}}^1$
- ⋮

with \mathbf{I}_{data} being the image recorded on the detector and $\mathbf{I}(\mathbf{X})$ a simulated image from our model.

The extension of the linearity domain with the number of iterations is shown on Fig. 1 for tip/tilt and focus. A few iterations are sufficient to improve linearity.

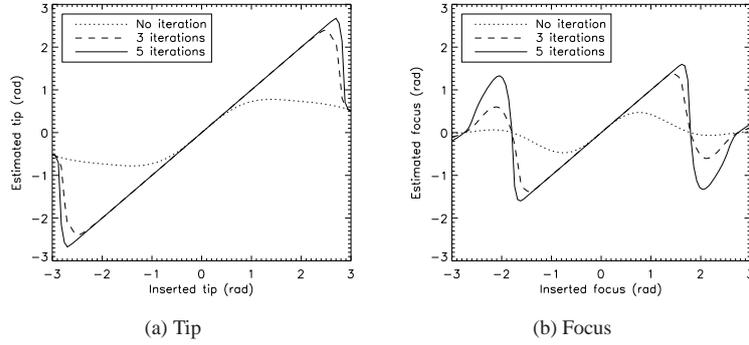


Fig. 1. Simulation of the linearity domain with different numbers of iterations for tip and focus. Phase offset: 0.4 rad of astigmatism. Monochromatic light.

We may discuss the stopping criterion used to determine the number of iterations. Without the approximation made by linearization, the maximum likelihood estimator $\hat{\mathbf{A}}_{\text{OPD}}$ verifies:

$$\hat{\mathbf{A}}_{\text{OPD}}^{\text{ML}} = \arg \min_{\mathbf{A}_{\text{OPD}}} ([\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})]^t \mathbf{R}_n^{-1} [\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})]) \quad (13)$$

We thus chose to evaluate, at each iteration k , the criterion:

$$C_k = [\mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^{k-1} + \mathbf{A}_{\text{off}})]^t \mathbf{R}_n^{-1} [\mathbf{I}_{\text{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\text{OPD}}^{k-1} + \mathbf{A}_{\text{off}})] \quad (14)$$

The algorithm iterates until the relative difference $|C_k - C_{k-1}|/C_k$ is lesser than 10^{-6} or until the mean of the difference between two successive estimations $\langle \|\hat{\mathbf{A}}_{\text{OPD}}^k - \hat{\mathbf{A}}_{\text{OPD}}^{k-1}\| \rangle$ is lesser than 10^{-3} rad.

3. Algorithm tuning

In this section, we discuss the choice of the algorithm parameters. In particular, we choose the phase offset in order to have an efficient focus estimation. We studied two possibilities: focus and astigmatism. A theoretical analysis of the minimization criterion and of the noise propagation helps us choose the appropriate aberration as well as its amplitude. In a major part of this section, we discuss the estimation of focus. The values of aberrations are given in radians rms in all the following.

3.1. Choice of the sampling

Too much oversampling would lead to important noise propagation, as flux is divided into more pixels. We thus chose to use a Shannon sampling, i.e. the pixel size is $\lambda/2D$ with D the diameter of the pupil. However, sampling is a parameter that can be changed to fit any system. A Shannon/2 sampling could also be considered, but is not discussed in this paper.

3.2. Choice of the offset mode

Without any phase offset, the focus estimation is not unique: two values of opposite sign lead to the same focal image. Adding a known phase to the aberrated wavefront will help us remove this indetermination.

A natural choice would be a focus offset, as used in phase diversity. But this only shifts the position of the focal plane. The symmetry with respect to the focused image is still present. For example, if we choose a 0.4 rad focus offset, a -0.3 rad or a -0.5 rad focus will lead to the exact same focal image with an absolute amount of focus equal to 0.1 rad. To understand better the algorithm behavior with respect to the inserted focus, we can analyze the structure of the criterion to be minimized (see Eq. (13)), which can be rewritten:

$$C(\mathbf{A}_{\text{OPD}}) = ([\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})]^t \mathbf{R}_n^{-1} [\mathbf{I}_{\text{data}} - \mathbf{I}(\mathbf{A}_{\text{OPD}} + \mathbf{A}_{\text{off}})]) \quad (15)$$

We recall that \mathbf{A}_{off} denotes the phase offset, for now a focus offset. \mathbf{A}_{OPD} is called the assumed focus, and the focus to be estimated is called true focus. Hence, \mathbf{I}_{data} is formed by addition of the focus offset and the true focus.

At first iteration, the criterion takes the value at 0 rad of assumed focus. We call it the starting point, which corresponds to an image with the focus offset only $\mathbf{I}(\mathbf{A}_{\text{off}})$. During minimization, the criterion follows the curve toward lower values. The estimated focus is the assumed focus for which $C(\mathbf{A}_{\text{OPD}})$ is minimum. If several minima co-exist, the estimated focus is at the minimum closest to the starting point.

Figure 2(a) plots in dotted line the criterion value depending on the assumed focus. The focus offset is 0.4 rad, and the true focus is 0.2 rad. The criterion should thus be minimum at 0.2 rad.

Two minima are visible on the curve. One is at 0.2 rad and the other at -1 rad. As expected, they are symmetrical with respect to -0.4 rad, where the true focus and the focus offset cancel each other out. In this case, as 0.2 rad is closer to the starting point, the focus estimation is correct. On the other hand, if the true focus is less than -0.4 rad, the closest minimum may be the wrong one. Figure 2(b) demonstrates that the estimated focus is -0.2 rad if we insert a -0.6 rad focus. Hence, adding a focus offset does not totally remove the indetermination.

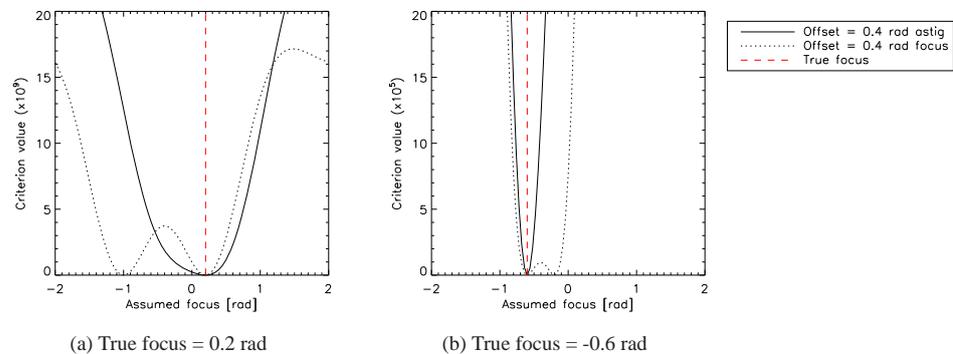


Fig. 2. Simulation of the minimization criterion value depending on the assumed focus. The curves are plotted for an offset phase at 0.4 rad, in focus or astigmatism. In the case of a focus offset, the estimated focus is 0.2 rad in (a) and -0.2 in (b). In the case of an astigmatism offset, the estimated focus is the true focus in (a) and (b).

Another intuitive choice is an astigmatism offset, since it gives particular shapes to the focal images when introducing a focus. The spot will stretch out in one direction with a positive focus, and in the perpendicular direction with a negative focus. Besides, astigmatism has already been used in focus control systems, and has shown to be effective [21]. If we compare it to a focus

offset in terms of minimization criterion, we observe only one minimum, at the true focus (Fig. 2). Hence, the astigmatism offset gives a better behavior which will lead to better linearity.

3.3. Choice of the offset amplitude

3.3.1. Impact on the minimization criterion

We have just seen that the criterion to be minimized has a better shape with an astigmatism offset than with a focus offset. But this shape actually depends on the amount of astigmatism introduced as a phase offset.

Figure 3(a) represents the criterion depending on the assumed focus and on the astigmatism offset. The true focus is still 0.2 rad. Each column corresponds to the criterion curve for a given astigmatism offset. For example, the column at the abscissa 0.4 rad gives the solid line in Fig. 2(a). The white dotted line at 0.2 rad of assumed focus is the criterion minimum that gives the true estimation for all astigmatism offsets, and the dashed line at 0 rad corresponds to the starting point. We confirm here that without any astigmatism, the estimation is undetermined, since we have two minima at 0.2 rad and -0.2 rad that are symmetrical with respect to the starting point.

This figure allows us to know the shape of the criterion for different astigmatism offsets. First of all, we checked that the offset has to be larger than 0.15 rad to avoid a minimum duality. The range 0.5-1.5 rad may seem attractive since the criterion is steeper, which should lead to better convergence of the algorithm. Though, a high astigmatism leads to a large spot, with less flux per pixel, and thus more sensitivity to noise. The following section shows how the noise propagation constrains the choice of the amplitude.

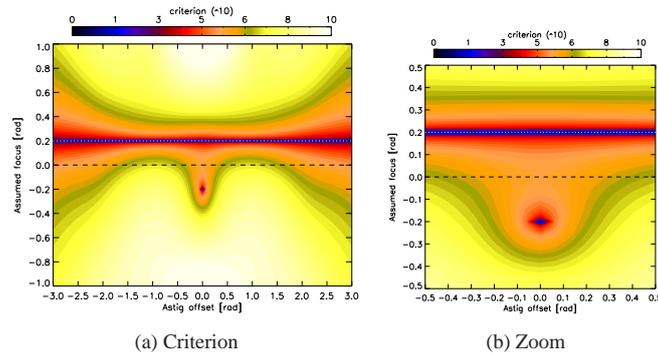


Fig. 3. Simulation of the minimization criterion value depending on the assumed focus and on the astigmatism offset. Dashed line: starting point. White dotted line: true focus to estimate.

3.3.2. Impact on noise propagation

The behavior of a wavefront sensor with respect to noise is described by the variance of its estimation error:

$$\sigma_{error}^2 = \sum_i \sigma^2(\hat{a}_i - a_i) = \overbrace{\left(\sum_i \alpha_i\right)}^{\alpha} \frac{1}{n_{ph}^{tot}} + \overbrace{\left(\sum_i \beta_i\right)}^{\beta} \left(\frac{\sigma_e}{n_{ph}^{tot}}\right)^2 \quad (16)$$

with n_{ph}^{tot} the total number of detected photons and σ_e the read out noise standard deviation (in $ph.e^-$). The variance is computed as a sum over the estimated modes. The coefficients α_i and

β_i are respectively the photon noise and read out noise coefficients for the Zernike polynomial Z_i .

For $\mathbf{A}_{OPD} = \mathbf{0}$, we can get the theoretical variance from Eq. (12), and then compute easily the coefficients. If we set the read-out noise to 0, we find

$$\alpha = \sum_i (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_d))^{-1} [i, i] \times n_{ph}^{tot} \quad (17)$$

and if we remove photon noise, we find

$$\beta = \sum_i (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_d))^{-1} [i, i] \times \left(\frac{n_{ph}^{tot}}{\sigma_e} \right)^2 \quad (18)$$

One can similarly compute the modal noise coefficients α_i and β_i . Low values mean a low sensitivity to noise.

Analytical calculations have shown that the amount of astigmatism we introduce can be optimized to lower α and β (Fig. 4) for the estimation of tip/tilt and focus. The coefficients decrease dramatically at low astigmatism, because of the indetermination removal. At higher astigmatism, we get less flux per pixels, hence the coefficients rise again. We find a different optimum for the two coefficients: 0.42 rad leading to $\alpha = 1.64$ and 0.6 rad leading to $\beta = 84$. The best amount of astigmatism to be introduced is thus between these two values, depending on the detector and signal characteristics. It is closer to 0.42 rad at high flux, where photon noise is predominant, and closer to 0.6 rad at low flux, where read-out noise is more important. For our simulations and experiments, we chose an astigmatism offset approximately at the middle of the bracket, at 0.5 rad, but this value can be tuned depending on needs. This value should also lead to good criterion properties, as discussed in section 3.3.1.

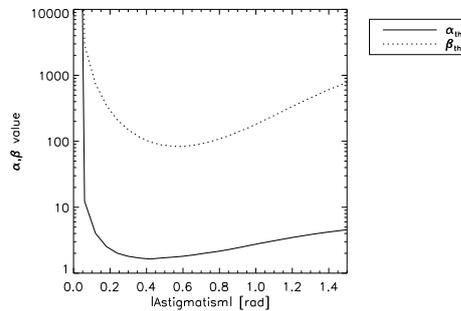


Fig. 4. α and β computed over 3 modes (tip/tilt/focus) depending on the amount of astigmatism introduced as phase offset (in absolute value).

We have justified the choice of astigmatism as a phase offset and demonstrated that it can be optimized depending on the conditions of experimentation. We now focus on the noise propagation analysis to compare LIFT with other sensors.

4. Comparison of LIFT's noise propagation with the Shack-Hartmann 2x2 and pyramid sensors

In the previous section, we related the variance of a WFS estimation error to noise through the coefficients α and β , which respectively quantify the sensitivity to photon noise and the sensitivity to read-out noise.

We compare the values of α and β computed over tip/tilt and focus for LIFT with a 0.5 rad astigmatism, for a Shack-Hartmann 2×2 and for a fixed (e.g. without modulation) pyramid sensor with a pupil sampled on 4 pixels (Table 1). The parameters and formulas for computing the Shack-Hartmann coefficients are those used in references [16, 22]. In particular, for the Weighted Center of Gravity (WCoG), we consider a fixed Gaussian weighting, and the estimation is corrected by an unbiasing factor, as described by Thomas *et al.* [23]. The Full Width at Half Maximum (FWHM) of the weighting is equal to the Airy spot FWHM, i.e. 2 pixels at Shannon sampling. In our pyramid simulation, we performed a maximum likelihood estimation, and were able to compute α and β with the same analysis we used for LIFT (see Eq. (12)).

Table 1 shows that LIFT is approximately 5 times less sensitive to photon noise and 4 times less sensitive to read-out noise than a SH 2×2 using a WCoG. In LIFT, the photon noise sensitivity is reduced by the full aperture gain. The SH 2×2 divides the flux by 4, and each spot is imaged at the resolution of a subaperture. LIFT thus gets more flux and better precision. Moreover, sensitivity to detector noise is limited thanks to the maximum likelihood estimation properties [16]. As a result, for the estimation of tip/tilt, LIFT takes the better from the Center of Gravity (CoG), which is efficient in photon noise regime and the weighted center of gravity, which is efficient in read-out noise regime [22]. This effect is clearly visible on Fig. 5, which plots the variance of estimation error of LIFT, a CoG and a WCoG for the estimation of a tip only.

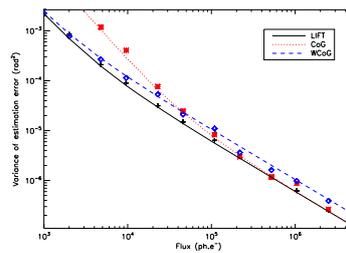


Fig. 5. Variance of the error on tip estimation for LIFT, the center of gravity and the weighted center of gravity at Shannon sampling on 8×8 pixels. Read-out noise = $10 e^-$. Lines : theory. Symbols : simulations. The weighting used for the WCoG is a Gaussian, and the estimation is corrected by an unbiasing factor. The parameters are $N_t = N_w = 2$ (N_t : FWHM of the aberrated spot, N_w : FWHM of the weighting function, see reference [23]).

The pyramid sensor benefits from the full aperture gain as well, and sensitivity to detector noise is reduced by the use of only 4 pixels. The similarity of noise coefficients could thus be expected.

Table 1. Comparison of noise propagation coefficients between LIFT, a Shack-Hartmann 2×2 (using either a center of gravity or a weighted center of gravity), and a fixed pyramid with a pupil sampled on 4 pixels for the estimation of tip/tilt and focus. Coefficients for the WCoG are computed in the case of an unbiased Gaussian weighting, with parameters $N_t = N_d = N_w = 2$ (see reference [23]).

	SH 2×2 (CoG)	SH 2×2 (WCoG)	LIFT	Pyramid
α	4.61	8.19	1.71	2.3
β	$> 10^4$	334	87	62

As an example of performance, we plot the variance of estimation error on tip/tilt and focus depending on flux for a read-out noise of $2 e^-$ (Fig. 6), which is the typical noise on high performance infrared cameras [24]. In this case, a Shack-Hartmann 2×2 using a weighted center of gravity would need a flux of 958 photons to reach a total variance of 0.01 rad^2 over the three modes, whereas LIFT and a pyramid sensor would respectively need 291 and 310 photons. The ratio between LIFT and the SH is 3.3, which amounts to more than 1 star magnitude. Table 2 sums up the gains in star magnitude obtained with LIFT and the pyramid compared to the SH, for the case above and an ideal case with photon noise only.

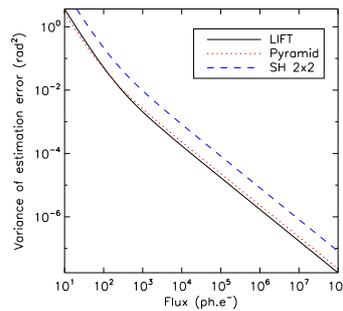


Fig. 6. Simulation of the total variance of tip/tilt/focus estimation error for LIFT, the SH 2×2 using a WCoG (with an unbiased Gaussian weighting) and the fixed pyramid. Read-out noise = $2 e^-$.

Table 2. Gain in star magnitude obtained with LIFT and the pyramid compared to the Shack-Hartmann for the estimation of tip/tilt/focus, with a read-out noise of $2 e^-$ and without read-out noise.

	LIFT	Pyramid
RON = $2 e^-$	1.3	1.2
RON = $0 e^-$	1.7	1.4

We were able to prove that LIFT is much more suited for low flux estimation than a Shack-Hartmann WFS. It has a general performance similar to a pyramid WFS, and is less sensitive to photon noise. In the following section, we validate our concept through lab experiments.

5. Experimental validation

The previously mentioned performance of LIFT assumes a perfect knowledge of the astigmatism offset, a point source and a perfectly linear camera with uniform read-out noise, gain and dark current, as well as a known sampling. Robustness to uncertainties on these parameters had to be demonstrated. We want to test the linearity of LIFT in narrow and wide spectral band, as well as the noise propagation. To do that, we must be able to control the flux, the aberrations, the read-out noise, the sampling and the spectrum of the source. We have therefore performed experiments in realistic conditions. In the following, we discuss the obtained results.

5.1. LIFT's test bed optical setup

Figure 7 describes the optical setup. The object is a single mode optical fiber connected to a laser diode light at 635 nm ($\Delta\lambda < 0.5$ nm). It can also be connected to a white light for wide band experiments. Two lenses collimate and then focus again the beam. Finally, we form the image with a CCD at Shannon sampling. The pupil is defined by a stop in the collimated beam, and the 0.5 rad astigmatism offset is inserted thanks to a tilted glass plate in the diverging beam coming from the source. The tip/tilt and focus are inserted by translating the camera along X, Y or Z. A Shack-Hartmann sensor was used to calibrate the astigmatism offset and the static aberrations. We neglect here differential aberrations.

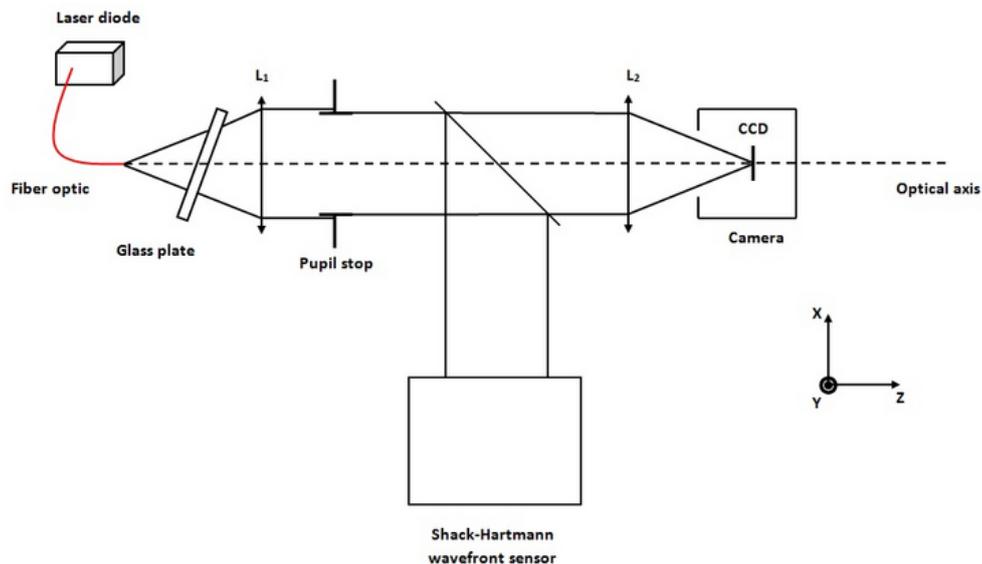


Fig. 7. Bench dedicated to the experimental validation of LIFT.

5.2. Linearity

5.2.1. Tip/tilt estimation

Without other aberrations, tip/tilt can easily be calibrated thanks to a center of gravity. This method is known to be effective at high flux, where the read-out noise becomes negligible [5]. The astigmatism offset does not bias the estimation, since the resulting spot is symmetrical with respect to both x and y axis. We thus formed images at high flux with different amounts of tip, in monochromatic light at 635 nm.

Figure 8 shows the tip estimated by LIFT as a function of the inserted tip, calibrated with a center of gravity. The solid line follows the equation $y=x$. The estimations given by LIFT and the center of gravity match each other over more than ± 2 rad. The maximum error is 0.04 rad (less than a $\lambda/100$ rms wavefront error) at -2.6 rad of tip.

This demonstrates that LIFT's estimation of tip is linear and accurate on a wide range.

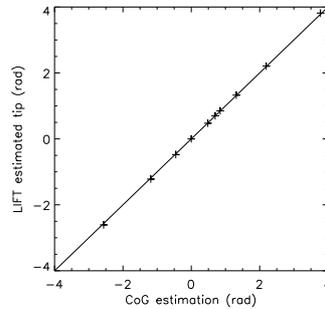


Fig. 8. Experimental estimation of a tip by LIFT versus a center of gravity estimation. Solid line: $y=x$.

5.2.2. Focus estimation

Monochromatic case The Shack-Hartmann WFS could also be used for the calibration of focus. At high flux, its estimation is reliable. Similarly to the previous experiment, we took images with different amounts of focus. The astigmatism offset measured by the Shack-Hartmann and used in our model is 0.16 rad on Z_5 and 0.43 rad on Z_6 . The phase offset is different from the optimum because of the experimental setup, which only allowed a rough adjustment. Nevertheless, as seen in Fig. 4, noise propagation coefficients do not vary significantly around the optimum. The combination of Z_5 and Z_6 is not far from 0.5 rad here, and is thus acceptable.

In Fig. 9(a), we have plotted the estimation of LIFT versus the Shack-Hartmann estimation. The dotted line is $y=x$ and the solid line is a simulation. We globally observe a good linearity, with a maximum error of 0.15 rad ($\sim \lambda/40$ rms).

The error is symmetrical with respect to 0 and can be explained by a difference between the astigmatism seen on the detector and the one used in our model. The data acquired for this experiment also allowed to apply a phase diversity algorithm [14], so as to obtain the aberrations on the imaging camera. By performing this method, we found that differential aberrations added 0.1 rad on both Z_5 and Z_6 . On this last mode, this means a model error of $\sim 25\%$. Figure 9(b) shows the estimation with correction of the astigmatism in the model. The maximum error is then 0.04 rad ($\sim \lambda/160$ rms).

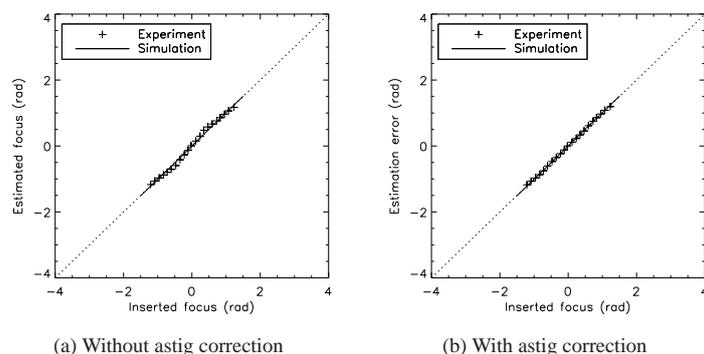


Fig. 9. Experimental estimation of a focus by LIFT with monochromatic light. In (a), the astigmatism offset used in the model is the one measured by the Shack-Hartmann. In (b), the model is corrected by adding the differential aberrations found by phase diversity.

The focus estimation is thus accurate in the range $[-1, 1]$ rad, even with a significant model error on the astigmatism offset.

Polychromatic case The estimation of focus was also made with a polychromatic source. Its FWHM is 200 nm for a central wavelength at 676 nm ($R=3.4$). As seen in section 2, LIFT can make an estimation in polychromatic light by summing the contributions from the different wavelengths. But simulations have shown that assuming solely a contribution of the central wavelength, it still gives a good performance [17], as long as the spectral bandwidth is narrow enough. Moreover, we checked that the coefficients of noise propagation do not vary significantly in this case, and we thus have the same optimum as in monochromatic light. We thus estimated the focus in this configuration.

The astigmatism offset measured by the Shack-Hartmann and used in our model is 0.11 rad on Z_5 and 0.5 rad on Z_6 at the central wavelength. The inserted focus in radians was also computed from the Shack-Hartmann measurements at the central wavelength.

Figure 10 shows the source spectrum and the result of the estimation. As for monochromatic light, a linear estimation is observed. The maximum error goes up to 0.12 rad ($\sim \lambda/50$ rms) in negative foci, and 0.03 rad ($\sim \lambda/200$ rms) in positive foci. The higher error for negative values is due to lateral chromatic aberration coming along with the tilted plate, which was not taken into account in our model. Again, the robustness of LIFT is demonstrated here.

As a conclusion, the estimation of LIFT in wide band can be as accurate as in monochromatic light, with no additive calculations.

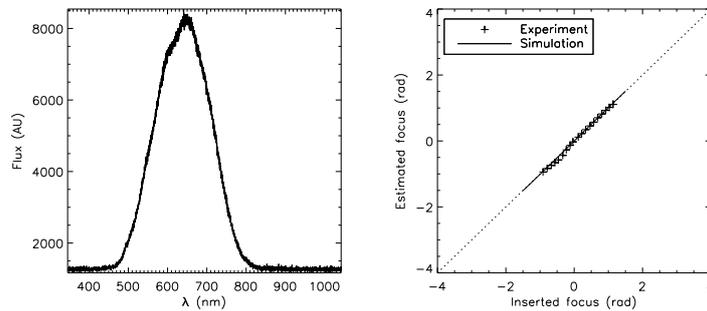


Fig. 10. Left: Spectrum of the source used in the estimation with large bandwidth light. Right: Experimental estimation of a focus by LIFT in large bandwidth.

5.3. Noise propagation

We intend here to validate the simulations shown in section 5.1 by observing the evolution of the estimation error variance with the flux. To get interpretable data, where we could see both the read-out noise and the photon noise regimes, we changed the conditions of experimentation. The astigmatism was raised to 1 rad to increase α and β . The equivalent read-out noise in this experiment is 14 ph.e^- . The estimation is made on a zone of 20×20 pixels.

We took several series of images at different flux and computed an empirical variance for tip/tilt and focus (Fig. 11). Given the conditions of experiment (sampling, phase offset...), we are able to find the noise propagation coefficients α and β thanks to analytical calculations (see section 3.3.2). Once we have these coefficients, we can compute a theoretical variance of estimation error at any flux.

The theoretical variance and the empirical one match over several decades of flux, confirming our model of noise propagation.

A saturation can be seen on both figures. It is estimated from the value at the highest flux, and is equal to $5.9 \pm 3.4 \times 10^{-5} \text{ rad}^2$ for tip and $7 \pm 3.3 \times 10^{-7} \text{ rad}^2$ for focus. We checked that the ratio between these saturations is compatible with vibrations of the bench.

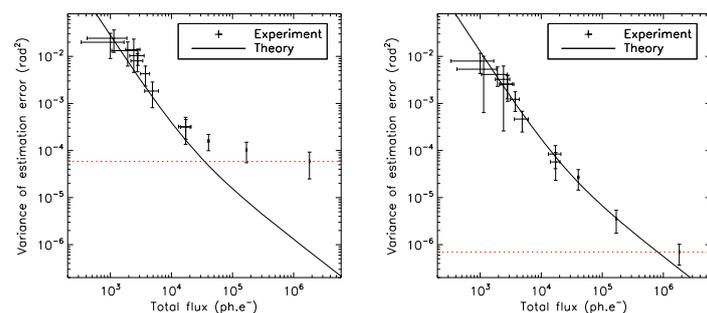


Fig. 11. Experimental variance of estimation error for tip/tilt (left) and focus (right). RON = 14 ph.e^- . Red dotted line: saturation due to vibrations.

6. Conclusion

We detailed here the concept of LIFT. An astigmatism offset is added to the incoming phase to remove the indetermination on focus, then a maximum likelihood algorithm is performed on the focal image. Iterations compensate for the approximation of small phase. A theoretical study of the minimization criterion and noise propagation has let us understand better the algorithm and justify the choice of an astigmatism offset rather than focus, as well as the choice of the offset amplitude.

We have demonstrated that LIFT is robust to setup uncertainties and provides a linear focus estimation, even with large spectral bandwidth light ($R=3.4$). The validation of noise propagation confirms the superiority of LIFT over a Shack-Hartmann 2×2 : LIFT has more than a 1 magnitude gain over a Shack-Hartmann 2×2 for a variance of 0.01 rad^2 and a read-out noise of $2 e^-$. This performance is similar to what one could have with a fixed pyramid sensor. Moreover, LIFT is light in hardware, easy to set up, and thanks to its algorithm, automatically adapts the weighting on pixels depending on noise properties. Its offset phase and sampling can be tuned to match most of systems. Finally, it relies on well sampled focal images that can be interpreted easily, which eases the debugging stage.

LIFT has already been tested on an operational system: Gemini Southern Observatory [25]. An on-sky validation is planned for spring 2013. Also, a deeper study of the algorithm will be made to get a further comparison with the pyramid and other wavefront sensors. However, LIFT can already be considered as an effective and robust low-order WFS in the context of wide field adaptive optics for very large and extremely large telescopes.

Acknowledgments

The authors would like to thank Bruno Fleury for his help on understanding some experimental issues, and Baptiste Paul for helping in the setup of the bench. The work was partly funded by the European Commission under FP7 Grant Agreement No. 312430 Optical Infrared Coordination Network for Astronomy.

2.3.3 Bilan et discussion

Bilan Dans cet article, nous avons procédé à l'optimisation des paramètres de LIFT, et validé le concept en laboratoire.

Dans un premier temps, nous avons cherché une phase de diversité optimale. Deux modes ont été considérés : le focus et l'astigmatisme. En observant la forme que prend le critère à minimiser pour l'estimation d'un focus, nous avons constaté l'apparition d'un minimum secondaire lorsque l'on choisit un focus comme phase de diversité. L'utilisation d'un focus de diversité ne lève donc l'indétermination de l'estimation du focus que sur un intervalle limité. L'astigmatisme, au contraire, conduit à un minimum unique, et s'est imposé comme mode de diversité. L'amplitude optimale d'astigmatisme a ensuite été trouvée grâce au calcul des coefficients de propagation du bruit dans LIFT. Pour un astigmatisme de 0,5 rad, ces coefficients sont minimaux.

A partir de cette étude de la propagation du bruit, nous avons comparé LIFT au Shack-Hartmann 2×2 et à la pyramide pour l'estimation du tip/tilt/focus. Les propagations ont été évaluées pour une estimation de maximum de vraisemblance et autour d'une phase nulle, ce qui est équivalent à une évaluation de la propagation via la matrice de Fisher (voir §1.4.6 p.60). Nous montrons ainsi que LIFT, comme la pyramide, permet de gagner plus d'une magnitude par rapport au Shack-Hartmann, en ayant l'avantage d'une installation simple. Une erreur, qui n'a pas d'impact sur cette comparaison, s'est introduite sur le coefficient de propagation du bruit de détecteur pour le Shack-Hartmann utilisant le centre de gravité simple. Dans la partie 2.4 p.89, nous reprenons ce résultat et le corrigeons.

D'autre part, la linéarité de l'estimation du tip/tilt et du focus a été validée expérimentalement, avec une erreur sur le front d'onde inférieure à $\lambda/100$. Nous avons de plus vérifié la robustesse de l'estimation par LIFT lorsque qu'il est soumis à une large bande spectrale ($\lambda/\Delta\lambda = 3, 4$). Enfin, nous avons montré que notre modèle de propagation de bruit évaluait correctement la variance d'erreur d'estimation sur une grande plage de rapports signal-à-bruit.

Discussion Dans l'article, nous avons d'abord voulu optimiser l'amplitude d'astigmatisme en cherchant les valeurs pour lesquelles le critère est très pentu autour de son minimum. Cela mène en effet à une bonne convergence de l'algorithme. Finalement, c'est l'étude de la propagation du bruit qui a permis d'obtenir une amplitude précise.

Cependant, ces deux aspects sont liés par l'information de Fisher. L'information de Fisher peut se calculer à partir de la courbure de la log-vraisemblance (cf §1.4.6 p.60). Au niveau d'un extremum, plus la courbure est élevée en valeur absolue, plus la log-vraisemblance est pentue autour de cet extremum. Or, le critère à minimiser est directement issu de la log-vraisemblance (cf §2.2 p.66). Regarder la pente du critère autour du minimum revient alors à regarder la courbure de la log-vraisemblance, donc l'information de Fisher. D'autre part, comme nous l'avons montré dans la partie 1.4.6 p.60, les coefficients de propagation de bruit déterminés pour LIFT sont directement égaux aux coefficients diagonaux de son information de Fisher. Les recherches de l'amplitude d'astigmatisme optimale via la forme du critère et via la propagation du bruit sont donc finalement équivalentes. En l'occurrence, le critère a été seulement calculé par rapport à l'estimation du focus. L'information était donc incomplète. C'est pourquoi l'étude de la propagation du bruit, qui prenait aussi en compte l'estimation du tip/tilt, a été plus efficace pour déterminer l'amplitude optimale.

Cette publication ne fait pas mention de l'étude de robustesse en simulation qui a permis de

préparer la validation expérimentale. Le paragraphe suivant montre les résultats de cette étude.

2.3.4 Étude de robustesse

La répartition du flux dans les pixels est liée à l'échantillonnage du système ainsi qu'à l'astigmatisme introduit. Or, l'estimation s'appuie sur cette répartition. Il est donc important de vérifier quelle erreur peut être admise sur ces paramètres dans le modèle utilisé par LIFT.

2.3.4.1 Robustesse à une méconnaissance sur l'échantillonnage

Imaginons que les pixels soient plus petits que la taille attendue (on a sous-évalué l'échantillonnage). Un tip/tilt sera alors vu par un déplacement de la tache sur plus de pixels que prévu. De même, si l'image est défocalisée, la tache sera répartie sur plus de pixels. Une sous-évaluation de l'échantillonnage entraîne donc une surestimation du tip/tilt et du focus, en valeur absolue. Avec le raisonnement inverse, on déduit qu'une surévaluation de l'échantillonnage entraîne au contraire une sous-estimation de ces trois modes.

S'il n'y a pas d'erreur sur l'échantillonnage, l'estimation est linéaire, avec un coefficient de linéarité unitaire. Avec une mauvaise évaluation de l'échantillonnage, on peut donc s'attendre à avoir un coefficient de linéarité différent de 1. La figure 2.3 montre le coefficient de linéarité en fonction de l'échantillonnage de l'image. L'échantillonnage vrai est à 2 Shannon, et on fait varier l'échantillonnage entré comme paramètre dans le modèle de LIFT.

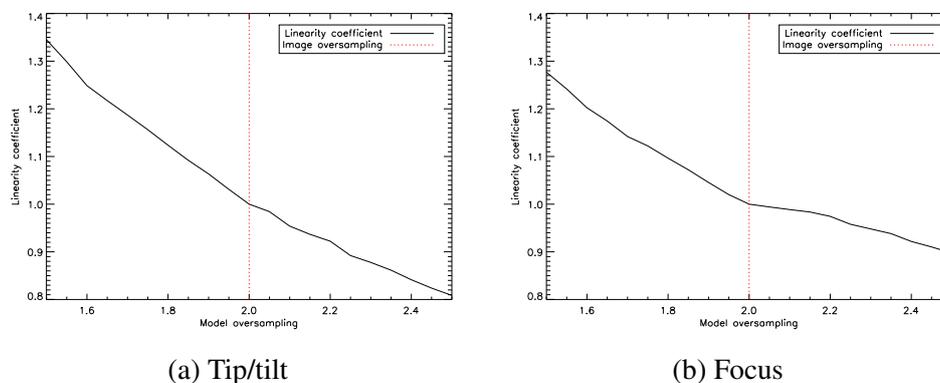


FIGURE 2.3 – Coefficient de linéarité en fonction de l'échantillonnage entré comme paramètre dans le modèle de LIFT. L'image est échantillonnée à 2 Shannon.

Une erreur d'échantillonnage de 5% entraîne une erreur d'estimation d'environ 2-3% sur chaque mode. Cela correspond à une erreur d'environ 0,01 rad pour un mode à 0,5 rad, donc une variance d'erreur de 0,0003 rad² sur les trois modes. En pratique, il est peu probable que la variance d'erreur d'estimation due au bruit descende en dessous de 0,01 rad², à cause du faible flux reçu. L'échantillonnage peut facilement être évalué avec une précision inférieure ou égale à 5%, LIFT sera donc robuste à une méconnaissance de l'échantillonnage.

2.3.4.2 Robustesse à une méconnaissance sur l'astigmatisme

La défocalisation d'une tache astigmatique entraîne l'allongement de cette tache dans une direction particulière. Plus on est défocalisé, plus la tache est longue. De la même manière, un allongement de la tache est obtenu si, pour une amplitude de focus donné, on augmente l'amplitude de l'astigmatisme. Ainsi, une sous-évaluation de l'astigmatisme entraîne une surestimation du focus, et inversement. L'effet sur le tip/tilt est négligeable, puisque l'astigmatisme change seulement la forme de la tache, et de manière symétrique.

On peut donc s'attendre à ce qu'une erreur sur l'astigmatisme introduit entraîne un changement de coefficient de linéarité pour le focus. La figure 2.4 montre le coefficient de linéarité en fonction de l'astigmatisme introduit dans l'image. L'astigmatisme vrai vaut 0,5 rad, et on fait varier l'astigmatisme entré comme paramètre dans le modèle de LIFT.

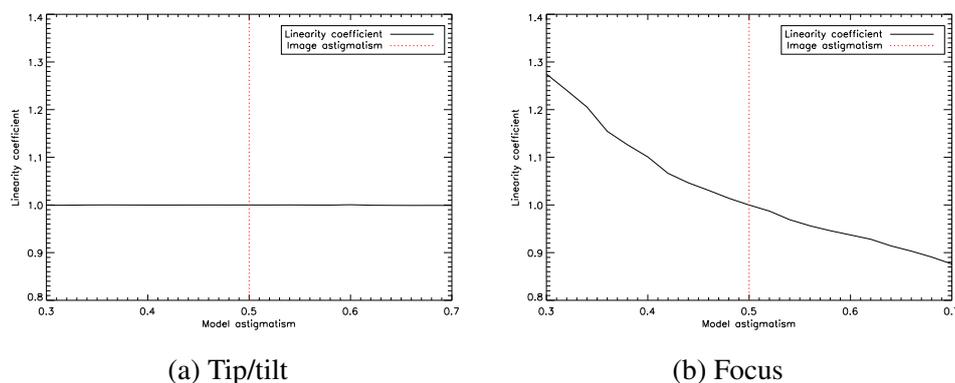


FIGURE 2.4 – Coefficient de linéarité en fonction de l'astigmatisme entré comme paramètre dans le modèle de LIFT. L'astigmatisme vrai vaut 0,5 rad.

On a bien vérifié que l'erreur d'estimation était négligeable pour le tip/tilt, malgré une grande erreur sur l'astigmatisme. En calibrant l'astigmatisme avec un Shack-Hartmann ou la diversité de phase à fort flux, on peut atteindre facilement une précision de l'ordre de 0,01 rad ($\sim \lambda/600$). Sachant qu'une erreur de 10% (0,05 rad) sur l'astigmatisme entraîne une erreur d'estimation d'environ 0,02 rad ($\sim \lambda/300$) pour un focus à 0,5 rad, on peut affirmer que LIFT est robuste à une erreur sur l'astigmatisme introduit.

2.4 Comparaison de LIFT avec d'autres analyseurs

Dans cette partie, nous proposons une comparaison de LIFT avec d'autres analyseurs en calculant leurs coefficients de propagation du bruit à l'aide de l'information de Fisher (voir §1.4.6 p.60). Dans chaque cas, on peut faire ce calcul à partir de pseudo-données (pentes du Shack-Hartmann par exemple) ou des données brutes, c'est-à-dire l'image sur le détecteur.

Une première comparaison de LIFT avec le Shack-Hartmann 2×2 et une pyramide fixe a été effectuée dans l'article présenté dans la partie 2.3 p.70. Les coefficients de propagation du bruit ont été calculés à partir de la matrice de covariance d'erreur d'un estimateur de maximum de vraisemblance (cf §1.4.4 p.59, §1.4.5 p.60). D'après le paragraphe 1.4.7 p.61, ces coefficients sont équivalents à ceux que l'on aurait obtenus par l'information de Fisher. Le calcul a été

fait sur des pseudo-données, sauf dans le cas de LIFT, qui fait son estimation directement sur l'image.

Ici, nous refaisons cette comparaison en appliquant le calcul des coefficients de propagation sur les données brutes. Nous ajoutons aussi la diversité de phase et le masque de Zernike. Nous décrivons tout d'abord les paramètres utilisés pour les simulations, puis nous comparons les coefficients obtenus.

2.4.1 Paramètres de simulation

Les paramètres de simulation des analyseurs sont présentés ici. Nous ne rappelons pas leurs principes, ni les méthodes de simulation, qui sont décrits dans la partie 1.4 p.46. Lorsque l'on considérera des données brutes, nous prendrons toujours le modèle de bruit gaussien décrit au paragraphe 1.2.1.1 p.22.

2.4.1.1 LIFT et diversité de phase

On considère ici un échantillonnage est à Shannon. On simule LIFT pour une amplitude d'astigmatisme de 0,5 rad, et la diversité pour des amplitudes de focus de 1,8 rad et 0,5 rad. L'image entière (128×128 pixels) est prise en compte dans tous les cas.

2.4.1.2 Shack-Hartmann

Les simulations sont faites pour un Shack-Hartmann avec 2×2 sous-pupilles carrées. L'échantillonnage est à Shannon, avec 16×16 pixels par sous-pupille. On suppose qu'on est à la limite de diffraction, les variances sur les pseudo-données sont donc calculées pour $N_t = N_d = N_w = 2$ pixels. Nous prenons $N_w = N_t$ car c'est le réglage optimal pour le centre de gravité pondéré [Thomas *et al.*, 2006]. N_s est arbitrairement fixé à 4 pixels, qui est un compromis entre la sensibilité au bruit de détecteur et une dynamique correcte pour le centre de gravité. Les données brutes correspondent à l'image complète au plan focal des micro-lentilles, comprenant toutes les imagettes.

2.4.1.3 Pyramide

La pupille est échantillonnée sur 2×2 pixels. Nous simulons une pyramide sans modulation. Notons que pour cette configuration, l'approximation d'un bruit négligeable sur le dénominateur est tout à fait discutable. Nous verrons cependant que l'approximation est correcte dans notre cas. Nous considérerons deux types de données brutes : l'image complète ou seulement les pixels des 4 images de pupille.

2.4.1.4 Masque de Zernike

La pupille est échantillonnée sur 4×4 pixels. En effet, comme l'image donne une estimation directe de la phase, c'est le minimum requis pour estimer un focus. Seuls les pixels compris dans l'image de la pupille sont pris en compte dans les données brutes.

2.4.2 Coefficients de propagation du bruit

Les coefficients de propagation mode à mode obtenus pour l'estimation du tip/tilt et du focus sont présentés dans le tableau 2.1. Ces coefficients sont calculés autour d'une phase nulle et pour une lumière monochromatique.

	tip/tilt ($\alpha_2 = \alpha_3$)	focus (α_4)	tip/tilt ($\beta_2 = \beta_3$)	focus (β_4)	$\alpha = \sum \alpha_i$	$\beta = \sum \beta_i$
LIFT	0,59	0,53	16	54	1,71	87
LIFT	0,47	0,51	16	54	1,45	87
Diversité (1,8 rad)	0,46	1,42	36	526	2,34	598
Diversité (0,5 rad)	0,34	1	23	64	1,68	110
Pyramide	0,87	0,56	23	16	2,3	62
Pyramide (pixels pupille)	0,81	0,73	23	15	2,35	61
Pyramide (image totale)	0,63	0,67	23	15	1,93	61
Shack-Hartmann (CdG)	2,03	0,55	240	65	4,61	545
Shack-Hartmann (CdGP)	3,6	0,98	147	40	8,19	334
Shack-Hartmann (image totale)	1,26	0,42	139	44	2,94	322
Masque de Zernike	1,29	1,1	33	28	3,68	94
Masque de Zernike (approximation)	0,27	0,37	4	5,5	0,91	13,5

TABLE 2.1 – Coefficients de propagation du bruit mode à mode des différents analyseurs considérés pour l'estimation du tip/tilt et du focus. α_i : coefficient de propagation du bruit de photons pour le mode i . β_i : coefficient de bruit de détecteur pour le mode i . Les coefficients correspondant aux pseudo-données et provenant de l'article (cf §2.3 p.70) sont montrés en rouge. CdG : centre de gravité. CdGP : centre de gravité pondéré. Les coefficients β_i pour le centre de gravité ont été corrigés car l'article comportait une erreur.

On rappelle que plus les coefficients de propagation du bruit sont petits, moins l'erreur d'estimation est grande. On peut tout d'abord voir que LIFT est très compétitif vis-à-vis des autres analyseurs, que ce soit en bruit de photons ou bruit de détecteur. La seule exception est le masque de Zernike lorsqu'on fait l'approximation que le signal est proportionnel à la phase. La simulation montre que cette approximation n'est pas valide pour le calcul de la propagation de bruit. En effet, seulement la moitié du flux est utile pour le signal à cause des effets de diffraction, et le signal obtenu n'est pas exactement proportionnel à la phase. LIFT est aussi légèrement moins bon que la pyramide en bruit de détecteur.

Si l'on compare ces résultats avec les coefficients obtenus dans l'article reproduit dans la partie 2.3 p.70, on se rend compte que LIFT est ici légèrement meilleur en bruit de photons. Cela est simplement dû au fait que nous avons pris en compte toute l'image ici, plutôt que la fenêtre de calcul. La pyramide, lorsqu'on considère les pixels dans les images de pupilles uniquement, a une propagation quasi-identique à celle dans l'article. Cela montre que les pseudo-données utilisées pour la pyramide extraient bien toute l'information. Le fait que la correspondance ne soit pas tout à fait exacte est sûrement dû aux approximations faites sur le modèle de bruit des pseudo-données. La performance en bruit de photons peut être améliorée en utilisant toute l'image, mais cela est un peu artificiel : le flux entre les images de pupille est soumis aux flux parasites tels que la diffusion par les arêtes imparfaites de la pyramide, et sont

donc difficilement exploitables en pratique. Enfin, le Shack-Hartmann affiche une meilleure performance, notamment en bruit de photons. En effet, les imagerie contiennent plus d'informations qu'un simple centre de gravité. A partir de l'image totale du Shack-Hartmann, on pourrait reconstruire le front d'onde beaucoup plus précisément. C'est ce qu'a montré Cannon [Cannon, 1995]. Seulement, cela requiert un temps de calcul beaucoup plus long, et n'est pas facilement compatible avec la cadence requise en optique adaptative.

Enfin, la diversité est évidemment très similaire à LIFT pour une amplitude de diversité équivalente (0,5 rad). Elle a une performance globale légèrement moins bonne à cause de la division du flux entre les deux images. L'amplitude de diversité habituelle, 1,8 rad, est optimisée pour l'estimation de plusieurs dizaines de modes, et n'est pas efficace pour l'estimation du tip/tilt et du focus uniquement.

Finalement, en considérant les données brutes, LIFT est le meilleur analyseur en bruit de photons, avec notamment un gain de plus de 1,5 sur la pyramide, de 2 sur le Shack-Hartmann et d'environ 2,5 sur le masque de Zernike. En bruit de détecteur, il perd un facteur 1,5 par rapport à la pyramide, malgré un gain supérieur à 3 par rapport au Shack-Hartmann. Cependant, LIFT présente l'avantage d'être simple à installer et à utiliser. LIFT semble donc être le meilleur choix d'analyseur dédié à l'estimation de bas ordres sur étoile naturelle.

2.5 Conclusion

Dans ce chapitre, nous avons présenté l'optimisation de LIFT et sa validation expérimentale. Sa compétitivité vis-à-vis des autres analyseurs en termes de sensibilité au bruit a aussi été démontrée.

La première étape de l'optimisation a été l'analyse du critère qui est minimisé dans LIFT. J'ai d'abord montré que l'estimation du focus était plus efficace en introduisant un astigmatisme de référence plutôt qu'un focus. J'ai déterminé un intervalle d'amplitudes d'astigmatisme pour lesquelles l'algorithme convergeait plus facilement. Grâce à une quantification de la propagation de bruit dans LIFT, j'ai affiné cet intervalle et j'ai défini une amplitude d'astigmatisme optimale pour limiter la sensibilité au bruit. Cette étude a aussi permis une première comparaison de LIFT avec d'autres analyseurs. Le gain de LIFT par rapport au Shack-Hartmann 2×2 est supérieur à 1 magnitude et sa performance est comparable à celle de la pyramide, en ayant l'avantage d'une installation et d'une utilisation simples.

Afin de préparer la validation expérimentale, je me suis assuré du bon fonctionnement de LIFT malgré des méconnaissances sur l'échantillonnage et l'astigmatisme grâce à une étude en simulations. Les tests sur banc ont ensuite confirmé la linéarité de l'estimation, en monochromatique ainsi qu'en large bande spectrale, et le modèle de propagation de bruit.

Enfin, j'ai repris la comparaison de la propagation de bruit dans LIFT avec d'autres analyseurs pour l'estimation du tip/tilt et du focus, en m'appuyant sur la matrice de Fisher. J'ai ainsi confirmé les résultats de la première étude de propagation de bruit. J'ai aussi considéré deux autres analyseurs : la diversité de phase et le masque de Zernike. J'ai montré que LIFT est légèrement meilleur que la diversité de phase grâce à l'utilisation d'une seule image plutôt que deux. D'autre part, le masque de Zernike s'est révélé peu performant en bruit de photons, notamment à cause des effets de diffraction, et LIFT offre un gain de 2,5 par rapport à cet analyseur. LIFT se révèle donc finalement très compétitif par rapport aux autres analyseurs considérés.

Cependant, en situation réelle, LIFT sera soumis aux hauts ordres qui n'ont pas été entièrement corrigés par l'optique adaptative. Ces hauts ordres résiduels peuvent induire une erreur sur l'estimation. On peut alors se demander si LIFT fonctionnera correctement dans un système opérationnel. Dans le chapitre suivant, nous présenterons la première validation sur ciel de LIFT, puis nous étudierons les effets des hauts ordres ainsi que les pistes d'amélioration pour optimiser la performance de LIFT dans une boucle fermée.

Chapitre 3

Vers une utilisation de LIFT dans une optique adaptative grand champ

3.1 Introduction

LIFT a été validé en laboratoire, sur des phases statiques ne présentant pas de hauts ordres. Dans un cas réel, LIFT sera utilisé pour estimer les bas ordres sur une étoile naturelle dans un système d'optique adaptative assisté par laser. Les hauts ordres seront déjà pré-correctés grâce à l'analyse sur les étoiles lasers. Cependant, la correction de l'optique adaptative n'est pas parfaite (cf §1.3.3 p.33). Ainsi, les images d'étoile naturelle vues par LIFT présentent des résidus de hauts ordres turbulents qui pourront altérer l'estimation des bas ordres. Dans ce chapitre, nous étudions le comportement de LIFT en présence de hauts ordres résiduels typiques d'une optique adaptative grand champ. Les termes du budget d'erreur qui vont nous intéresser sont l'erreur de repliement et la propagation du bruit sur le tip/tilt et le focus (cf §1.3.3 p.33).

Repliement On peut distinguer deux types de repliement, que l'on nomme "repliement physique" et "repliement de sous-modélisation". Le repliement physique peut être observé lorsque les informations provenant des basses fréquences et des hautes fréquences sont indiscernables dans les données. Il s'agit d'un repliement fondamental, qui ne peut être atténué qu'en changeant la configuration de l'analyseur. Dans un Shack-Hartmann, par exemple, la pupille est divisée en sous-pupilles et une pente moyenne du front d'onde est estimée dans chaque sous-pupille (cf §1.4.1.1 p.46). Le front d'onde global est reconstruit à partir de ces pentes locales. Le Shack-Hartmann ne peut donc pas "résoudre" des variations plus petites que la taille des sous-pupilles. Ainsi, les variations de phase de haute fréquence au sein d'une sous-pupille interviennent dans la mesure de pente, mais ne peuvent pas être reconstruites à partir de cette même mesure. Le repliement de sous-modélisation apparaît lorsqu'on restreint le nombre de modes estimés. On a alors un modèle qui ne tient pas compte de l'influence des hauts ordres même s'ils peuvent être estimés. On dit que le problème est sous-modélisé, d'où le nom de ce type de repliement.

Propagation du bruit La présence de hauts ordres entraîne également un élargissement de la tache. En effet, les hautes fréquences spatiales de la phase forment des speckles éloignés du centre de la tache image. Par conséquent, le flux est dispersé sur les pixels et le rapport

signal-à-bruit est dégradé (Fig. 3.1).

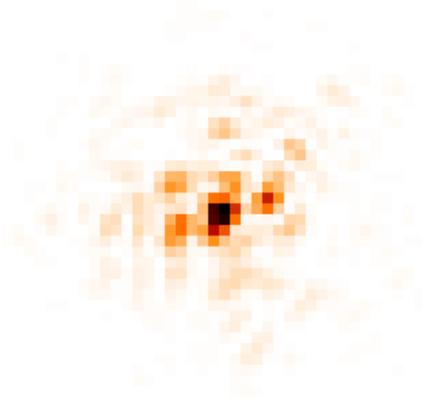


FIGURE 3.1 – Tache présentant des résidus de hauts ordres turbulents.

Ces effets se traduisent par une erreur de reconstruction de la phase plus importante. Cependant, des simulations faites dans le cadre d'ATLAS ont montré que LIFT permettait la correction du tip/tilt et du focus malgré la présence de hauts ordres [Meimon *et al.*, 2010a]. Afin de confirmer expérimentalement ces simulations, nous avons effectué la première validation en boucle ouverte de LIFT sur ciel grâce au système d'optique adaptative multiconjuguée du télescope Gemini Sud, GeMS (§3.2 p.96). Après quelques tests de préparation sur des sources de calibration interne, nous montrons les résultats d'estimation du focus par LIFT sur une étoile partiellement corrigée, qui ont fait l'objet d'une publication reproduite dans ce chapitre (§3.2.3 p.100).

Dans un deuxième temps (§3.3 p.110), nous étudions le repliement sur l'estimation du tip/tilt/focus, et nous proposons une piste d'amélioration de l'algorithme afin de minimiser son impact sur le budget d'erreur. Ensuite, en simulant des hauts ordres résiduels typiques du système GeMS, nous déterminons l'erreur due à la propagation du bruit en fonction des conditions de turbulence. Nous montrons enfin la première simulation d'une fermeture de boucle, et faisons une comparaison qualitative avec une récente variation de LIFT.

Pour finir, nous discutons des perspectives pour de futurs tests et pour l'utilisation de LIFT dans un système d'optique adaptative tomographique (§3.4 p.128).

3.2 Validation sur ciel en boucle ouverte avec le télescope Gemini Sud

Nous détaillons ici le déroulement des tests effectués sur le télescope Gemini Sud, qui ont mené à la première validation sur ciel de LIFT en boucle ouverte. Nous présentons tout d'abord le système d'optique adaptative du télescope Gemini Sud, GeMS (§3.2.1 p.97). Nous décrivons ensuite les premiers tests, sur source de calibration, qui ont permis de préparer les tests sur ciel (§3.2.2 p.97). Enfin, nous montrons la validation de la linéarité de l'estimation du focus par LIFT sur une étoile naturelle, réalisée le 31 janvier 2013. (§3.2.3 p.100).

3.2.1 Présentation du système GeMS

GeMS (*Gemini Multi-conjugated adaptive optics System*) est le système d'optique adaptative du télescope Gemini Sud. Il fournit des images en proche infrarouge avec une résolution proche de la diffraction sur un champ de plus d'1 minute d'arc de côté. Le système inclut 5 étoiles guides laser de type sodium, 3 étoiles guides naturelles, 3 miroirs déformables conjugués à 0, 4,5 et 9 km et 1 miroir de tip-tilt [Rigaut *et al.*, 2013]. Au moment de nos tests, le miroir conjugué à 4,5 km était hors-service.

Le banc d'optique adaptative de GeMS, nommé Canopus (Fig. 3.2), a été conçu pour l'optique adaptative multi-conjuguée. Un miroir dirige la lumière collectée par le télescope vers le banc (ouverture en haut à droite sur la figure). Le faisceau est ensuite collimaté par un miroir parabolique hors-axe sur les trois miroirs déformables (d'abord le 9 km, puis le 4,5 km, puis le 0 km) et le miroir de tip-tilt. Une lame séparatrice, dite "scientifique", transmet la lumière infrarouge à un correcteur de dispersion atmosphérique. Enfin, le faisceau corrigé est refocalisé à $f/32$ par un autre miroir parabolique hors-axe et sort du banc pour aller vers la caméra scientifique [Rigaut *et al.*, 2013] (ouverture au milieu à droite sur la figure).

La lumière utilisée pour l'analyse de front d'onde est réfléchiée par la lame scientifique. La lumière à 589 nm provenant des 5 étoiles laser est réfléchiée par une deuxième lame séparatrice, dite "laser", et envoyée vers les analyseurs de front d'onde dédiés aux étoiles laser : cinq Shack-Hartmann 16×16 alignés sur les 5 faisceaux laser. Des composants motorisés permettent un contrôle actif du focus.

La lumière visible des trois étoiles guides naturelles passe de même dans un correcteur de dispersion atmosphérique et arrive sur les analyseurs de front d'onde dédiés aux étoiles naturelles. Ces analyseurs sont au nombre de 3 et sont indépendants. Ils sont motorisés pour se positionner sur leurs étoiles respectives. Chaque analyseur possède quatre fibres reliées à des photodiodes à avalanche à la manière d'un analyseur à quatre cadrants. L'analyseur numéro 3 envoie 25% de la lumière qu'il reçoit vers un Shack-Hartmann 2×2 qui estime le focus à une fréquence faible (~ 1 Hz).

Les tests suivants ont été effectués en bande Ks ($\lambda_0 = 2.2 \mu m$, $\Delta\lambda = 0.32 \mu m$), tout d'abord sur des sources de calibration, puis sur une étoile naturelle. Les images sont prises sur le détecteur scientifique, nommé GSAOI [McGregor *et al.*, 2004]. Il s'agit d'une matrice de 2×2 détecteurs Rockwell HAWAII-2RG.

Chacun des détecteurs de GSAOI peut être lu sur une région d'intérêt uniquement, nommée On Detector Guide Window (ODGW) [Young *et al.*, 2012]. Les ODGW permettent d'estimer le tip-tilt à partir de 4 étoiles naturelles. Elles peuvent être lues à une fréquence allant jusqu'à 800 Hz, et avoir une taille de 2×2 à 128×128 pixels. La taille du pixel est 20 mas.

Ce système est donc parfait pour une validation de LIFT. D'une part, on peut détecter des images dans l'infrarouge, ce qui permet de bénéficier d'une bonne correction. En effet, la force de la turbulence et l'amplitude de la phase diminuent toutes les deux avec la longueur d'onde. D'autre part, on peut sélectionner l'étoile à analyser sur une petite fenêtre qui peut être lue rapidement.

3.2.2 Tests sur source de calibration

Afin de préparer la validation sur ciel, des tests préliminaires ont été effectués avec les sources de calibration du banc Canopus. Dans un premier temps, nous avons vérifié la linéarité



FIGURE 3.2 – Le banc Canopus.

de l'estimation du focus en l'absence de hauts ordres, pour s'assurer de la bonne prise en main du système. Nous avons ensuite simulé une turbulence résiduelle cyclique dans le but de démontrer la répétabilité de l'estimation du focus en présence de hauts ordres, et ainsi confirmer un comportement normal de l'analyseur dans ces conditions.

3.2.2.1 Validation de la linéarité

Les pentes de référence des analyseurs dédiés aux étoiles laser peuvent être changées afin d'ajouter un focus ou un astigmatisme connu. D'autre part, la boucle d'optique adaptative peut être gelée. Les images obtenues sont alors très stables. Pour confirmer la linéarité de l'estimation de LIFT, nous avons introduit l'astigmatisme de diversité de 0,5 rad, puis différentes amplitudes de focus, de -500 nm (1,4 rad) à 200 nm (0,6 rad). Pour chaque valeur de focus, nous avons acquis une série d'images.

L'estimation de LIFT en fonction du focus introduit est présentée sur la figure 3.3. L'écart-type de l'estimation est de 3 nm (0,01 rad) pour chaque série. De -100 nm à 200 nm, les estimations sont parfaitement égales aux valeurs introduites, et on constate une erreur de 10 nm (0,03 rad) à -200 nm et -500 nm. Cette erreur doit être due à une légère différence dans la calibration du focus entre ces séries et les autres.

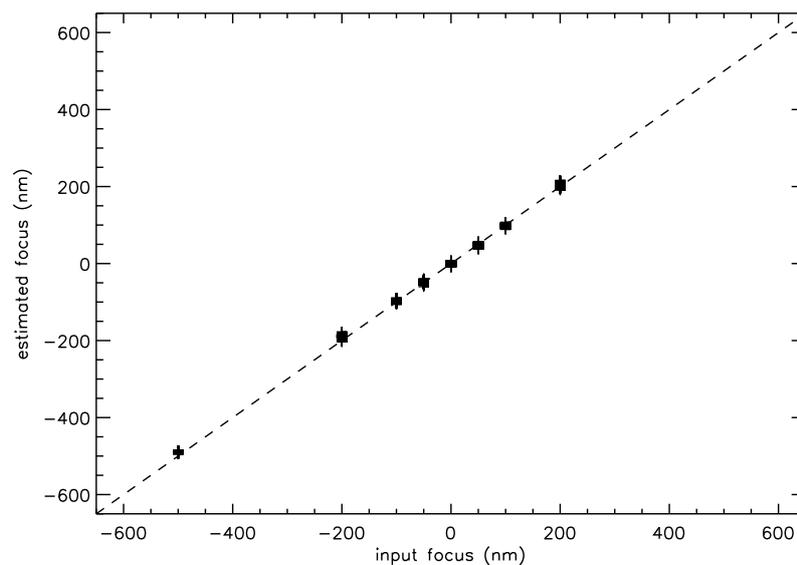


FIGURE 3.3 – Validation de la linéarité de l'estimation du focus sur les sources de calibration de Gemini. Chaque estimation est représentée par une croix. Pour chaque focus, il y a une série d'estimations.

Nous avons donc redémontré la précision de l'estimation d'un focus jusqu'à une amplitude de 1,4 rad sur les sources de calibration d'un système opérationnel.

3.2.2.2 Répétabilité de l'estimation

Afin de s'assurer de la stabilité de l'estimation du focus, nous avons acquis des images en présence d'une turbulence simulée. Une fois la boucle gelée, nous avons envoyé de manière

répétée un cycle de tensions perturbatives sur le miroir 0 avec une période de 10 secondes. Ces tensions étaient représentatives d'une turbulence résiduelle après correction par l'optique adaptative, et contenaient toutes les fréquences spatiales que peut reproduire le miroir déformable. Cette configuration reproduit donc, à l'erreur de sous-modélisation du miroir près, les conditions opérationnelles de LIFT.

L'estimation du focus sur les images résultantes est tracée en fonction du temps sur la figure 3.4. On observe, comme prévu, une périodicité de 10 secondes. Nous avons superposé trois cycles afin de les comparer (Fig. 3.5). Le cycle a été reproduit avec un écart-type d'erreur de 7 nm (0,02 rad) seulement. L'estimation du focus est donc bien répétable.

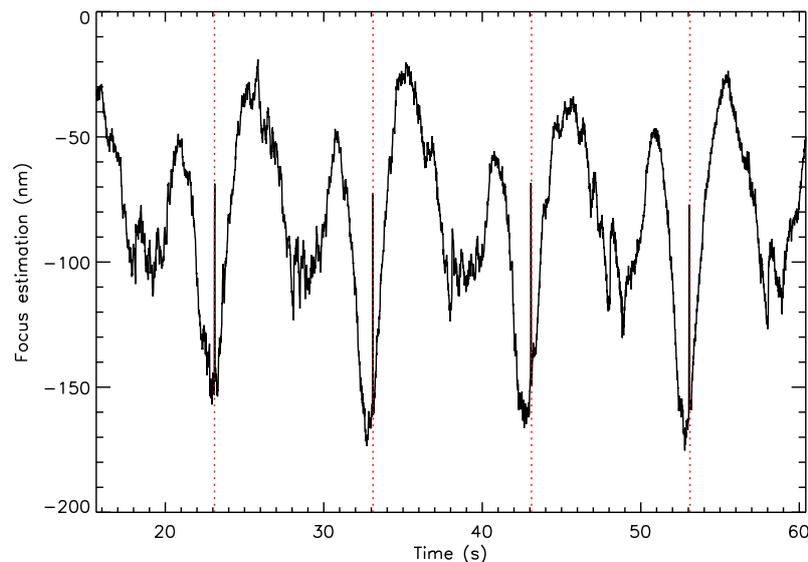


FIGURE 3.4 – Estimation du focus sur source de calibration interne de GeMS avec une turbulence résiduelle simulée grâce aux miroirs déformables.

A ce stade, nous avons donc démontré la linéarité et la répétabilité de l'estimation du focus sur des sources de calibration internes. Dans la suite, nous cherchons à démontrer la linéarité de l'estimation en présence d'une vraie turbulence résiduelle.

3.2.3 Validation de la linéarité sur une étoile

3.2.3.1 Guide de lecture

Ces tests ont fait l'objet d'un acte de conférence reproduit dans le paragraphe suivant. La première partie (§2) est un rappel du fonctionnement de LIFT, qui a déjà été expliqué dans les parties 1.4.2 p.56 et 2.2 p.66 de la thèse. Nous y avons ajouté une comparaison qualitative de LIFT avec une technique similaire proposée récemment, le ILPD [Smith *et al.*, 2013], pour *Iterative Linear Phase Diversity* (§2.3). Cette comparaison est reprise au paragraphe 3.3.4 p.126.

La validation sur ciel est ensuite abordée (§3). Le système GeMS est d'abord présenté (§3.1), de la même manière qu'au 3.2.1 p.97. Le paragraphe 3.2 de l'acte décrit le protocole suivi et les résultats obtenus pour la validation sur ciel de l'estimation du focus par LIFT.

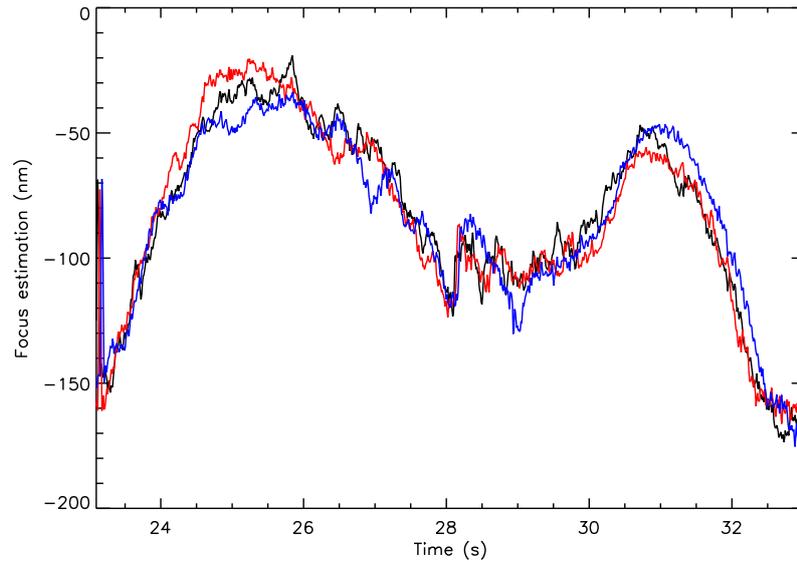
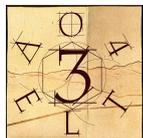


FIGURE 3.5 – Superposition des estimation de focus pour trois cycles de turbulence simulée. Erreur RMS entre chaque cycle : 7 nm.

Enfin, nous discutons de l'utilisation de LIFT dans un système de MCAO (§4). En particulier, nous faisons une comparaison de la propagation du bruit pour l'estimation du tip/tilt entre LIFT et le *Quad-Cell*, qui est la méthode utilisée pour estimer le tip/tilt sur étoile naturelle dans GeMS.

3.2.3.2 Acte de conférence

L'acte de conférence est reproduit dans les pages suivantes.



On-sky validation of LIFT on GeMS

Cédric Plantet^{1a}, Serge Meimon¹, Jean-Marc Conan¹, Benoit Neichel², and Thierry Fusco^{1,2}

¹ Onera, the French Aerospace Lab, 92322 Chatillon, France

² Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France

Abstract. Laser assisted adaptive optics systems rely on Laser Guide Star (LGS) Wave-Front Sensors (WFS) for high order aberration measurements, and rely on Natural Guide Stars (NGS) WFS to complement the measurements on low orders such as tip-tilt and focus. The sky-coverage of the whole system is therefore related to the limiting magnitude of the NGS WFS. We have recently proposed LIFT, a novel phase retrieval WFS technique, that allows a 1 magnitude gain over the usually used 2x2 Shack-Hartmann WFS. Its noise propagation is comparable to a 4-pixel pyramid sensor without modulation. Besides, it requires a much more simple hardware, making it a reliable and easy to set up solution. Early this year, LIFT came out of the lab and has been tested on GeMS, the multiconjugate adaptive optics system of Gemini South. We present here the first on-sky IR wave-front sensing data obtained with LIFT. We show that these results constitute a clear on-sky demonstration of the LIFT concept.

1 INTRODUCTION

Tomographic adaptive optics uses laser guide stars to measure the turbulence volume. Nevertheless, natural guide stars are still needed to sense low order modes [1, 2].

We recently proposed a wavefront sensor called LIFT (Linearized Focal-plane Technique) dedicated to faint natural guide stars [3]. The estimation by LIFT is based upon a single full aperture image taken at the focal plane of the telescope. The relation between the aberrations and the intensity pattern is linearized to make computations easier and faster. Besides, we add a known phase offset to avoid indetermination. First experiments on ONERA's test bench and on the calibration sources of GeMS, the multi-conjugate adaptive optics (MCAO) system of Gemini South Telescope [4], have already provided a first validation of LIFT [5, 6]. The next step was an on-sky validation that is the object of the present paper. We present here on-sky results obtained with GeMS on a natural star. We also briefly discuss a similar concept proposed recently, the Iterative Linear Phase Diversity (ILPD)[7].

The sensor concept is shortly explained and compared to ILPD in the first part. We then show the first on-sky validation of LIFT on GeMS. Finally, we discuss the implementation of LIFT on a MCAO system.

2 LIFT : A PHASE-RETRIEVAL SENSOR

LIFT is a focal plane WFS, comprising a focusing optical system introducing a known amount of astigmatism, a CCD camera at Shannon sampling ¹, and an original phase retrieval algorithm (Fig. 1). We now recall its principle and basic equations.

^a cedric.plantet@onera.fr

¹ A Shannon/2 sub-sampled version is also under study.

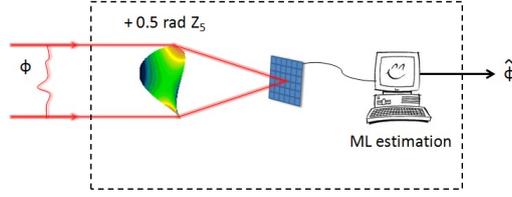


Fig. 1: LIFT schematics.

2.1 Data formation

The intensity pattern on the imaging sensor is :

$$I(\phi) = \int_{\lambda} F(\lambda) |\text{FT}\{\underbrace{P \exp[i\phi_d(\lambda)]}_{P_d} \times \exp[i\phi(\lambda)]\}|^2 d\lambda + n = \int_{\lambda} F(\lambda) I_{\lambda}(\phi) d\lambda + n \quad (1)$$

with F the flux, ϕ the aberrated phase to be estimated, ϕ_d the known phase offset and n the noise. Since we work with numerical data, we consider a discrete sum over the wavelengths :

$$I(\phi) = \sum_{\lambda} F_{\lambda} I_{\lambda}(\phi) + n \quad (2)$$

ϕ can be decomposed on Zernike modes Z_i so that $\phi = \sum_i a_i Z_i$. Let \mathbf{A} be the vector of coefficients a_i , and I be the function giving the intensity pattern in the focal plane. We assume that the aberrations, in optical path difference ($\mathbf{A}_{\text{OPD}} = \mathbf{A} \times \lambda/2\pi$), do not depend on the wavelength.

Eq. 2 becomes

$$\mathbf{I}(\mathbf{A}_{\text{OPD}}) = \sum_{\lambda} F_{\lambda} \mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}}) + \mathbf{n} \quad (3)$$

$\mathbf{I}(\mathbf{A}_{\text{OPD}})$ is a vector, its p -th component $\mathbf{I}(\mathbf{A}_{\text{OPD}})[p]$ being the p -th pixel of the image.

We use a first order Taylor expansion around $\phi = 0$ to linearize Eq. 3, as described by Gonçalves for small phase estimation [8]:

$$\mathbf{I}(\mathbf{A}_{\text{OPD}}) - \mathbf{I}(\mathbf{0}) \simeq \sum_k a_k \mathbf{I}'_k(\mathbf{0}) + \mathbf{n} \quad (4)$$

with $\mathbf{I}'_k(\mathbf{0}) = \left. \frac{\partial \mathbf{I}(\mathbf{A}_{\text{OPD}})}{\partial a_k} \right|_{\mathbf{A}_{\text{OPD}}=\mathbf{0}}$ and \mathbf{n} the noise vector.

Let $\Delta \mathbf{I}$ be the vector so that $\Delta I[p] = I(\mathbf{A}_{\text{OPD}})[p] - I(\mathbf{0})[p]$, and \mathbf{H} be the matrix of which the element at the p -th row and k -th column is defined by $H[p, k] = I'_k[p]$. Then Eq. (4) becomes:

$$\sum_{\lambda} F_{\lambda} [\mathbf{I}_{\lambda}(\mathbf{A}_{\text{OPD}}) - \mathbf{I}_{\lambda}(\mathbf{0})] = \left(\sum_{\lambda} \mathbf{H}_{\lambda}(\mathbf{0}) \right) \mathbf{A}_{\text{OPD}} + \mathbf{n}$$

or

$$\Delta \mathbf{I} = \mathbf{H}(\mathbf{0}) \mathbf{A}_{\text{OPD}} + \mathbf{n} \quad (5)$$

Eq. 5 is solved using a maximum likelihood (ML) estimation. If the optical bandwidth is narrow enough, we can consider the central wavelength only.

2.2 Algorithm

We approximate the photon noise by an additive zero-mean Gaussian noise, with a variance equal to the mean flux \bar{I}_k for the k-th pixel. Read-out noise follows a zero-mean Gaussian distribution with a standard deviation σ_e considered uniform on the detector, hence the global noise variance for the pixel k:

$$\sigma_{n_k}^2 = \bar{I}_k + \sigma_e^2 \quad (6)$$

We call the noise covariance matrix $\mathbf{R}_n = \langle \mathbf{nn}^t \rangle$. The solution of the maximum likelihood estimation is

$$\hat{\mathbf{A}}_{ML} = \mathbf{P}_{ML}(\mathbf{0})\Delta\mathbf{I}; \quad \mathbf{P}_{ML}(\mathbf{0}) = (\mathbf{H}(\mathbf{0})^t\mathbf{R}_n^{-1}\mathbf{H}(\mathbf{0}))^{-1}\mathbf{H}(\mathbf{0})^t\mathbf{R}_n^{-1} \quad (7)$$

and the variance of the estimation error is given by:

$$\sigma_{error}^2 = Tr\{\langle \mathbf{EE}^t \rangle\} = Tr\{(\mathbf{H}^t\mathbf{R}_n^{-1}\mathbf{H})^{-1}\} \quad (8)$$

with $\mathbf{E} = \mathbf{A} - \hat{\mathbf{A}}_{ML}$. Each diagonal element of $\langle \mathbf{EE}^t \rangle$ is equal to the variance of the estimation error for one mode.

Of course, first order Taylor expansion has a limited validity range, which defines the sensor linearity domain. However, this domain can be extended by iterating the estimation. We use the following algorithm:

- Iteration 0: $\Delta\mathbf{I} = \mathbf{I}(\mathbf{A}) - \mathbf{I}(\mathbf{0})$; $\hat{\mathbf{A}}_{ML_0} = \mathbf{P}_{ML}(\mathbf{0})\Delta\mathbf{I}$
- Iteration 1: $\Delta\mathbf{I} = \mathbf{I}(\mathbf{A}) - \mathbf{I}(\hat{\mathbf{A}}_{ML_0})$; $\hat{\mathbf{A}}_{ML_1} = \mathbf{P}_{ML}(\hat{\mathbf{A}}_{ML_0})\Delta\mathbf{I} + \hat{\mathbf{A}}_{ML_0}$
- Iteration 2: $\Delta\mathbf{I} = \mathbf{I}(\mathbf{A}) - \mathbf{I}(\hat{\mathbf{A}}_{ML_1})$; $\hat{\mathbf{A}}_{ML_2} = \mathbf{P}_{ML}(\hat{\mathbf{A}}_{ML_1})\Delta\mathbf{I} + \hat{\mathbf{A}}_{ML_1}$
- ⋮

The extension of the linearity domain with the number of iterations is shown on Fig. 2 for tip/tilt and focus. A few iterations are sufficient to improve linearity.

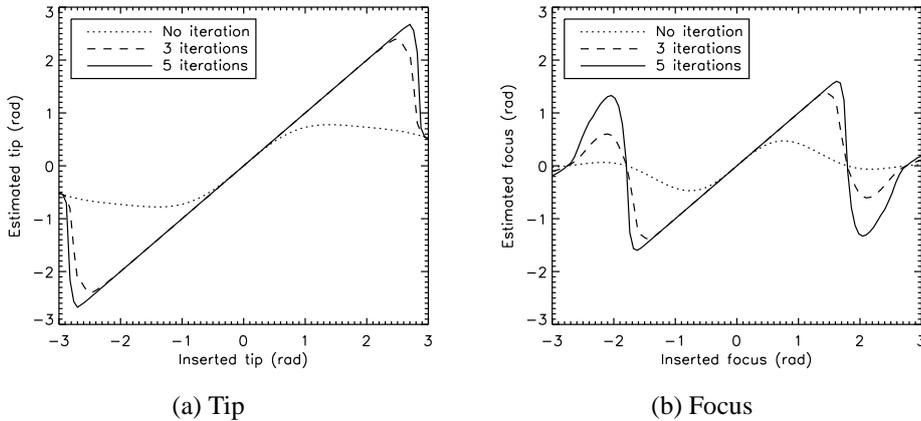


Fig. 2: Linearity domain with different numbers of iterations for tip and focus. Phase offset: 0.4 rad of astigmatism.

2.3 Analysis of a recent variation of LIFT

Another linearized phase retrieval concept has been recently proposed [7], so called Iterative Linear Phase Diversity (ILPD). LIFT and ILPD are very similar approaches in the sense that:

- both algorithms are based on the first order Taylor development in phase of a full aperture image
- the image is assumed to result from a known phase offset (so-called diversity phase) plus the small phase to be estimated
- Taylor development is of course performed around the known diversity phase, as suggested for instance in [9], since this phase offset is not necessarily small.

The methods, and their evaluation, however differ for some aspects:

- LIFT is designed for small amplitude low order wave-front sensing (tip-tilt-focus) on very faint sources, hence the choice of a small (~ 0.5 rad) astigmatism phase offset [6]
- LIFT has been studied for now through a detailed analysis in terms of wave-front sensing error (sensitivity to noise) when applying the algorithm to a single full aperture image. The closed loop behavior is currently under study.
- ILPD seems to be designed for high flux sensing (up to 10^6 photons in [7]) of medium order phase (14 modes in [7]) of possibly rather large amplitude (up to 1 rad rms in [7]), hence a large 2 rad focus phase offset to improve robustness rather than sensitivity
- ILPD is evaluated in a closed loop scheme, it therefore benefits of a new full aperture image at each phase estimation step, and of an optical correction that quickly reduces the amplitude of the sensed residual wavefront.

The two approaches therefore follow a similar logic, and are of interest for their respective field of application. The LIFT-ILPD comparative study found in [7] is however not relevant for the following reasons:

- it compares LIFT based on a single uncorrected full aperture image with ILPD using up to five closed-loop images, with the same SNR per image, hence a huge amount of additional information for ILPD
- it uses LIFT out of the operating domain it was designed for (small phase amplitudes, low order sensing), hence issues of divergence, biases...
- the discussion concerning computation time is not relevant for LIFT, at least for its original application: for tip-tilt-focus estimation, the computation cost is indeed negligible, especially compared to the rest of adaptive optics calculations (high orders control...).

Next section presents GeMS and the results of on-sky tests with LIFT.

3 On-sky validation with GeMS

GeMS (the Gemini Multi-conjugated adaptive optics System) is a facility instrument for the Gemini-South telescope. It delivers a uniform, nearly diffraction-limited image quality at near-infrared wavelengths over an extended field of view of more than 1 arcmin across [10]. The system includes 5 laser guide stars, 3 natural guide stars, 3 deformable mirrors optically conjugated at 0, 4.5 and 9km and 1 tip-tilt mirror [4].

3.1 Optical system

The adaptive optics bench of the Gemini South Telescope (Fig. 3), named Canopus, is designed for multi-conjugate adaptive optics. A fold mirror directs the light collected by the telescope at the Cassegrain focus to the upper entrance shutter. The beam is folded by a flat mirror and collimated by an off-axis parabola onto three deformable mirrors (DM) conjugated at different elevation (9km, 4.5km and 0km respectively) and a tip/tilt mirror (TTM). A science beam splitter transmits the infra-red light onto an atmospheric dispersion corrector. The corrected beam is folded by another flat mirror and refocused at $f/32$ by an off axis parabola to exit through the bottom shutter [11].

The light used for wavefront sensing is reflected by the science beam splitter. The 589nm wavelength from the five laser beacons is reflected by the LGS beam splitter and sent to the Laser Guide Star Wavefront Sensor. The LGSWFS consists of five 16x16 Shack-Hartmann sensors aligned with the five beams of the LGS constellation. The LGSWFS assembly has motorized components to actively control zoom and magnification corrections.

The visible light from the three natural guide stars passes through an atmospheric dispersion corrector onto the Natural Guide Star Wavefront Sensing Unit. The NGSWFS consists of three independent probes, motorized to allow tracking on their respective guide star on the sky. Each probe is fitted with 4 fiber optics feeds attached to quad cell avalanche photo-diodes sensors. ~30% of NGS probe #3's signal is sent to a slow focus sensor. At the time of our experiments, the DM conjugated at 4.5km was not available for technical reasons.



Fig. 3: Canopus system.

For these first on-sky tests, we do not close the loop with LIFT. GeMS runs on its own with its NGS and LGS WFSs while we acquire infrared images with the science detector called GSAOI [12]. These images are processed with LIFT to estimate tip/tilt and focus on a guide star available in the field of view. Since turbulence is weak in infrared light, we have small phases to estimate and we are in LIFT's domain of linearity.

Each of the GSAOI detectors can support one programmable on detector guide window (ODGW) [13]. The ODGWs are designed to select up to four natural guide stars. The ODGWs can be read out at a rate of up to 800 Hz and the sizes of these window range from 2×2 to 128×128 . The GSAOI pixel size is 20mas.

The following tests are made on a natural star with a Ks filter ($\lambda_0 = 2.2\mu\text{m}$, $\Delta\lambda = 0.32\mu\text{m}$). Its magnitude in Ks band is ~ 12.5 .

3.2 Linearity of focus estimation on a natural star

Reference slopes of LGSWFS can be changed to add a chosen offset of astigmatism. To generate images that could be processed by LIFT, we therefore introduced an astigmatism offset of 0.5 rad. The choice of the offset was determined in a former study to minimize LIFT's sensitivity to noise [6].

We wanted to confirm the linear estimation of focus, but it is impossible to compare LIFT's estimation on a NGS with the LGSWFS measurements, mostly because they do not see the same turbulent focus as in the NGS direction. We thus chose to add several offsets of focus thanks to the reference slopes of the LGSWFS and to take a series of images for each offset. The mean of the estimations on one series should be close to the corresponding offset, since dynamical effects eventually average out.

We added several offsets of focus from -100 nm (-0.3 rad) to 200 nm (0.6 rad). 100 images were taken successively for each series. Since we only wanted to demonstrate the linearity of focus estimation, we chose an exposure time of 125 ms to get rid of high order residuals and get a good signal-to-noise ratio. Besides, this frame rate should be enough to correct the fluctuations of focus due to the sodium layer [14].

Figure 4 shows the mean value of estimations on each series. The fluctuations of the sodium layer and an insufficient number of samples can be the causes of small biases. Nevertheless, we obtain a good match with the inserted focus, the maximum error being 40 nm (0.1 rad) on this data set.

We have thus proved the capacity of LIFT to sense focus on real images. In the following section, we discuss how LIFT could be used on a large telescope system.

4 LIFT on a MCAO system

LIFT, as any NGSWFS, could be set up on a specific channel, to provide tip/tilt and focus estimation. It only needs an optical element to introduce astigmatism and an infrared camera. It would benefit from the good MCAO performance in infrared. LIFT has similar performance to a pyramid optimized for sensing tip/tilt/focus and offers a 1 magnitude gain over a Shack-Hartmann 2×2 [6]. In terms of photon noise, its sensitivity is similar to a quad-cell for the estimation of tip/tilt, as shown by table 1.

In a nutshell, LIFT is an efficient solution to sense focus with high precision on NGS without losing performance on tip/tilt estimation. Besides, it is easy to set up and very reliable.

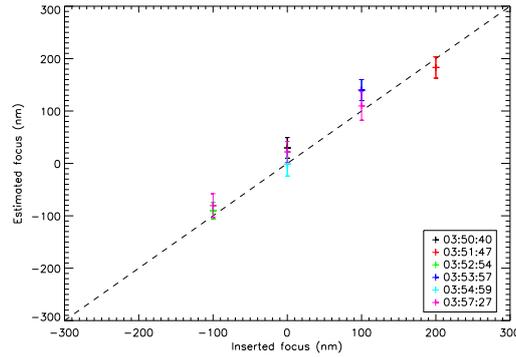


Fig. 4: Focus estimation on a natural star. Each point is the mean value on a series of estimations at a given focus offset. Error bars at one sigma.

	LIFT	QC
α	1.18	1.12
β	32	4

Table 1: Comparison of noise propagation coefficients between LIFT and a quad-cell at Shannon/4 sampling for the estimation of tip and tilt. α is the variance of estimation error for 1 photon without read out noise. β is the variance of estimation error for 1 photon and 1 electron of read out noise without photon noise. The total variance is $\alpha/N_{ph} + \beta \times (\sigma_e/N_{ph})^2$, with N_{ph} the flux in photons and σ_e is the standard deviation of read out noise in electrons.

5 CONCLUSION

LIFT is a single image focal-plane sensor for low orders with high robustness with respect to noise. The concept had already been demonstrated, firstly in lab and then on calibration sources on GeMS, the MCAO system of Gemini South. We then used GeMS to demonstrate the linearity of focus estimation on a natural star. We took infrared images on the scientific detector, GSAOI, with a fast frame rate by reading only regions of interest. The adaptive optics system let us add the astigmatism offset of LIFT and chosen offsets of focus on corrected images. The offsets of focus were successfully retrieved by LIFT.

In a MCAO system, as any NGSWFS, it can be set up on a specific channel, and it would provide a tip/tilt/focus estimation with a 1 magnitude gain over a Shack-Hartmann 2x2.

We now aim to perform a tip/tilt/focus closed loop with LIFT in a wide field adaptive optics system. In particular, it requires a detailed study of aliasing effects in this context. This should lead in the near future to a first closed loop on-sky demonstration, for instance on GeMS.

References

1. T. Fusco, S. Meimon, Y. Clenet, M. Cohen, J. Paufigue, and H. Schnetler, Proc. SPIE **7736**, (2010) 77360D
2. E. Diolaiti et al, Proc. AO4ELT, (2007)
3. S. Meimon, T. Fusco, and L. M. Mugnier, Optics Letters **35**, (2010)
4. B. Neichel et al, Proc. SPIE **7736**, (2010) 773606

5. C. Plantet, B. Neichel, S. Meimon, T. Fusco, and J.-M. Conan, Proc. SPIE, (2012)
6. C. Plantet, S. Meimon, J.-M. Conan, and T. Fusco, Optics Express **21**, (2013) 16337–16352
7. C. S. Smith, R. Marinica, A. J. den Dekker, M. Verhaegen, V. Korhikoski, C. U. Keller and N. Doelman, JOSAA **30**, (2013) 2002–2011
8. Robert A. Gonsalves, Optics Letters **26**, (2001)
9. I. Mocoer, L. M. Mugnier, and F. Cassaing, Optics Letters **34**, (2009) 3487–3489
10. F. Rigaut et al, Monthly Notices of the Royal Astronomical Society, (2013)
11. M. Bec et al, Proc. SPIE **7015**, (2008) 701568
12. P. McGregor et al, Proc. SPIE **5472**, (2004) 1033–1044
13. P. Young, P. Mc Gregor, J. van Harmelena, and B. Neichel, Proc. SPIE **8451**, (2012)
14. B. Neichel, C. D’Orgeville, J. Callingham, F. Rigaut, C. Winge and G. Trancho, MNRAS, (2013)

3.2.3.3 Bilan et discussion

Bilan Cet acte de conférence montre la première validation en boucle ouverte de LIFT sur des images réelles présentant des hauts ordres résiduels, grâce au système de MCAO GeMS. Nous avons démontré que l'estimation du focus restait alors linéaire au moins jusqu'à 0,6 rad avec une erreur maximale de 0,1 rad. Les écarts à la valeur vraie constatés sur l'estimation du focus peuvent être dus à plusieurs facteurs. D'une part, nos références de focus étaient basées sur les étoiles laser, nous étions donc sensibles à des variations d'altitude de la couche de sodium. D'autre part, les images ont été prises sur un temps de pose de 125 ms, qui est relativement grand devant le temps caractéristique d'évolution de la turbulence (quelques dizaines de milli-secondes). On fait donc une erreur de modèle puisque LIFT suppose que le temps de pose est petit devant ce temps d'évolution. En ajoutant le fait qu'il y a des hauts ordres résiduels dans les images, il n'est donc pas étonnant d'avoir eu une précision d'estimation très inférieure à celle obtenue en laboratoire.

D'autre part, nous avons montré que l'estimation du tip/tilt par LIFT était aussi performante que celle d'un quad-cell en bruit de photons. LIFT pourrait donc très bien remplacer les quad-cells, utilisés pour l'analyse du tip/tilt sur étoile naturelle, et offrirait de plus une estimation du focus aussi précise qu'un Shack-Hartmann 2×2 (cf §2.3 p.70).

Discussion Cette validation a été faite à fort flux, avec un réglage fin des paramètres de LIFT, et sur une seule nuit. LIFT, une fois intégré dans une boucle d'optique adaptative, sera soumis à des faibles flux et des conditions de turbulence variées. Nous aimerions que LIFT dispose d'un réglage robuste aux différentes conditions et qui offre une performance optimale. Dans la suite, je m'intéresse donc aux effets des hauts ordres sur l'estimation du tip/tilt et du focus.

3.3 Vers une utilisation de LIFT en boucle fermée

Afin de s'assurer que LIFT sera opérationnel dans un système d'optique adaptative, nous étudions dans cette partie le comportement de LIFT en présence de hauts ordres.

Dans une première partie (§3.3.1 p.110), nous évaluons les effets des hauts ordres et nous proposons des pistes d'amélioration de l'algorithme. Une fois la configuration optimale de LIFT choisie, nous nous intéressons à la perte de performance due à la présence des hauts ordres (§3.3.2 p.117), et nous la comparons à celle d'un Shack-Hartmann 2×2 dans les mêmes conditions. Nous montrons ensuite la première simulation d'une fermeture de boucle sur le tip/tilt et focus avec LIFT (§3.3.3 p.123). Enfin, nous discutons d'une récente variation de LIFT, le ILPD (§3.3.4 p.126)

3.3.1 Améliorer l'algorithme pour limiter les effets des hauts ordres

3.3.1.1 Évaluation des effets de repliement sur l'estimation du tip/tilt et du focus

On rappelle que, par l'intermédiaire d'un développement de Taylor, on exprime la différence entre l'image acquise et une image de référence simulée comme une fonction linéaire des coefficients de Zernike. En reprenant les notations de la partie 2.2 p.66, et en négligeant le bruit, le modèle s'écrit :

$$\Delta I = \mathbf{H} \mathbf{A}_{\text{OPD}} \quad (3.1)$$

On suppose que 500 modes suffisent à décrire correctement les données. On a alors un modèle que l'on considère vrai :

$$\Delta \mathbf{I} = \mathbf{H}_{500} \mathbf{A}_{\text{OPD}, 500} \quad (3.2)$$

et un modèle tronqué pour l'estimation de N modes :

$$\Delta \mathbf{I} = \mathbf{H}_N \mathbf{A}_{\text{OPD}, N} \quad (3.3)$$

L'estimation se fait en appliquant le restructeur de maximum de vraisemblance (cf §2.2 p.66) :

$$\mathbf{P}_{\text{MV}} = \mathbf{H}_{\text{MV}, N}^\dagger \triangleq (\mathbf{H}_N^t \mathbf{R}_n^{-1} \mathbf{H}_N)^{-1} \mathbf{H}_N^t \mathbf{R}_n^{-1} \quad (3.4)$$

avec \mathbf{R}_n la matrice de covariance du bruit. On suppose que le bruit est décorrélé entre les pixels, cette matrice est donc diagonale.

L'estimateur s'écrit alors :

$$\hat{\mathbf{A}}_{\text{OPD}, N} = \mathbf{H}_{\text{MV}, N}^\dagger \Delta \mathbf{I} = \overbrace{\mathbf{H}_{\text{MV}, N}^\dagger \mathbf{H}_{500}}^{\mathbf{R}_{\text{alias}}} \mathbf{A}_{\text{OPD}, 500} \quad (3.5)$$

$\mathbf{R}_{\text{alias}}$ traduit la façon dont les modes non reconstruits perturbent l'estimation des modes reconstruits. On peut noter, dans l'équation 3.4, qu'en régime de bruit de détecteur, $\mathbf{R}_n = \sigma_e^2 \times Id$, avec σ_e^2 la variance de bruit de détecteur et Id l'identité. Le restructeur est donc indépendant du niveau de bruit dans ce cas. En régime de bruit de photons, \mathbf{R}_n est proportionnelle au flux. Sachant que \mathbf{H}_N et \mathbf{H}_{500} sont de même proportionnels au flux, la reconstruction est donc aussi indépendante du niveau de bruit lorsqu'on est en régime de bruit de photons. On calculera alors un unique $\mathbf{R}_{\text{alias}}$ selon le régime de bruit.

Prenons l'exemple d'une estimation sur $N=3$ modes, l'estimateur s'écrit :

$$\overbrace{\begin{pmatrix} \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \end{pmatrix}}^{\hat{\mathbf{A}}_{\text{OPD}}} = \overbrace{\begin{pmatrix} r_{22} & r_{23} & r_{24} & \dots & r_{2\infty} \\ r_{32} & r_{33} & r_{34} & \dots & r_{3\infty} \\ r_{42} & r_{43} & r_{44} & \dots & r_{4\infty} \end{pmatrix}}^{\mathbf{R}_{\text{alias}}(\mathbf{A}_{\text{off}})} \overbrace{\begin{pmatrix} a_2 \\ a_3 \\ a_4 \\ a_5 \\ \vdots \\ a_\infty \end{pmatrix}}^{\mathbf{A}_{\text{OPD}}} \quad (3.6)$$

D'où :

$$\hat{a}_i = \sum_k r_{ik} a_k \quad (3.7)$$

Chaque ligne de $\mathbf{R}_{\text{alias}}(\mathbf{A}_{\text{off}})$ donne donc des coefficients de repliement pour le mode correspondant. Le produit du k -ième coefficient de Zernike a_k par son coefficient de repliement r_{ik} est ajouté à l'estimation de a_i . La sous-matrice 3×3 délimitée par les coefficients r_{22} , r_{42} , r_{24} et r_{44} est une matrice identité. En effet, le tip/tilt et le focus, en l'absence de hauts ordres, sont parfaitement estimés et leurs estimations sont décorrélées. Il faut noter que le modèle est construit dans l'approximation d'une phase faible, donc ces calculs ne sont valables qu'avec des petits a_k .

En boucle fermée, la variance d'erreur due au bruit sera sûrement supérieure ou égale à $0,05 \text{ rad}^2$ pour l'estimation du tip/tilt et du focus (variance d'erreur pour un flux d'environ

100 photons et un bruit de détecteur de 2 électrons à la diffraction). Cela correspond à un écart-type d'erreur d'environ 0,13 rad sur chaque mode. On vise une variance résiduelle sur le tip/tilt/focus de $0,1 \text{ rad}^2$ (90% de Strehl si les hauts ordres sont parfaitement corrigés), c'est-à-dire une amplitude résiduelle de 0,18 rad par mode. Pour avoir une erreur suffisamment faible devant celle provoquée par le bruit, nous fixons donc comme **objectif** un repliement maximal de l'ordre de **10% par mode** (30% au total sur tip/tilt et focus) dans chaque régime de bruit.

La figure 3.6 donne les coefficients de repliement pour l'estimation du tip/tilt et focus ($N=3$). L'estimation est faite sur l'image entière (128×128 pixels).

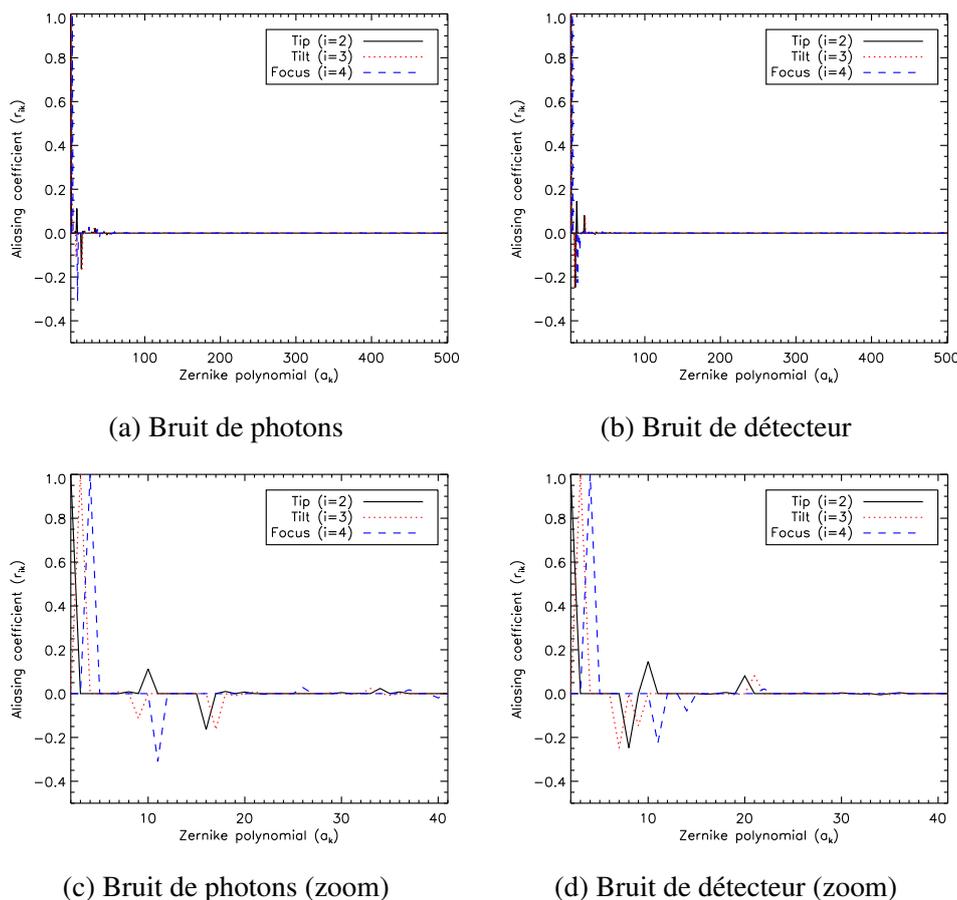


FIGURE 3.6 – Coefficients de repliement r_{ik} sur l'estimation du tip/tilt et du focus en fonction du numéro de mode de Zernike, en régime de bruit de photons et en régime de bruit de détecteur. Le premier mode est Z_2 , le tip. Astigmatisme de référence : 0,5 rad de Z_5 . Échantillonnage à Shannon. Fenêtre de calcul : image complète (128×128 pixels).

On remarque tout d'abord qu'au dessus d'une quarantaine de modes, les effets du repliement sont négligeables. D'autre part, on vérifie que les coefficients de repliement du tip/tilt/focus sur eux-mêmes sont bien égaux à 1.

Repliement sur le tip/tilt Le tip/tilt subit majoritairement le repliement de la coma (Z_7 et Z_8) en bruit de détecteur. En effet, dans ce régime de bruit, l'estimation du tip/tilt est similaire à une estimation de type centre de gravité pondéré (voir §1.4.2 p.56), et la coma est une aberration

qui déplace le centre de gravité de la tache. En régime de bruit de photons, plus de pixels sont pris en compte, donc seules les comas d'ordre supérieur (Z_{16} et Z_{17}), qui ont plus d'effet sur les bords de la tache, se replient sur le tip/tilt.

Repliement sur le focus L'estimation du focus est majoritairement touchée par l'aberration sphérique (Z_{11}) dans les deux régimes de bruit. En effet, ces deux aberrations déforment la tache de manière très similaire.

Au final, le repliement total, obtenu en sommant les valeurs absolues des coefficients de repliement en excluant le tip/tilt et focus, est de l'ordre de 50% en bruit de photons et de 100% en bruit de détecteur pour l'estimation d'un tip ou d'un tilt. Pour l'estimation d'un focus, il est de l'ordre de 35% en bruit de photons et en bruit de détecteur. L'objectif fixé n'est clairement pas atteint ici, il faut donc améliorer l'algorithme. Je présente dans la suite les solutions envisagées.

3.3.1.2 Solutions envisagées et amélioration apportée

Troncature de la fenêtre de calcul Le calcul effectué dans LIFT se fait sur une zone carrée limitée autour du centre de gravité de la tache (voir §1.4.2 p.56 et §2.2 p.66). Le but, outre de gagner en temps de calcul, est de limiter les effets de repliement. En effet, les hautes fréquences de phase engendrent des speckles éloignés du centre de la tache image. En n'incluant pas ces speckles dans l'image traitée, on retire l'influence des hauts ordres sur l'estimation.

Le raisonnement est analogue au filtrage par un trou au plan focal proposé par Poyneer et Macintosh pour le Shack-Hartmann [Poyneer et Macintosh, 2004]. Dans ce paragraphe, j'évalue l'effet du trou filtre proposé par Poyneer et Macintosh sur le filtrage des modes de Zernike, et l'assimile au filtrage par la limitation de la zone de calcul de LIFT. Bien que ces deux approches soient de nature différente, la première étant physique et l'autre numérique, nous avons retrouvé des propriétés de filtrage équivalentes en simulation.

Dans leur article, Poyneer et Macintosh montrent que les fréquences spatiales de la phase inférieures à $1/2d - 1/2D$, avec λ/d la taille du filtre et λ/D la largeur à mi-hauteur de la PSF, passent totalement et que les fréquences spatiales supérieures à $1/2d + 1/2D$ sont totalement rejetées. Entre ces deux bornes, les fréquences sont partiellement rejetées. On peut déduire de cela une relation entre la taille de la fenêtre et les modes de Zernike qui sont filtrés.

Reprenons ces notations pour la fenêtre de calcul de LIFT. Soit N la taille de la fenêtre de calcul en diamètres de tache d'Airy. On a $N = \frac{\lambda D}{d \lambda} = D/d$. Une fréquence spatiale pure f dans la phase engendre une série de speckles à une distance $\pm k \lambda f$ ($k = 1, 2, 3, \dots$) du centre de l'image, dont l'amplitude décroît avec la distance. Cette fréquence passe totalement si $f < 1/2d - 1/2D$, ce qui est équivalent à $f < (N - 1)/2D$. De même, elle est totalement rejetée si $f > (N + 1)/2D$.

Or, d'après la thèse de Jean-Marc Conan [Conan, 1994], on peut définir une fréquence spatiale équivalente pour un mode de Zernike d'ordre radial n_{rad} :

$$f_{eq} \simeq 0.3(n_{rad} + 1)/D \quad (3.8)$$

Autrement dit, $n_{rad} \simeq f_{eq}D/0.3 - 1$. On peut donc déterminer un ordre radial maximal n_{pass} en dessous duquel les modes passent entièrement, et un ordre minimal n_{filt} au-dessus duquel

les modes sont totalement rejetés. On a alors :

$$n_{pass} = \lfloor (N - 1)/0.6 - 1 \rfloor \quad (3.9)$$

$$n_{filt} = \lceil (N + 1)/0.6 - 1 \rceil \quad (3.10)$$

avec $\lfloor \cdot \rfloor$ l'arrondi à l'entier inférieur et $\lceil \cdot \rceil$ l'arrondi à l'entier supérieur. Quel que soit N , on aura $n_{filt} - n_{pass} = 4$ ordres radiaux partiellement rejetés.

Nous évaluons maintenant l'amélioration apportée par ce filtrage. Les coefficients de repliement sur le tip/tilt/focus obtenus pour différentes tailles de fenêtre de calcul sont représentés dans la figure 3.7. Les coefficients sur l'image complète sont montrés à nouveau en guise de rappel. Les tailles de fenêtre que nous avons considérées correspondent à 4 et 3 diamètres de taches d'Airy.

On n'observe aucun changement en bruit de détecteur. En effet, dans ce régime de bruit, seuls les pixels centraux sont pris en compte (cf §1.4.2 p.56), et la taille de la fenêtre importe peu. En bruit de photons, on obtient plutôt l'effet contraire de celui escompté : le repliement augmente. Cela est dû à deux effets :

- Le filtrage par la fenêtre de calcul ne permet pas de rejeter totalement les modes se repliant majoritairement, qui sont d'ordre radial 3 et 4. En effet, l'ordre radial 3 ne peut pas être totalement rejeté, et pour rejeter l'ordre 4, il faudrait prendre une fenêtre de calcul qui fait la taille de la tache d'Airy, ce qui n'est pas envisageable pour une estimation correcte du tip/tilt et du focus.
- En diminuant la taille de la fenêtre, on se prive alors des pixels qui permettraient de faire la différence entre un tip/tilt et une coma, ou entre un focus et une aberration sphérique. Cette solution n'est donc pas satisfaisante.

Estimation régularisée Une autre méthode pour réduire le repliement est de prendre en compte les modes qui se replient dans le modèle. Cela implique d'estimer des modes au-delà du tip/tilt et du focus avec LIFT.

Si l'on veut pouvoir estimer plus de modes, il est préférable d'ajouter une régularisation. La régularisation permet, à l'aide d'une connaissance préalable sur les données, de guider l'estimation et ainsi éviter des divergences de l'algorithme. En effet, si aucune supposition n'est faite sur l'amplitude des modes à estimer, un algorithme plan focal peut facilement diverger à cause du nombre élevé de combinaisons d'aberrations qui mènent à la même tache image.

Pour mettre en place une régularisation, j'ai utilisé l'estimateur Maximum A Posteriori (MAP) à la place du maximum de vraisemblance. Le MAP s'appuie sur la probabilité *a priori* de la phase à estimer. Cette probabilité est décrite par la variance et la moyenne des coefficients de Zernike qui caractérisent la phase à estimer.

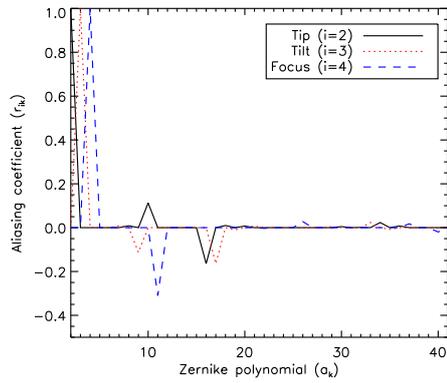
Supposons que la statistique des aberrations résiduelles est connue. La solution du MAP s'écrit comme suit :

$$\hat{\mathbf{A}}_{\text{MAP}} = (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_d) + \mathbf{C}_\phi^{-1})^{-1} (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \Delta \mathbf{I} + \mathbf{C}_\phi^{-1} \bar{\mathbf{A}}) \quad (3.11)$$

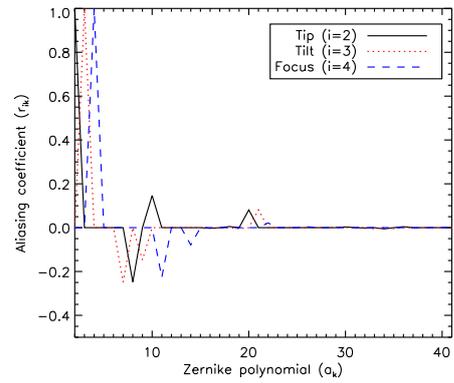
avec \mathbf{C}_ϕ la matrice de covariance des coefficients de Zernike et $\bar{\mathbf{A}}$ leur moyenne.

Cependant, l'algorithme de LIFT est itératif. Lorsqu'on itère, le point de fonctionnement est modifié. Comme nous l'avons vu dans la partie 2.2 p.66, l'estimation à l'itération k peut s'écrire :

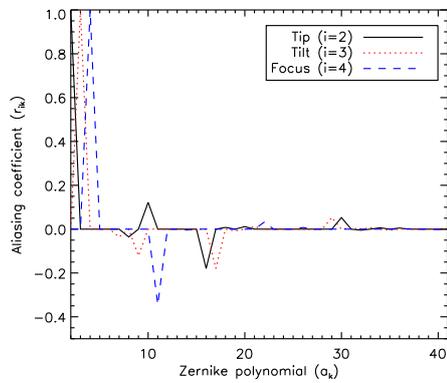
$$\hat{\mathbf{A}}^k = \Delta \hat{\mathbf{A}}^k + \hat{\mathbf{A}}^{k-1} \quad (3.12)$$



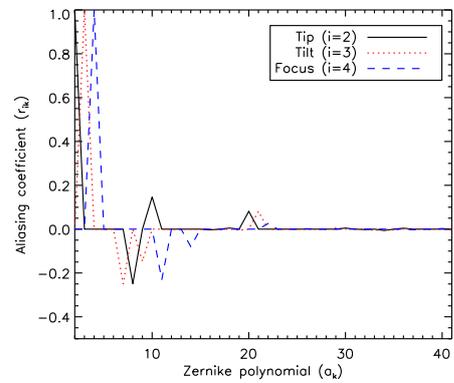
(a) Image complète, bruit de photons



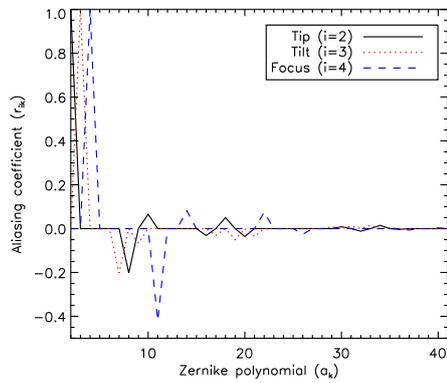
(b) Image complète, bruit de détecteur



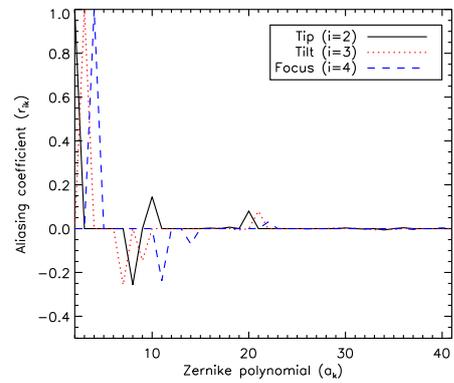
(c) 8×8 pixels, bruit de photons



(d) 8×8 pixels, bruit de détecteur



(e) 6×6 pixels, bruit de photons



(f) 6×6 pixels, bruit de détecteur

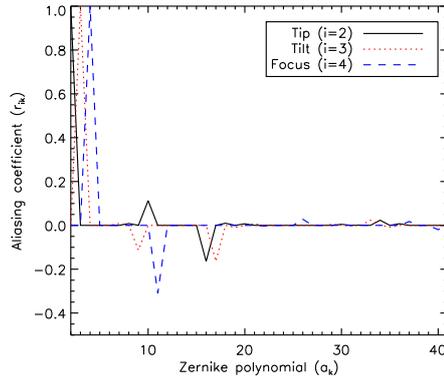
FIGURE 3.7 – Coefficients de repliement en fonction du numéro de mode de Zernike et de la taille de la fenêtre de calcul, en régime de bruit de photons et en régime de bruit de détecteur. Le premier mode est Z_2 , le tip. Astigmatisme de référence : 0,5 rad de Z_5 . Échantillonnage à Shannon.

où $\Delta \hat{\mathbf{A}}^k$ est l'estimateur de $\Delta \mathbf{A}^k = \mathbf{A} - \hat{\mathbf{A}}^{k-1}$ et $\hat{\mathbf{A}}^0 = 0$. Dans notre cas, il faut donc plutôt écrire l'estimateur MAP de cette façon :

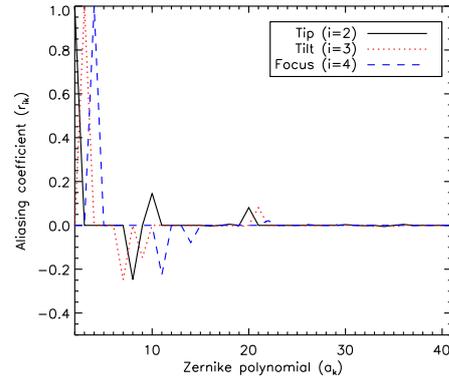
$$\hat{\mathbf{A}}_{\text{MAP}}^k = (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{A}_d) + \mathbf{C}_\phi^{-1})^{-1} (\mathbf{H}(\mathbf{A}_d)^t \mathbf{R}_n^{-1} \Delta \mathbf{I} + \mathbf{C}_\phi^{-1} \Delta \bar{\mathbf{A}}^k) + \hat{\mathbf{A}}_{\text{MAP}}^{k-1} \quad (3.13)$$

avec $\Delta \bar{\mathbf{A}}^k = \bar{\mathbf{A}} - \hat{\mathbf{A}}^{k-1}$. Sachant que la moyenne des aberrations résiduelles $\bar{\mathbf{A}}$ est nulle, on a donc $\Delta \bar{\mathbf{A}}^k = 0$ à la première itération et $\Delta \bar{\mathbf{A}}^k = -\hat{\mathbf{A}}^{k-1}$ à l'itération k .

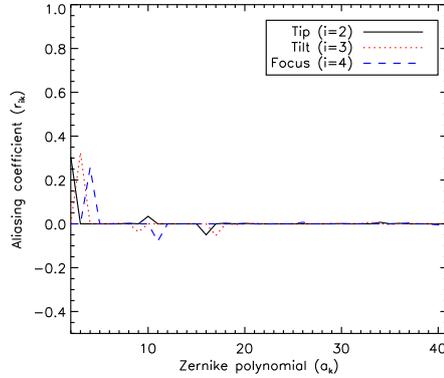
Nous évaluons maintenant l'amélioration apportée par cette solution. Les coefficients de repliement obtenus pour une estimation MAP sont représentés dans la figure 3.8. Les coefficients pour l'estimation MV sont rappelés sur la figure. La régularisation est basée sur des résidus d'aberrations typiques d'un système tel que GeMS.



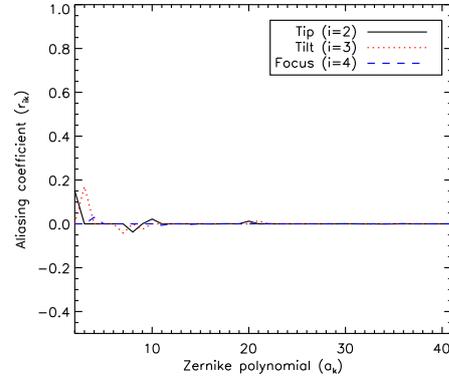
(a) Estimation MV, bruit de photons



(b) Estimation MV, bruit de détecteur



(c) Estimation MAP, bruit de photons



(d) Estimation MAP, bruit de détecteur

FIGURE 3.8 – Coefficients de repliement en fonction du numéro de mode de Zernike avec une estimation MAP, en régime de bruit de photons et en régime de bruit de détecteur. Le premier mode est Z_2 , le tip. Astigmatisme de référence : 0,5 rad de Z_5 . Échantillonnage à Shannon.

On observe cette fois une nette diminution des coefficients de repliement. Les coefficients du tip/tilt et du focus sont aussi modifiés car l'estimation MAP est une estimation biaisée : l'estimateur ne converge pas vers la valeur vraie si l'on moyenne les estimations sur beaucoup d'occurrences de bruit, mais, si l'*a priori* utilisé est correct, l'erreur d'estimation est en moyenne moins élevée.

Avec cette méthode, on a un repliement total d'environ 20% en bruit de photons et 15% en bruit de détecteur sur l'estimation du tip/tilt, et d'environ 10% en bruit de photons et 0% en bruit de détecteur sur l'estimation du focus. Cela donne 30% au total sur le tip/tilt et focus en bruit de photons et 15% en bruit de détecteur. L'objectif initial est donc atteint avec cette méthode.

Un autre avantage de l'estimation MAP est qu'elle est quasiment indépendante de la taille de la fenêtre et du nombre de modes estimés, ce qui permet de garder un réglage fixe pour n'importe quel cas. Dans toute la suite, nous utilisons une estimation MAP.

3.3.2 Sensibilité au bruit en présence de hauts ordres résiduels

LIFT utilisera des images pré-corrigées par une optique adaptative assistée par laser. La qualité d'estimation de LIFT dépendra donc de la correction effectuée grâce aux étoiles laser. Selon les conditions atmosphériques et le dimensionnement du système, cette correction sera plus ou moins efficace.

Dans cette partie, nous expliquons le choix de paramètres que nous avons fait et montrons comment calculer des images partiellement corrigées dans différentes conditions. Nous évaluons ensuite l'impact des hauts ordres, sans chercher à le limiter, sur la propagation du bruit pour l'estimation du tip/tilt/focus dans le cas d'un Shack-Hartmann 2×2 , qui nous servira de référence, puis dans le cas de LIFT. Nous nous attendons à ce que l'étalement du flux dû aux hauts ordres soit équivalent à avoir un flux efficace $n_{ph,eff} = n_{ph} \times SR$, avec n_{ph} le flux réel et SR le rapport de Strehl. Cela implique, pour les coefficients de propagation du bruit de photons α et du bruit de détecteur β (cf §1.4.5 p.60), les relations suivantes :

$$\begin{aligned}\alpha &\sim \alpha^0 / SR \\ \beta &\sim \beta^0 / SR^2\end{aligned}\tag{3.14}$$

avec α^0 et β^0 les coefficients calculés à la limite de diffraction dans la partie 2.3 p.70.

Sur un système réel, les mesures des analyseurs bas ordres dédiés aux étoiles naturelles sont concaténées à celles des analyseurs hauts ordres dédiés aux étoiles lasers, ce qui permet de réduire les effets de repliement dans la reconstruction finale. Pour une comparaison simple, nous considérons ici que l'estimation des bas ordres sur étoile naturelle est faite sans aucune prise en compte de l'estimation des hauts ordres sur les étoiles laser. Il s'agit donc d'un cas pessimiste.

3.3.2.1 Hypothèses et outils de travail

Paramètres du système d'optique adaptative LIFT a été validé pour la première fois sur le ciel grâce au système d'optique adaptative du télescope Gemini Sud, GeMS. Afin de déterminer les paramètres de notre système d'optique adaptative, nous avons donc décidé de nous inspirer de GeMS. Le système a été décrit au paragraphe 3.2.1 p.97.

La première étoile laser se trouve sur l'axe optique. Les 4 autres étoiles sont aux sommets d'un carré d'une minute d'arc de côté selon le schéma 3.9.

Les différents paramètres de simulation sont listés dans le tableau 3.1.

Création de phases résiduelles Pour calculer des images issues d'un tel système, on a besoin de créer des occurrences de phase résiduelle. Il faut alors connaître la statistique de la phase

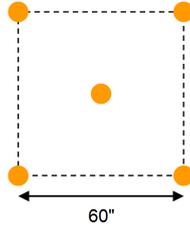


FIGURE 3.9 – Disposition des étoiles laser. L'étoile au centre se trouve sur l'axe optique.

Paramètre	Valeur
Diamètre du télescope	8 m
Fréquence d'échantillonnage ASO étoiles laser	400 Hz
Retard de la boucle	5 ms
Altitude conjuguée du miroir 1	0 km
Altitude conjuguée du miroir 2	9 km
Pas des actionneurs du miroir 1	0,5 m
Pas des actionneurs du miroir 2	1 m
Zeropoint V	10^{10} ph/m ² /s
Magnitude V étoiles laser	10
Longueur d'onde des étoiles laser	0,6 μ m
Longueur d'onde d'imagerie	2,2 μ m
Nombre de sous-pupilles pour l'analyse sur étoile laser	16 \times 16

TABLE 3.1 – Paramètres du système d'optique adaptative considéré dans les simulations.

résiduelle, qui est caractérisée par sa Densité Spectrale de Puissance (DSP). La DSP permet de connaître l'énergie associée à chaque fréquence spatiale de la phase. Benoît Neichel a montré dans sa thèse qu'il était possible de calculer des DSP de phase résiduelle pour un système donné grâce à une analyse dans le domaine de Fourier [Neichel, 2008, Neichel *et al.*, 2009]. La méthode qu'il décrit est celle utilisée pour les simulations qui suivent. Elle inclut la prise en compte des erreurs de reconstruction et de sous-modélisation du miroir, ainsi que l'erreur temporelle. L'erreur de repliement et l'effet de cône sont ici négligés.

Le calcul d'une DSP de phase résiduelle nécessite évidemment des hypothèses sur les conditions de turbulence. Les caractéristiques qui nous intéressent sont : le profil de C_n^2 , le profil de vent, le seeing, la grande échelle et l'angle zénithal d'observation. Les profils de C_n^2 et de vent sont modélisés sur 4 couches. Le tableau 3.2 montre les valeurs choisies pour ces profils.

Altitude (km)	C_n^2	Vitesse du vent (m/s)
0	40 %	10
2	30 %	15
9	20 %	20
17	10 %	25

TABLE 3.2 – Profils de C_n^2 et de vent. Le C_n^2 par couche est exprimé en pourcentage du C_n^2 intégré sur toutes les couches. La direction du vent dans chaque couche est aléatoire.

La grande échelle est fixée à 50 m, et l'angle zénithal d'observation à 30° . Afin d'avoir des DSP représentatives de différentes forces de turbulence, on fait varier le seeing entre 0,45" et 1,1". Pour chaque seeing, on considère 5 directions d'analyse pour LIFT. La première est sur l'axe et les suivantes s'éloignent régulièrement de l'axe avec un pas de 10", selon la disposition décrite par le schéma 3.10.

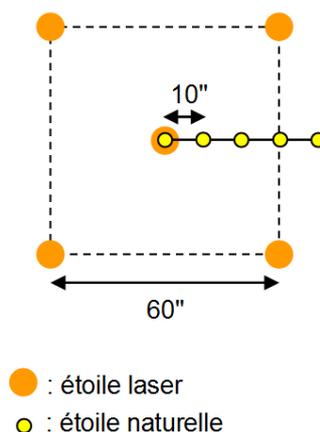


FIGURE 3.10 – Directions d'analyse par LIFT considérées pour le calcul des DSP, supposées correspondre à des positions d'étoiles naturelles.

Le tableau 3.3 récapitule les cas simulés et le seeing correspondant, ainsi que la performance

de la correction exprimée en rapport de Strehl. Dans chaque cas, on calcule ce rapport pour 50 tirages de phase et on en fait la moyenne.

Le calcul d'une phase résiduelle se fait de la manière suivante. Un tableau 2D ayant les mêmes dimensions que la DSP est créé. La taille de la DSP est au moins 2 fois supérieure à celle choisie pour la pupille, afin de respecter le critère de Shannon. Le tableau est ensuite rempli de valeurs aléatoires qui suivent une statistique gaussienne de moyenne nulle, comme la phase turbulente. Pour pondérer les valeurs de manière à respecter la DSP, on fait la transformée de Fourier du tableau, ce qui permet de passer dans l'espace des fréquences spatiales. Les valeurs à calculer sont des amplitudes de phase en radians, tandis que la DSP représente des énergies¹, en radians², on effectue donc la pondération en multipliant le tableau par la racine de la DSP. Enfin, on revient dans l'espace de la phase par transformée de Fourier inverse. On a alors un écran de phase présentant des résidus de hauts ordres en accord avec la DSP. On en extrait une partie, de la taille de la pupille, pour obtenir une occurrence de phase résiduelle.

Cas n°	Seeing (")	Rapport de Strehl sur axe (%)	Rapport de Strehl à 10" (%)	Rapport de Strehl à 20" (%)	Rapport de Strehl à 30" (%)	Rapport de Strehl à 40" (%)
1	0,45	76	71	62	56	55
2	0,7	57	49	38	30	30
3	0,9	42	33	23	17	15
4	1,1	27	18	10	7	6

TABLE 3.3 – Seeings considérés et performances associées, exprimées en rapport de Strehl (hors tip/tilt).

Nous avons maintenant les éléments nécessaires pour évaluer l'impact des hauts ordres sur la propagation du bruit. Dans la suite, nous comparons la performance du Shack-Hartmann 2×2 et de LIFT en présence de hauts ordres résiduels.

3.3.2.2 Impact des hauts ordres sur la performance du Shack-Hartmann 2x2

Lorsque que le seeing augmente, le rapport de Strehl diminue et les contributions des différentes aberrations se mélangent de plus en plus dans l'image. D'autre part, le flux est de plus en plus étalé sur les pixels. On peut donc s'attendre à ce que la performance de l'analyseur diminue. Nous évaluons ici l'impact des hauts ordres sur la propagation du bruit de l'analyseur classique pour l'estimation du tip/tilt/focus sur étoile naturelle : le Shack-Hartmann 2×2. Cela nous servira de référence pour la performance de LIFT. Le tableau 3.4 donne la variance d'erreur d'estimation du tip/tilt et du focus pour les différentes valeurs du seeing avec un flux total de 100 photons. Les pentes sont estimées par centre de gravité pondéré, afin de minimiser la propagation du bruit de détecteur (voir §2.3 p.70), qui est dominant à très faible flux, et on propage le bruit sur l'estimation grâce aux coefficients de Rigaut-Gendron [Rigaut et Gendron, 1992].

A partir de ces valeurs, on peut déduire de nouveaux coefficients de propagation du bruit (voir §1.4.5 p.60) pour les différents niveaux de Strehl. On retire pour cela l'erreur due aux

1. En fait, l'unité de la DSP est le radian²/m² dans notre cas, car la DSP représente des énergies par unité de fréquence spatiale 2D. Cela implique seulement l'ajout d'un facteur de normalisation dans le calcul décrit ici.

Cas n°	Rapport de Strehl sur axe (%)	Variance d'erreur sur tip/tilt/focus (rad ²) en bruit de photons uniquement	Variance d'erreur sur tip/tilt/focus (rad ²) en bruit de détecteur uniquement	Variance d'erreur sur tip/tilt/focus due aux effets de repliement (rad ²)
1	76	0,16	0,13	0,071
2	57	0,28	0,27	0,15
3	42	0,39	0,47	0,21
4	27	0,63	0,85	0,28

TABLE 3.4 – Variance d'erreur d'estimation du tip/tilt/focus par un Shack-Hartmann 2×2 pour les différents seeings considérés, calculée par la méthode de Monte-Carlo avec 1000 occurrences. Méthode d'estimation des pentes : centre de gravité pondéré. Direction d'observation : axe optique. Flux total = 100 photons, écart-type de bruit de détecteur = 1 électron. La variance est indiquée pour deux cas : Bruit de photons uniquement ou bruit de détecteur uniquement.

effets de repliement, qui est indépendante du bruit, à la variance d'erreur d'estimation, puis on applique le calcul des coefficients décrit dans le paragraphe 1.4.5 p.60. Le tableau 3.5 montre ces coefficients pour les différents cas, ainsi que les coefficients théoriques donnés par l'équation 3.14. Les coefficients calculés à la limite de diffraction dans la partie 2.3 p.70 sont rappelés ici.

Cas n°	Rapport de Strehl sur axe (%)	α	β	α_{th}	β_{th}
0	100	8,19	334	8,19	334
1	76	8,9	590	10,78	578
2	57	13	1200	14,36	1028
3	42	18	2600	19,5	1893
4	27	35	5700	30,33	4582

TABLE 3.5 – Coefficients de propagation du bruit de photons (α) et du bruit de détecteur (β) du Shack-Hartmann 2×2 pour l'estimation du tip/tilt et du focus en fonction du rapport de Strehl. Cas 0 : cas limité par la diffraction calculé dans la partie 2.3 p.70. α_{th} et β_{th} sont les coefficients déduits de l'équation 3.14.

Les coefficients obtenus sont du même ordre de grandeur que ceux attendus. Même si la différence semble parfois grande, cela peut correspondre à une faible erreur dans l'estimation du rapport de Strehl. En particulier, le β obtenu dans le cas n° 4 correspondrait à un rapport de Strehl de 24% au lieu de 27%. La loi définie par l'équation 3.14 est donc bien vérifiée.

3.3.2.3 Impact des hauts ordres sur la performance de LIFT

Nous évaluons maintenant la performance de LIFT en présence de hauts ordres résiduels. On compare tout d'abord la performance de LIFT sur axe avec celle du Shack-Hartmann, puis nous discutons des directions hors-axe.

Performance sur axe et comparaison avec le Shack-Hartmann Le tableau 3.6 donne la variance d'erreur sur le tip/tilt et focus pour différentes valeurs du seeing avec un flux de 100 photons.

Cas n°	Rapport de Strehl sur axe (%)	Variance d'erreur sur tip/tilt/focus en bruit de photons uniquement (rad ²)	Variance d'erreur sur tip/tilt/focus en bruit de détecteur uniquement (rad ²)	Variance d'erreur sur tip/tilt/focus due aux effets de repliement (rad ²)
1	76	0,022	0,012	0,00037
2	57	0,038	0,035	0,0058
3	42	0,081	0,086	0,031
4	27	0,23	0,29	0,16

TABLE 3.6 – Variance d'erreur d'estimation du tip/tilt et focus par LIFT pour les différents seeings, calculée par la méthode de Monte-Carlo avec 100 occurrences. Direction d'observation : axe optique. Flux = 100 photons, écart-type de bruit de détecteur = 1 électron. La variance est indiquée pour deux cas : bruit de photons uniquement ou bruit de détecteur uniquement. Taille de la fenêtre : 8 taches d'Airy. 54 modes estimés.

Comme au paragraphe précédent, on peut déduire de ces valeurs de nouveaux coefficients de propagation du bruit, qui sont donnés dans le tableau 3.7.

Cas n°	Rapport de Strehl sur axe (%)	α	β	α_{th}	β_{th}
0	100	1,71	87	1,71	87
1	76	2,16	116	2,25	151
2	57	3,22	292	3	268
3	42	5	550	4,07	493
4	27	7	1300	6,33	1193

TABLE 3.7 – Coefficients de propagation du bruit de photons (α) et du bruit de détecteur (β) dans LIFT pour l'estimation du tip/tilt et du focus en fonction du rapport de Strehl. Cas 0 : cas limité par la diffraction calculé dans la partie 2.3 p.70. α_{th} et β_{th} sont les coefficients déduits de l'équation 3.14.

De même que précédemment, on vérifie la loi décrite par l'équation 3.14. Ainsi, le gain de performance offert par LIFT par rapport au Shack-Hartmann 2×2 en termes de propagation de bruit reste celui montré dans la partie 2.3 p.70 quel que soit le rapport de Strehl. Bien sûr, les effets de repliement interviennent aussi dans le budget d'erreur, mais on peut voir que LIFT y est beaucoup moins sensible. LIFT surpasse donc le Shack-Hartmann 2×2 sur les deux termes du budget d'erreur considérés ici. On peut toutefois noter que la performance de LIFT diminue très fortement lorsqu'on passe de 42% à 27% de Strehl. En effet, en dessous de 36% de Strehl, la phase a une variance supérieure 1 rad², et on dépasse le cadre de l'approximation de faible phase. On devient alors très sensible au repliement des hauts ordres.

Performance hors-axe Nous avons évalué la performance de LIFT sur axe, mais pour un même seeing, la correction de l'optique adaptative se dégrade aussi lorsque l'axe d'observation s'éloigne de l'axe optique. Nous voulons ici vérifier que la performance est la même quelle que soit l'origine de la dégradation du rapport de Strehl (seeing ou direction d'observation). Nous traçons alors la variance d'erreur d'estimation pour différentes directions d'observation et différents seeings en fonction du rapport de Strehl (Fig. 3.11).

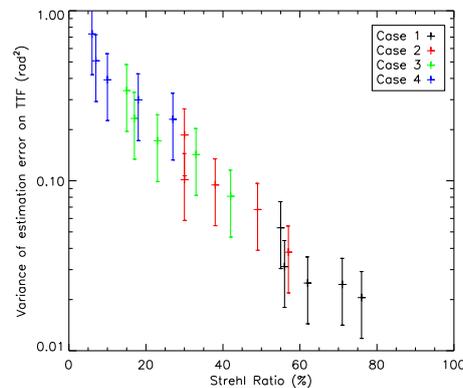


FIGURE 3.11 – Variance d'erreur d'estimation sur tip/tilt/focus pour différentes directions d'observations et différents seeings en fonction du rapport de Strehl, calculée par la méthode de Monte-Carlo avec 100 occurrences. Flux = 100 photons. Bruit de photons uniquement.

La diminution de la performance semble bien n'être due qu'à la diminution du rapport de Strehl. Les conclusions tirées précédemment sur la propagation de bruit sont donc légitimes pour toutes les conditions d'observation.

Bilan Dans cette partie, nous avons défini un lien entre le rapport de Strehl et la propagation du bruit, et confirmé sa validité pour une estimation du tip/tilt/focus par un Shack-hartmann 2×2 ou par LIFT. Nous avons alors montré que l'erreur due au bruit était moins importante pour LIFT que pour le Shack-Hartmann 2×2 quel que soit le rapport de Strehl. Le gain apporté par LIFT sur ce terme d'erreur est bien celui annoncé dans la partie 2.3 p.70, c'est-à-dire un gain en magnitude limite de l'ordre de 1,5.

Nous avons maintenant tous les éléments pour utiliser LIFT dans une boucle d'optique adaptative. Dans le paragraphe suivant, nous présentons la première simulation de LIFT en boucle fermée.

3.3.3 Simulation d'une boucle fermée

3.3.3.1 Paramètres de simulation

La première étape est de générer la phase vue par le télescope en fonction du temps. Pour cela, on génère un écran de phase grâce à une de nos DSP et on déplace linéairement la pupille du télescope dans cet écran de phase, ce qui est équivalent à avoir un vent qui fait défiler la phase devant le télescope.

A partir des profils de C_n^2 et de vitesse de vent décrits dans le tableau 3.2, on peut déterminer la vitesse à laquelle la phase doit défilier devant le télescope. Dans notre cas, on constate que la majorité de la turbulence est présente dans les deux premières couches. Nous pouvons donc approximer le vent moyen par une moyenne sur ces deux couches uniquement. Cela nous donne un vent de 12,5 m/s.

Afin de faciliter les simulations, on aimerait déplacer la pupille d'un pixel par image acquise pour LIFT. Le produit vitesse de vent \times temps de pose doit donc être égal à la longueur équivalente d'un pixel. On prend un temps de pose équivalent à celui des analyseurs étoile laser. Ces analyseurs travaillent à 400 Hz (cf 3.1), on a donc un temps de pose de 2,5 ms. La longueur équivalente d'un pixel doit alors valoir 3,125 cm. Sachant que le télescope fait 8 m de diamètre, on échantillonne la pupille sur 256 pixels. On déplace horizontalement cette pupille dans un écran de 1024×1024 pixels.

En situation réelle, LIFT verra un tip/tilt turbulent et un focus lié aux variations d'altitude de la couche de sodium. Une série temporelle de tip/tilt turbulent peut être obtenue à partir d'une DSP temporelle théorique, dont la formule est donnée par Jean-Marc Conan [Conan *et al.*, 1995], correspondant à notre configuration. Afin d'inclure cette série dans la boucle, à chaque itération, la phase est décomposée sur les modes de Zernike puis recomposée en remplaçant les coefficients de tip/tilt par ceux de la série temporelle.

La couche de sodium évolue lentement : on peut s'attendre à une variation de focus de l'ordre de quelques dizaines de nm ($\lesssim 0,1$ rad) en 1 minute [Neichel *et al.*, 2013]. Ces variations sont totalement négligeables dans notre cas d'application. Cependant, des météores peuvent traverser l'atmosphère et changer rapidement l'altitude de la couche de sodium dans la direction d'analyse [Herriot et Irvin, 2012]. L'altitude peut varier de 1 km en moins d'une seconde. Pour un télescope de 8 m et une altitude moyenne de la couche de sodium de 90 km, cela correspond à un focus de 0,8 rad rms à une longueur d'onde de $2,2 \mu\text{m}$. La variation reste donc faible à l'échelle de quelques ms. Il faut cependant noter que, pour un télescope de 30 m, ce résultat est à multiplier par 14 (par 24 pour un télescope de 39 m). Dans le contexte des nouveaux grands télescopes, une correction du focus en temps réel sera donc nécessaire. Dans ces simulations, nous avons choisi de garder le focus résiduel donné par les DSP, qui évolue de 0,1-0,2 rad à chaque itération de boucle, afin de s'assurer que LIFT reste opérationnel dans le cas des futurs grands télescopes.

Enfin, on considère un miroir déformable parfait qui corrige le tip/tilt/focus estimé avec un retard de deux trames. La commande est de type intégrateur, c'est-à-dire que le tip/tilt/focus utilisé pour mettre en forme le DM est le tip/tilt/focus estimé multiplié par un coefficient inférieur à 1, nommé gain intégrateur. Le gain choisi est 0,3. La boucle s'est avérée stable avec ce gain.

Le tableau 3.8 récapitule les paramètres de simulation utilisés dans la suite.

3.3.3.2 Exemple d'accrochage

La figure 3.12 montre l'évolution de la variance sur le tip/tilt/focus en fonction du temps lorsqu'on ferme la boucle. La DSP utilisée est celle du cas n° 2 sur axe. Le flux est de 100 photons, et il n'y a pas de bruit de détecteur.

La boucle se ferme correctement. La variance résiduelle sur le tip/tilt est le focus est de $0,045 \text{ rad}^2$ en moyenne lorsque la boucle est fermée. Cette performance est cohérente avec les résultats obtenus au paragraphe 3.3.2.3 p.121.

Paramètre	Valeur
Taille de l'écran	1024 × 1024 pixels
Diamètre de la pupille	256 pixels
Taille du pixel	3,125 cm
Vitesse du vent	12,5 m/s
Temps de pose	2,5 ms
Retard de la boucle	5 ms
Gain intégrateur	0,3

TABLE 3.8 – Paramètres de la simulation d'une boucle fermée.

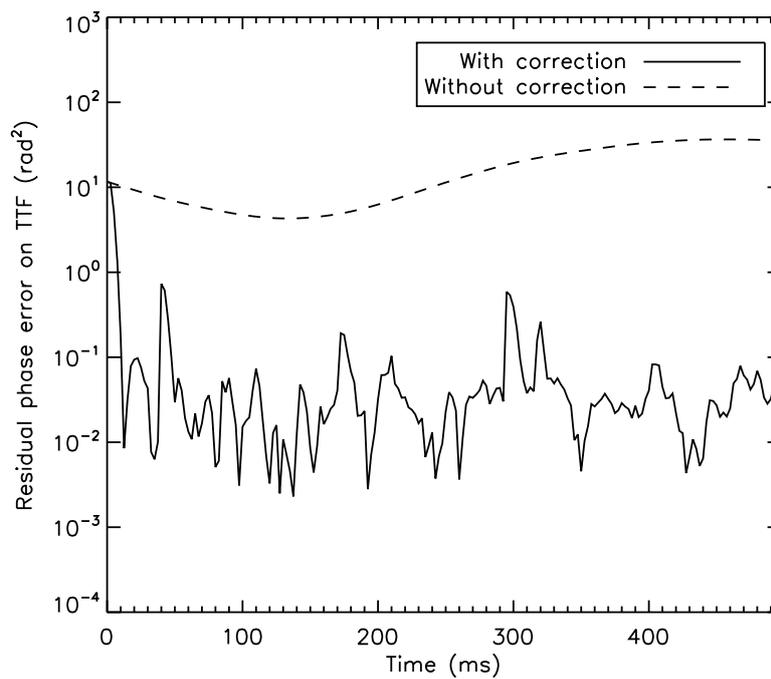


FIGURE 3.12 – Variance résiduelle du tip/tilt/focus avec et sans correction en fonction du temps. Flux = 100 photons. Pas de bruit de détecteur. Taille de la fenêtre : 8 taches d'Airy. 54 modes estimés.

3.3.4 Discussion autour d'une variation de LIFT

Un autre concept d'analyse plan focal linéarisée a été proposée récemment, le *Iterative Linear Phase Diversity* (ILPD), et il a été comparé à LIFT vis-à-vis de la performance de correction en boucle fermée [Smith *et al.*, 2013]. LIFT et le ILPD sont en effet très similaires :

- Les deux algorithmes sont basés sur un développement de Taylor au premier ordre sur la phase d'une image pleine pupille.
- Ils supposent que l'image est le résultat de l'addition d'une phase connue (phase de diversité) et d'une faible phase à estimer.
- Le développement de Taylor est effectué autour de la phase de diversité, comme suggéré dans [Mocœur *et al.*, 2009], puisque que cette phase n'est pas nécessairement faible.

Cependant, les deux méthodes diffèrent sur les applications visées :

- LIFT est conçu pour mesurer des bas ordres de faible amplitude sur des sources très peu lumineuses, d'où le choix d'une phase de diversité de faible amplitude (0,5 rad) (voir §2.3 p.70).
- Le ILPD semble être conçu pour une estimation à fort flux (jusqu'à 10^6 photons), d'ordres plus élevés (14 modes) et d'amplitude relativement élevée (1 rad rms) [Smith *et al.*, 2013], d'où le choix d'une amplitude de phase de diversité de 2 rad pour améliorer la robustesse, au détriment de la sensibilité.

Les deux approches ont donc une logique similaire, et sont intéressantes pour leurs domaines d'application respectifs. La comparaison ILPD/LIFT que l'on peut trouver dans [Smith *et al.*, 2013] n'est toutefois pas légitime pour les raisons suivantes :

- Elle compare une estimation de LIFT basée sur une seule image non corrigée avec une estimation du ILPD qui utilise 5 images consécutives d'une boucle fermée, avec le même rapport signal-à-bruit par image, donc beaucoup plus d'information pour le ILPD.
- Elle utilise LIFT hors de son domaine d'application (faible phase, estimation de bas ordres), ce qui entraîne des problèmes de divergence et de biais.
- La discussion concernant le temps de calcul n'a aucun sens vis-à-vis de LIFT, en particulier dans son domaine d'application : pour l'estimation du tip/tilt et du focus, le temps de calcul est évidemment négligeable, surtout comparé au reste des calculs de la boucle d'optique adaptative (contrôle des hauts ordres...).

D'autre part, l'étude semble insister sur l'intérêt de ne faire qu'une seule itération numérique entre deux itérations de la boucle d'optique adaptative. On peut effectivement se demander s'il est utile de faire les itérations numériques dans LIFT lorsqu'on est en boucle fermée. Pour répondre à cela, j'ai tracé la variance d'erreur résiduelle sur le tip/tilt/focus en fonction du nombre d'itérations numériques pour différents cas (Fig. 3.13).

Le résultat est clair : avec une seule itération, on peut perdre un facteur supérieur à 10 en termes de variance d'erreur d'estimation. On peut cependant noter qu'il n'est pas utile d'itérer au-delà de 4 fois. Les itérations sont rapides, et le gain de temps obtenu en passant de 4 à une seule itération serait inutile dans une boucle d'optique adaptative (cf raisonnement ci-dessus). Il ne faut donc pas se priver du gain apporté par les itérations numériques.

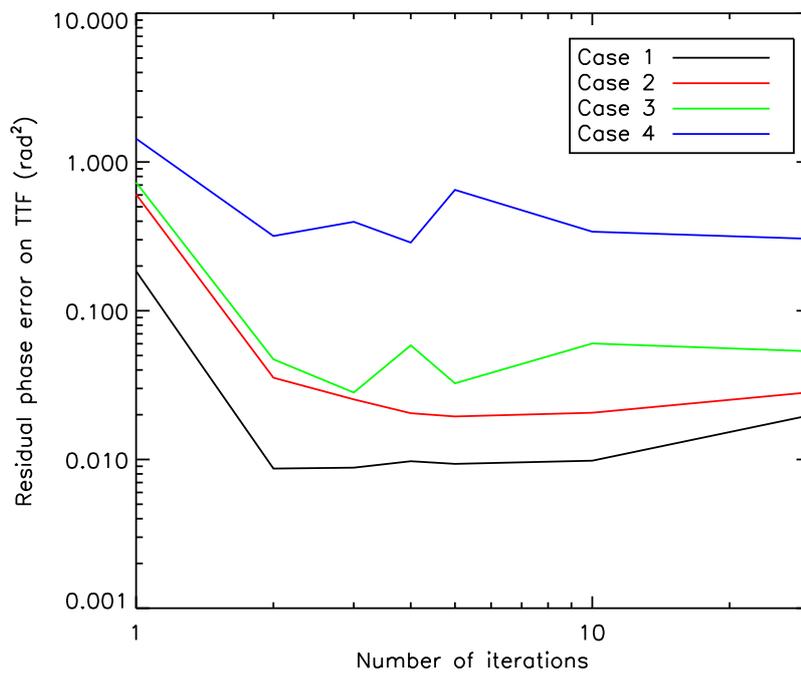


FIGURE 3.13 – Erreur résiduelle sur le tip/tilt/focus en fonction du nombre d’itérations numériques. Flux : 500 photons. Bruit de détecteur : 2 électrons. Variance résiduelle calculée par la méthode Monte-Carlo avec 20 occurrences.

3.4 Perspectives sur GeMS et utilisation de LIFT en optique adaptative

LIFT a été validé en boucle ouverte et bénéficie maintenant d'une méthode de traitement plus robuste, qui a été validée en simulation. Nous regardons maintenant les possibilités de tests pour finaliser la validation de LIFT, et discutons de l'implémentation de LIFT sur un système d'optique adaptative assistée par laser.

Perspectives immédiates : futurs tests en boucle fermée sur GeMS La méthode MAP permet une estimation par LIFT plus robuste en présence de hauts ordres résiduels. Nous pourrions donc essayer de l'appliquer aux données GeMS afin d'obtenir de meilleurs résultats. Cependant, cela ne présente qu'un intérêt réduit, puisque les images ont été acquises à fort flux, cas dans lequel le gain à utiliser un MAP par rapport à un maximum de vraisemblance est limité. À mon sens, le prochain test à faire serait plutôt de vérifier le bon fonctionnement de LIFT en boucle fermée sur un système tel que GeMS, à faible flux. Cela achèverait la validation expérimentale de LIFT. Un plan de test pour cette dernière validation est en cours de construction, en collaboration avec l'équipe de Gemini.

A moyen terme : implémentation de LIFT en LTAO ou MCAO Au delà de la validation en boucle fermée de LIFT sur étoile, cet analyseur a été imaginé dans le cadre des modules de LTAO et MCAO du futur E-ELT. L'implémentation de LIFT dans un système de MCAO est discutée dans l'article reproduit au paragraphe 3.2.3 p.100. Dans ce cas, LIFT bénéficierait de la bonne correction de la MCAO en infrarouge, quelle que soit l'étoile du champ visée, et seule une optique introduisant de l'astigmatisme serait requise en plus de la caméra d'acquisition.

Dans un système de LTAO, l'optique adaptative ne corrige les hauts ordres que dans la direction de l'objet d'intérêt. Dans ce cas, LIFT aura besoin d'un miroir déformable dédié pour bénéficier de cette correction dans la direction de l'étoile naturelle. Le miroir pourra, dans le même temps, introduire l'astigmatisme nécessaire, et pourquoi pas corriger en plus le résidu de focus anisoplanétique entre l'étoile naturelle et la source d'intérêt astronomique (pourvu que la déformée miroir en tip tilt focus soit bien maîtrisée).

Une variante : estimation des bas ordres directement sur le détecteur scientifique ? Étant donnée la simplicité de LIFT, il paraît tentant de l'insérer au plus proche du capteur scientifique. Effectivement, cela permettrait d'éviter les difficultés habituellement rencontrées lors de l'utilisation d'analyseurs spécifiques (aberrations différentielles, chromatisme...). LIFT bénéficierait en outre de la bonne correction dans l'infrarouge.

L'avènement de nouveaux détecteurs incluant une ou plusieurs fenêtres de lecture rapide, tel que le détecteur GSAOI de Gemini Sud [Young *et al.*, 2012] (cf §3.2.1 p.97), nous a inspiré l'idée d'utiliser LIFT directement sur le capteur scientifique. On pourrait alors sélectionner une étoile naturelle proche de l'objet d'intérêt pour analyser le tip/tilt à haute cadence. De plus, en ajoutant un astigmatisme, on aurait aussi une estimation du focus. Il serait cependant difficile de n'ajouter l'astigmatisme que sur l'étoile à analyser. Pour s'assurer de ne pas dégrader la qualité de l'image finale, l'astigmatisme peut alors être ajouté seulement pendant 1% du temps de pose (10 ms toutes les secondes par exemple), en alternant son signe afin de garder une bonne isotropie de la FEP. La perte moyenne en rapport de Strehl serait inférieure à 0,2% (Fig. 3.14),

ce qui est négligeable par rapport aux autres sources d'erreur classiques. De la même manière, la perte d'énergie encadrée sur un carré de taille $\lambda/D \times \lambda/D$ (avec D le diamètre de la pupille) serait aussi inférieure à 0,2%.

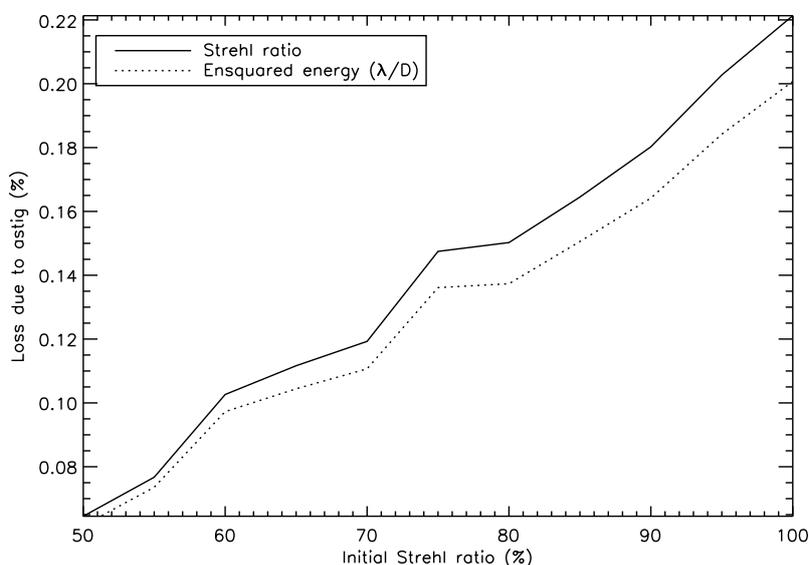


FIGURE 3.14 – Perte sur le rapport de Strehl et l'énergie encadrée lorsqu'on ajoute un astigmatisme de 0.5 rad pendant 1% du temps de pose, en fonction du rapport de Strehl initial. Le côté du carré pour le calcul de l'énergie encadrée est λ/D . Les valeurs sont calculées en moyenne sur 1000 occurrences sans bruit.

On aurait alors un analyseur de tip/tilt, qui analyserait de plus le focus à faible cadence (~ 1 Hz), sans aucun matériel supplémentaire, ni de dégradation de l'image, ni de problème d'aberrations non communes.

3.5 Conclusion

Dans ce chapitre, nous avons montré la première validation de LIFT sur ciel en boucle ouverte avec le système de MCAO du télescope Gemini Sud, GeMS, et nous avons étudié le comportement de LIFT en présence de hauts ordres.

Afin de préparer la validation sur ciel, des tests préliminaires sur sources internes ont permis de s'assurer du bon comportement de LIFT sur le système GeMS. Nous avons ainsi démontré à nouveau la linéarité de l'estimation du focus et vérifié sa répétabilité en présence de hauts ordres résiduels simulés, avec une erreur moyenne inférieure à 0,03 rad dans chaque cas. Nous avons alors acquis des images d'étoile naturelle, le 31 janvier 2013, pour confirmer la linéarité de l'estimation du focus en présence de hauts ordres résiduels réels. Le focus a été estimé en boucle ouverte avec une erreur maximale de 0,1 rad. LIFT a ainsi été validé pour la première fois sur le ciel.

Pour préparer une validation en boucle fermée, je me suis ensuite intéressé de plus près aux effets des hauts ordres sur l'estimation du tip/tilt et du focus.

J'ai tout d'abord montré, grâce à une analyse théorique, que LIFT était majoritairement sensible au repliement des modes d'ordre peu élevé. Afin d'atténuer le repliement, j'ai étudié

le filtrage des hauts ordres par la troncature de la zone de calcul. Cette étude a finalement révélé qu'il était impossible de bien filtrer les modes non estimés sans dégrader l'estimation du tip/tilt et focus. J'ai finalement opté pour une estimation régularisée, qui permet de prendre en compte l'effet des hauts ordres sur la tache et ainsi atténuer le repliement. Cette méthode permet d'avoir une performance d'estimation quasi indépendante de la taille de la zone de calcul et du nombre de modes estimés. Elle est donc robuste

J'ai ensuite déterminé l'effet des hauts ordres sur la propagation du bruit de LIFT et d'un Shack-Hartmann 2×2 pour différentes conditions de turbulence, en simulant des phases résiduelles typiques d'un système de type GeMS. J'ai ainsi montré que LIFT offrait le même gain sur le Shack-Hartmann en propagation du bruit pour n'importe quelles conditions de turbulence. D'après la partie 2.3 p.70, cela implique un gain en magnitude limite de l'ordre de 1,5. La première simulation d'une boucle fermée a ensuite permis de vérifier le bon fonctionnement de LIFT lorsqu'il est soumis à un tip/tilt turbulent et des variations fortes de la couche de sodium, ainsi que l'intérêt de faire plusieurs itérations numériques dans l'estimation.

Nous avons finalement discuté des derniers tests à faire, à savoir une validation en boucle fermée, ainsi que des possibilités d'utilisation de LIFT dans un système de LTAO ou MCAO. En particulier, nous avons évoqué une utilisation directement sur le détecteur scientifique, qui permettrait une estimation du tip/tilt/focus sans problème d'aberrations non communes, tout en ayant une dégradation imperceptible de l'image.

En attendant l'occasion de faire la validation en boucle fermée puis une éventuelle implémentation de LIFT dans un système, quelques questions se posent sur les autres possibilités d'utilisation de LIFT, notamment pour mesurer des hauts ordres. C'est le sujet de la partie qui suit.

Mesurer des hauts ordres avec LIFT ?

Lors de la présentation des performances de LIFT aux différentes conférences ou ateliers auxquels j'ai participé, deux questions revenaient inmanquablement : Comment LIFT se compare-t-il à la Pyramide ? LIFT peut-il mesurer plus de modes que tip, tilt et focus, et si oui, combien ?

Les chercheurs qui posaient ces questions avaient probablement derrière la tête la nécessité d'un "truth sensor", permettant dans les systèmes assistés par étoiles Laser de mesurer sur une étoile naturelle la qualité finale de correction, et ainsi de pouvoir évaluer les nombreux contributeurs au budget d'erreur de tels systèmes. Un capteur de Tip Tilt Focus permettant de jouer aussi le rôle de "truth sensor" serait un excellent candidat. La question était donc de voir si LIFT pouvait tenir ce rôle, et comment il se comparait avec l'analyseur "à la mode" du moment, l'analyseur pyramidal, encore auréolé des beaux résultats du LBT.

Comme souvent, notre réponse a été : "ça dépend". En l'occurrence, cela dépend du niveau d'aberration résiduel présenté à LIFT. Cette méthode est simple et rapide grâce à la linéarisation du modèle de formation des données. Cette linéarisation n'est valide que si la variance de phase est faible, ou que l'on dispose d'une bonne estimée initiale de cette phase. Bien sûr, il existe des méthodes plan-focal ne reposant pas sur la linéarisation, et dont le domaine de validité est supérieur à celui de LIFT. Cependant, cet accroissement du domaine de validité se fait au prix d'une complexité et d'un coût algorithmique accrus.

Notre démarche a alors été de chercher un moyen d'étendre LIFT à des cas forte phase en gardant la linéarisation. Une solution consiste à "diviser pour mieux régner", c'est-à-dire à découper la pupille en sous-zones au sein desquelles la variance de phase est compatible avec l'utilisation de LIFT. En d'autres termes, cela consiste à mettre LIFT derrière chaque sous-pupille d'un analyseur Shack-Hartmann aux micro-lentilles légèrement astigmatiques. Nous avons alors étudié le potentiel d'un tel "Shack-Hartmann LIFTé".

Avec ce nouveau venu dans le bestiaire déjà fourni des analyseurs, il nous a semblé d'autant plus nécessaire de trouver une méthode de comparaison équitable, permettant notamment de répondre à la question posée : comment cette déclinaison hauts-ordres de LIFT se compare-t-elle à l'analyseur pyramidal ?

Chapitre 4

Shack-Hartmann LIFTé : une adaptation de LIFT pour estimer les hauts ordres

4.1 Introduction

Nous savons que, contrairement à la pyramide, LIFT ne permet pas de mesurer des hauts ordres si la phase à mesurer était trop perturbée (cf chapitre 3 p.95). En effet, au-delà d'une variance de phase de 1 rad^2 (hors tip/tilt), une approximation de faible phase ne permet plus d'estimer correctement les aberrations. Cependant, la variance de phase diminue lorsque l'on considère une pupille plus petite. Afin d'estimer de fortes aberrations, on peut donc diviser la pupille en sous-pupilles, à la manière d'un Shack-Hartmann. Pour des conditions de turbulence données, on définit alors un diamètre de sous-pupille limite, noté r_1 , pour lequel la variance de phase à la longueur d'onde d'analyse est toujours inférieure ou égale à 1 rad^2 , hors tip/tilt. En respectant cette taille limite, il est possible d'estimer une phase locale dans chaque sous-pupille avec LIFT puis de reconstruire le front d'onde sur la pupille entière. C'est sur ce principe que repose le Shack-Hartmann LIFTé, qui consiste à utiliser un réseau de sous-pupilles astigmatiques et à estimer plusieurs modes par sous-pupilles avec LIFT.

Le Shack-Hartmann LIFTé est un intermédiaire entre le Shack-Hartmann et les méthodes d'analyse en plan focal. Le Shack-Hartmann a une bonne dynamique, il est simple à utiliser et il peut fonctionner en temps réel. Cependant, l'échantillonnage de la pupille en sous-pupilles mène à des erreurs de repliement et une propagation élevée du bruit. Les méthodes d'analyse en plan focal bénéficient d'une propagation de bruit plus avantageuse que le Shack-Hartmann grâce au gain pleine pupille (cf §1.4.8 p.61), mais la complexité de la relation entre la répartition d'intensité dans l'image et la phase à estimer rend les calculs fastidieux. Afin d'augmenter la vitesse de l'estimation, cette relation peut être linéarisée par une approximation de faible phase, comme dans LIFT (cf §1.4.2 p.56), mais la dynamique s'en retrouve diminuée.

Le Shack-Hartmann LIFTé garde une bonne vitesse d'exécution grâce à l'approximation de faible phase, ainsi qu'une grande dynamique grâce à la division en sous-pupilles. D'autre part, en estimant plusieurs modes par sous-pupille, on obtient une reconstruction du front d'onde plus fine qu'avec un Shack-Hartmann classique, qui estime deux pentes par sous-pupille. Cela permet soit :

- de reconstruire plus de modes sur la pupille globale, à nombre de sous-pupilles égal. On réduit ainsi les effets de repliement et il devient possible de piloter un miroir déformable avec plus d'actionneurs.

- de diminuer le nombre de sous-pupilles, ce qui revient à avoir des sous-pupilles plus grandes, en reconstruisant autant de modes sur la pupille globale. On profite alors du gain pleine pupille et la sensibilité au bruit est fortement diminuée.

Le Shack-Hartmann LIFTé pourrait donc se révéler comme une amélioration notable du Shack-Hartmann. Dans ce chapitre, je m'intéresse au gain apporté en propagation de bruit par rapport au Shack-Hartmann classique.

Dans un premier temps (§4.2 p.134), nous détaillons le principe du Shack-Hartmann LIFTé, puis nous présentons une première comparaison théorique avec le Shack-Hartmann classique, qui a fait l'objet d'une publication reproduite dans ce chapitre (§4.3 p.136). La sensibilité au bruit dans une sous-pupille, issue des analyses du chapitre 2 p.65, ainsi que la possibilité d'estimer de plus hauts ordres y sont étudiées.

J'ai ensuite approfondi cette analyse en m'appuyant sur une reconstruction de type maximum de vraisemblance à partir des estimations de LIFT dans chaque sous-pupille (§4.4 p.141). J'en déduis des coefficients de propagation du bruit pour l'estimation de modes de Zernike et les compare à ceux du Shack-Hartmann classique (§4.5 p.144).

Enfin, je discute des cas d'application du Shack-Hartmann LIFTé (§4.5.1 p.144).

4.2 Principe du Shack-Hartmann LIFTé

Le Shack-Hartmann classique échantillonne la pupille à l'aide d'une matrice de micro-lentilles, formant ainsi des sous-pupilles (cf 1.4.1.1 p.46). Une pente locale du front d'onde est calculée au niveau de chaque micro-lentille en calculant le centre de gravité de sa tache focale. Pour chaque sous-pupille, on a alors une pente dans la direction x , et une pente dans la direction y . Le Shack-Hartmann LIFTé échantillonne la pupille de la même manière, mais avec des micro-lentilles astigmatiques. L'astigmatisme introduit permet de mettre en oeuvre LIFT sur chaque tache focale de micro-lentille. Pour conserver l'approximation de faible phase, la taille des micro-lentilles doit être inférieure à r_1 . Afin d'être cohérent avec la géométrie des micro-lentilles, l'algorithme a été adapté pour estimer des modes de Zernike orthonormalisés sur une pupille carrée. On a ainsi, au lieu de deux pentes locales, une combinaison de coefficients de Zernike locaux, que l'on appellera dans la suite *hyperpentes*. La réponse d'un Shack-Hartmann LIFTé à un front d'onde peut donc s'écrire comme un vecteur d'hyperpentes.

Prenons l'exemple d'un Shack-Hartmann ayant N_{pup} sous-pupilles. Le Shack-Hartmann classique mesure deux pentes $[p_{x,i}, p_{y,i}]$ pour la i -ème sous-pupille. Le Shack-Hartmann LIFTé mesure, par exemple, 5 coefficients de Zernike par sous-pupille $[a_{2,i}, a_{3,i}, a_{4,i}, a_{5,i}, a_{6,i}]$ (tip, tilt, focus et les deux astigmatismes). Le vecteur de pentes du Shack-Hartmann classique \mathbf{Y} et le

vecteur d'hyperpentes du Shack-Hartmann LIFTé \mathbf{Y}^{LIFT} s'écrivent alors :

$$\mathbf{Y} = \begin{bmatrix} p_{x,1} \\ p_{y,1} \\ p_{x,2} \\ p_{y,2} \\ \vdots \\ p_{x,N_{pup}} \\ p_{y,N_{pup}} \end{bmatrix} ; \quad \mathbf{Y}^{\text{LIFT}} = \begin{bmatrix} a_{2,1} \\ a_{3,1} \\ a_{4,1} \\ a_{5,1} \\ a_{6,1} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \\ a_{5,2} \\ a_{6,2} \\ \vdots \\ a_{2,N_{pup}} \\ a_{3,N_{pup}} \\ a_{4,N_{pup}} \\ a_{5,N_{pup}} \\ a_{6,N_{pup}} \end{bmatrix} \quad (4.1)$$

On note $m_1, m_2, \dots, m_{N_{ab}}$ les N_{ab} modes à estimer sur la pupille. L'approche classique pour faire cette estimation nécessite de construire une matrice d'interaction, qui contient la réponse du Shack-Hartmann classique ou Shack-Hartmann LIFTé à chaque mode :

$$M_{int} = [\mathbf{Y}(m_1) | \dots | \mathbf{Y}(m_{N_{ab}})] ; \quad M_{int}^{\text{LIFT}} = [\mathbf{Y}^{\text{LIFT}}(m_1) | \dots | \mathbf{Y}^{\text{LIFT}}(m_{N_{ab}})] \quad (4.2)$$

soit :

$$\begin{aligned}
 M_{int} &= \left(\begin{array}{c} \left[\begin{array}{c} p_{x,1} \\ p_{y,1} \\ p_{x,2} \\ p_{y,2} \\ \vdots \\ p_{x,N_{pup}} \\ p_{y,N_{pup}} \end{array} \right] \\ (m_1) \end{array} \quad \left[\begin{array}{c} p_{x,1} \\ p_{y,1} \\ p_{x,2} \\ p_{y,2} \\ \vdots \\ p_{x,N_{pup}} \\ p_{y,N_{pup}} \end{array} \right] \\ (m_2) \quad \dots \quad \left[\begin{array}{c} p_{x,1} \\ p_{y,1} \\ p_{x,2} \\ p_{y,2} \\ \vdots \\ p_{x,N_{pup}} \\ p_{y,N_{pup}} \end{array} \right] \\ (m_{N_{ab}}) \end{array} \right) \\
 \\
 M_{int}^{LIFT} &= \left(\begin{array}{c} \left[\begin{array}{c} a_{2,1} \\ a_{3,1} \\ a_{4,1} \\ a_{5,1} \\ a_{6,1} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \\ a_{5,2} \\ a_{6,2} \\ \vdots \\ a_{2,N_{pup}} \\ a_{3,N_{pup}} \\ a_{4,N_{pup}} \\ a_{5,N_{pup}} \\ a_{6,N_{pup}} \end{array} \right] \\ (m_1) \end{array} \quad \left[\begin{array}{c} a_{2,1} \\ a_{3,1} \\ a_{4,1} \\ a_{5,1} \\ a_{6,1} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \\ a_{5,2} \\ a_{6,2} \\ \vdots \\ a_{2,N_{pup}} \\ a_{3,N_{pup}} \\ a_{4,N_{pup}} \\ a_{5,N_{pup}} \\ a_{6,N_{pup}} \end{array} \right] \\ (m_2) \quad \dots \quad \left[\begin{array}{c} a_{2,1} \\ a_{3,1} \\ a_{4,1} \\ a_{5,1} \\ a_{6,1} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \\ a_{5,2} \\ a_{6,2} \\ \vdots \\ a_{2,N_{pup}} \\ a_{3,N_{pup}} \\ a_{4,N_{pup}} \\ a_{5,N_{pup}} \\ a_{6,N_{pup}} \end{array} \right] \\ (m_{N_{ab}}) \end{array} \right) \tag{4.3}
 \end{aligned}$$

Une estimation de type maximum de vraisemblance est alors faite à partir de cette matrice. La matrice d'interaction du Shack-Hartmann LIFTé contient plus de données que celle du Shack-Hartmann classique. On peut donc prédire avec une bonne assurance que l'estimation du front d'onde sera plus fine avec le Shack-Hartmann LIFTé. Notamment, dans cet exemple, on est capable d'estimer l'astigmatisme au sein d'une sous-pupille. On peut ainsi estimer le mode gauffre du miroir, ce qui est impossible avec le Shack-Hartmann classique. Le mode gauffre est créé lorsque chaque actionneur a une tension opposée à celle de tous ses voisins. Comme la taille des pupilles du Shack-Hartmann correspond généralement au pas de répartition des actionneurs, on obtient un astigmatisme sur chaque sous-pupille.

4.3 Première analyse théorique

4.3.1 Guide de lecture

Cette étude fait l'objet de l'article reproduit dans le paragraphe suivant. Après une description du principe du Shack-Hartmann LIFTé, l'article montre comment le bruit se propage dans l'estimation de LIFT pour une sous-pupille (Fig. 2 p.2). Cette sensibilité au bruit est comparée à celle des méthodes classiques d'estimation des pentes dans le Shack-Hartmann, le centre de gravité et le centre de gravité pondéré.

Le contrôle d'un miroir déformable est ensuite étudié. La possibilité d'estimer plus finement le front d'onde dans les sous-pupilles implique, par exemple, la possibilité de contrôler un miroir avec plus d'actionneurs. En calculant les valeurs propres de la matrice d'interaction, les différentes possibilités de contrôle d'un miroir déformable sont évaluées. Les valeurs propres quantifient en effet la sensibilité de l'analyseur aux différents modes du miroir.

4.3.2 Article

L'article est reproduit dans les pages suivantes.

Sensing more modes with fewer sub-apertures: the LIFTed Shack–Hartmann wavefront sensor

Serge Meimon,^{1,*} Thierry Fusco,^{1,2} Vincent Michau,¹ and Cédric Plantet¹

¹ONERA—the French Aerospace Laboratory, F-92322 Châtillon, France

²Laboratoire d'Astrophysique de Marseille, 38 rue F. Joliot Curie, 13388 Marseille cedex 13, France

*Corresponding author: serge.meimon@onera.fr

Received February 20, 2014; revised April 5, 2014; accepted April 6, 2014;
posted April 8, 2014 (Doc. ID 206648); published May 5, 2014

We propose here a novel way to analyze Shack–Hartmann wavefront sensor images in order to retrieve more modes than the two centroid coordinates per sub-aperture. To do so, we use the linearized focal-plane technique (LIFT) phase retrieval method for each sub-aperture. We demonstrate that we can increase the number of modes sensed with the same computational burden per mode. For instance, we show the ability to control a 21×21 actuator deformable mirror using a 10×10 lenslet array. © 2014 Optical Society of America

OCIS codes: (010.1080) Active or adaptive optics; (010.1330) Atmospheric turbulence; (110.6770) Telescopes; (030.4070) Modes.

<http://dx.doi.org/10.1364/OL.39.002835>

The Shack–Hartmann (SH) wavefront sensor (WFS) [1] is ubiquitous, from astronomy [2] to biomedical imaging [3], to optical component characterization [4]. It has used widely because it is simple, easy to use, robust to model errors, adapted to a wide range of wavefront amplitudes, and because it works in wide spectral band configurations. It consists of sampling the pupil in sub-apertures by means of a lenslet array and measuring the center of gravity (CoG) of the images at the focus of each lenslet. Measuring the local centroids has two major advantages: the relationship between the centroid and the local wavefront slopes is linear, which makes the exploitation simple and robust; the pixels can be processed in parallel, spot by spot, which leads to a faster pixel-to-wavefront computation (or a faster computation of the deformable mirror control voltages from the pixel values in adaptive optics systems). However, the sampling of the pupil by the lenslet array leads to aliasing errors and lack of sensitivity [5].

On the other hand, focal-plane sensors [6–9] reconstruct the phase from one or several focal images obtained with the full pupil. This leads to a higher sensitivity compared with a SH WFS (full aperture gain), but the relationship between pixel values and phase is more complex. Whereas this relationship can be assumed as linear in a small phase approximation [7], a more sophisticated data formation model is needed when dealing with high spatial frequency phase distortion, high-amplitude wavefronts, or a wide spectral band. This induces a higher computational burden (and slower algorithms) and limited robustness to model errors.

In this Letter, we propose to process SH data by a focal-plane sensor approach. A first approach for performing phase retrieval on SH WFS data has been proposed by Cannon [10], with a Bayesian estimation of the wavefront from the whole SH detector frame. This approach exploits in particular the interference effects between the focal spots, at the cost of a high computational burden because all the pixels have to be processed together. However, these interference effects are usually hardly noticeable.

With our method, the image formation model is considered independent from one sub-aperture to the next.

When the phase variance at the level of a sub-aperture is low enough, it is possible to use a fast and simple focal-plane algorithm in a small phase approximation for each focal spot. The processing of each focal spot is performed separately as in the original SH scheme. A few local phase modes are sensed, instead of only the local mean slope on z and the local mean slope on y . This allows us to measure more information on the phase at the level of the sub-aperture than the local slopes, and thus, to reduce aliasing effects. These LIFT (linearized focal-plane technique) local phase mode estimates are then used in the same way as the local slopes in the classical SH scheme to reconstruct a wavefront or to control a deformable mirror.

The focal-plane image analysis technique we use is the LIFT method [9,11]. LIFT consists of adding a known astigmatism to the wavefront and analyzing the image produced assuming a linear relationship between the image and the aberrations. In the present work, this means using a custom lenslet array with astigmatism on each lenslet. The astigmatism offset avoids the sign ambiguity on even modes. The aberration estimation is performed using a maximum likelihood inversion. We have shown previously that this procedure is adapted to wide spectral band operation [11]. Once the sensor is calibrated, the computation burden is similar to a weighted CoG (WCoG), i.e., to a weighted sum of the image, for each local phase mode sensed.

As any wavefront-sensing technique, the LIFTed SH has a limited working domain. In particular, the sub-aperture size shall have to be small enough to be compatible with the small phase approximation; then, for a given lenslet array, the number of local modes measurable will be limited by noise considerations. We consider here five local modes to demonstrate the feasibility of the LIFTed SH. The local modes chosen here are the first five Zernike modes (tip, tilt, de-focus, both astigmatisms) orthonormalized on a squared sub-aperture. As shown previously [9], the weighing maps used in LIFT for local tip and tilt are similar to a WCoG in the detector noise regime and a CoG in the photon noise regime. Except for LIFT, there exist such maps for not only tip and tilt but also higher modes.

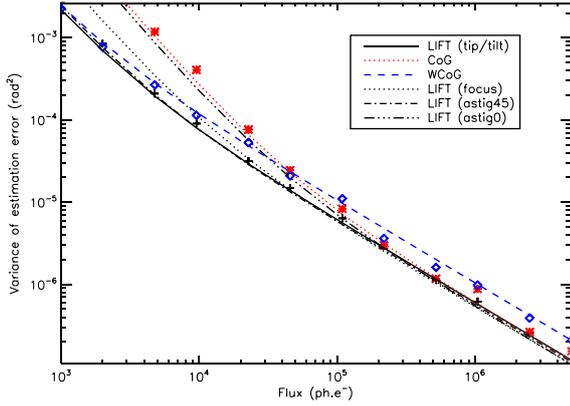


Fig. 1. Noise propagation for a given sub-aperture on local slope estimates with a CoG (red dotted line) and a WCoG (blue dashed line). The noise propagation on local mode estimates with LIFT are plotted in black. Detector readout noise: $10e^-$. For LIFT configuration, a 0.5 rad RMS astigmatism offset is added. The CoG and WCoG curves correspond to diffraction-limited spots (without any astigmatism offset). The cross symbols correspond to a diffractive simulation, whereas the solid lines correspond to the analytic formulas (from [13] for CoG/WCoG).

In a first step, we examine the noise propagation on the local mode estimates. Figure 1 presents the noise propagation on the LIFT local mode estimates. The performance of the WCoG is given for comparison. We can state that whatever the signal-to-noise ratio, it is always beneficiary to use a LIFTed SH with two modes estimated per sub-aperture, rather than a classical SH. Indeed, as LIFT performs a maximum likelihood estimation of the spot position, the noise propagation is better with LIFT than with CoG or WCoG methods. The added astigmatism offset, required for LIFT computation, is small enough (0.5 rad RMS) to not affect the spot full width at half-maximum. Indeed, our results show no performance loss at high flux for LIFT compared with a CoG centroid on a diffraction-limited spot. In the photon noise regime, LIFT noise propagation is the same whatever the mode (as it was observed with the classical phase diversity [12]) and is lower than WCoG noise propagation. When the detector noise is predominant, the noise propagation is higher for focus and astigmatism. Nevertheless, the pixels used for LIFT estimate remain few and the detector noise propagation remains low. Therefore, LIFT focus and astigmatism estimates should not inject noise dramatically in the phase reconstruction process.

Therefore, LIFT allows sensing more modes per sub-aperture than the local wavefront slopes, provided there exists a known astigmatism on each sub-aperture. Retrieving more local information for each sub-aperture leads to estimating more global modes with a given number of sub-apertures, or conversely, to measuring the same number of modes with less sub-apertures. This means a more accurate measurement of the wavefront in aberrometry applications. In adaptive optics applications, it leads to more accurate control of the deformable mirror state, and potentially, to better sensitivity. For instance, in astronomy, loss of sensitivity in SH occurs if the sub-aperture size becomes smaller than r_0 (considering a noiseless detector). There is, therefore, for a SH

WFS a painful trade-off between number of modes and sensitivity. LIFT can be used to regain sensitivity by using fewer large sub-apertures for the same number of modes corrected. Conversely, in cases where the sub-aperture size is bigger than r_0 , a deformable mirror with a smaller pitch could be used to sharpen the spots, thus leading once again to a gain in sensitivity. To properly control all the modes of such a small-pitch mirror, a LIFTed SH method could be used.

This improvement can be quantified by analyzing the interaction matrix M_{int} between a set of aberration modes and the WFS measurement set. Let N_{pup} be the total number of enlightened sub-apertures. For the classical SH, the set of measurements Y includes for each sub-aperture i the two CoG local slopes, denoted by $(s_{x,i}, s_{y,i})$. For the LIFTed SH, the set of measurements Y^{LIFT} includes for each sub-aperture i the five LIFT local mode estimates, denoted by $(a_{1,i}, a_{2,i}, a_{3,i}, a_{4,i}, a_{5,i})$:

$$Y = \begin{bmatrix} s_{x,1} \\ s_{y,1} \\ \vdots \\ s_{x,N_{\text{pup}}} \\ s_{y,N_{\text{pup}}} \end{bmatrix}, \quad Y^{\text{LIFT}} = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ a_{4,1} \\ a_{5,1} \\ \vdots \\ a_{1,N_{\text{pup}}} \\ a_{2,N_{\text{pup}}} \\ a_{3,N_{\text{pup}}} \\ a_{4,N_{\text{pup}}} \\ a_{5,N_{\text{pup}}} \end{bmatrix}. \quad (1)$$

Let $\{m_1, m_2, \dots, m_{N_{\text{ab}}}\}$ be the N_{ab} aberration modes to sense. For an aberrometry application, the modes m_j will be the vectors of a generic wavefront decomposition basis, such as the Zernike polynomials [14]; for an adaptive optics context, the modes m_j will correspond to the deformable mirror modes, i.e., the shapes obtained for each actuator. Then the interaction matrices for the classical SH and the LIFTed SH can be written as follows:

$$M_{\text{int}} = [Y(m_1) | \dots | Y(m_{N_{\text{ab}}})]; \\ M_{\text{int}}^{\text{LIFT}} = [Y^{\text{LIFT}}(m_1) | \dots | Y^{\text{LIFT}}(m_{N_{\text{ab}}})]. \quad (2)$$

The ability of a WFS to accurately sense a given set of modes is conveyed by the eigenvalues of this interaction matrix. In a high-flux operation, the number of modes actually sensed is the number of significant eigenvalues. In Fig. 2, we show the eigenvalues of the interaction matrix corresponding to the influence functions (i.e., the deformable mirror shape modes) of a 21×21 actuator deformable mirror. The dashed black line is obtained for a 10×10 sub-aperture SH ($N_{\text{pup}} \simeq 78$) with classical centroiding. The number of modes estimated is between N_{pup} and $2N_{\text{pup}}$, with N_{pup} being the total number of enlightened sub-apertures (the slope measurements are not independent). The eigenvalues obtained with the same hardware when estimating by LIFT only the local tip and tilt ($a_{1,i}, a_{2,i}$) are shown by the solid black line. The performance of classical SH and LIFTed SH with

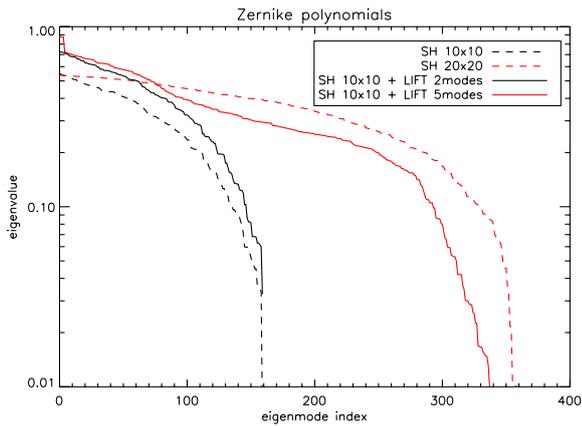


Fig. 2. Eigenvalues of the interaction matrix on a deformable Gaussian influence function on a 21×21 grid (with 25% coupling between modes) for a 10×10 sub-aperture SH WFS with classical centroiding [see Eqs. (1) and (2)], with a two mode per sub-aperture and a five mode per sub-aperture LIFT estimation. For comparison, we have plotted the eigenvalues obtained for a 20×20 sub-aperture SH WFS with classical centroiding.

two modes estimated per sub-aperture is very close. The difference comes from the fact that the local slopes ($s_{x,i}, s_{y,i}$) are not exactly the local tip and tilt on the sub-aperture as estimated by LIFT. Indeed, for instance, the local coma has a contribution to the local slopes, but is orthogonal to the local tip and tilt.

When retrieving with LIFT five modes per sub-aperture (red solid line) instead of two, the performance obtained is very close to a classical SH with 20×20 sub-apertures (red dashed line). With enough flux, it is even possible to control a 21×21 actuator deformable mirror with a LIFTed 10×10 SH WFS.

On the other hand, considering a 21×21 actuator deformable mirror, this procedure allows measuring high-order modes and cancels out the aliasing coming from these modes. For instance, the waffle mode is an unseen mode for a classical SH, as it produces a local astigmatism on each sub-aperture, but no local tip-tilt. With a LIFTed SH procedure, all mirror modes, including the waffle, can be seen.

In conclusion, we have demonstrated that the so-called LIFTed SH WFS allows sensing more modes with less sub-apertures, with a comparable computational burden per mode to sense. For instance, this allows reducing

aliasing by sensing higher order modes with the same hardware or controlling with the same lenslet array a deformable mirror with more actuators. Up to five modes can be sensed per sub-aperture, and for small aberrations, it compares favorably to classical SH in terms of noise propagation. This approach provides a new degree of freedom in designing a SH WFS, as with the procedure we have described, the number of modes sensed is not linked rigidly to the number of sub-apertures. Experimental validation, using a custom-made lenslet array with local astigmatism on each sub-aperture, is planned for 2014. An on-sky demonstration is planned in 2015.

This work was partly funded by the European Commission under FP7 Grant Agreement No. 312430 Optical Infrared Coordination Network for Astronomy and the Office National d'Etudes et de Recherches Aérospatiales (ONERA) in the frame of the NAIADÉ Research Project.

References

1. R. Shack and B. Platt, *J. Opt. Soc. Am. A* **61**, 656 (1971).
2. G. Rousset, in *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge University, 1999), Chap. 5, pp. 91–130.
3. J. Liang, B. Grimm, S. Goelz, and J. F. Bille, *J. Opt. Soc. Am. A* **11**, 1949 (1994).
4. J. A. Koch, R. W. Presta, R. A. Sacks, R. A. Zacharias, E. S. Bliss, M. J. Dailey, M. Feldman, A. A. Grey, F. R. Holdener, J. T. Salmon, L. G. Seppala, J. S. Toepfen, L. V. Atta, B. M. V. Wonterghem, W. T. Whistler, S. E. Winters, and B. W. Woods, *Appl. Opt.* **39**, 4540 (2000).
5. F. Rigaut and E. Gendron, *Astron. Astrophys.* **261**, 677 (1992).
6. R. G. Paxman, T. J. Schulz, and J. R. Fienup, *J. Opt. Soc. Am. A* **9**, 1072 (1992).
7. R. A. Gonsalves, *Opt. Lett.* **26**, 684 (2001).
8. A. Tokovinin and S. Heathcote, *Publ. Astron. Soc. Pac.* **118**, 1165 (2006).
9. S. Meimon, T. Fusco, and L. M. Mugnier, *Opt. Lett.* **35**, 3036 (2010).
10. R. C. Cannon, *J. Opt. Soc. Am. A* **12**, 2031 (1995).
11. C. Plantet, S. Meimon, J.-M. Conan, and T. Fusco, *Opt. Express* **21**, 16337 (2013).
12. L. Meynadier, V. Michau, M.-T. Velluet, J.-M. Conan, L. M. Mugnier, and G. Rousset, *Appl. Opt.* **38**, 4967 (1999).
13. M. Nicolle, T. Fusco, G. Rousset, and V. Michau, *Opt. Lett.* **29**, 2743 (2004).
14. R. J. Noll, *J. Opt. Soc. Am.* **66**, 207 (1976).

4.3.3 Bilan et discussion

Bilan Cet article constitue une première comparaison du Shack-Hartmann LIFTé avec le Shack-Hartmann classique. On y perçoit notamment les avantages du Shack-Hartmann LIFTé vis-à-vis du repliement.

Tout d'abord, nous avons montré les bonnes propriétés de sensibilité au bruit au sein d'une sous-pupille. Sur l'estimation du tip/tilt, LIFT est toujours au moins aussi performant qu'un centre de gravité ou un centre de gravité pondéré. De plus, la variance d'estimation du focus et de l'astigmatisme est identique à celle du tip/tilt en régime de bruit de photons, et n'est que légèrement plus élevée en régime de bruit de détecteur. Le fait d'estimer 5 modes ne devrait donc pas dégrader significativement la propagation du bruit.

Dans un deuxième temps, nous avons montré les possibilités de contrôle d'un miroir déformable avec un Shack-Hartmann LIFTé. Pour cela, nous avons construit des matrices d'interaction pour le contrôle d'un miroir ayant 21×21 actionneurs. Plusieurs configurations ont été considérées : un Shack-Hartmann classique 10×10 , un Shack-Hartmann LIFTé 10×10 estimant 2 ou 5 hyperpentes par sous-pupille et un Shack-Hartmann classique 20×20 . Le calcul des valeurs propres de ces matrices a permis de démontrer que le Shack-Hartmann LIFTé estimant 5 hyperpentes par sous-pupille était plus efficace que le Shack-Hartmann classique 10×10 et quasiment aussi efficace que le Shack-Hartmann classique 20×20 pour contrôler ce miroir. Le Shack-Hartmann LIFTé voit en effet plus de hautes fréquences que le Shack-Hartmann classique 10×10 et les modes du miroir sont mieux vus. Ainsi, les effets de repliement sont fortement atténués.

Discussion Dans cet article, une estimation de 5 hyperpentes a été considéré. On peut facilement imaginer qu'avec plus d'hyperpentes estimées dans chaque sous-pupille, on puisse estimer plus de modes sur la pupille globale. Ainsi, on pourrait éventuellement être capable de contrôler un miroir déformable ayant 21×21 actionneurs avec un Shack-Hartmann LIFTé 10×10 aussi bien, voire mieux, qu'avec un Shack-Hartmann classique 20×20 .

L'avantage vis-à-vis du repliement ayant été démontré, j'étudie dans la suite la propagation de bruit dans un Shack-Hartmann LIFTé pour l'estimation d'une phase pleine pupille et je la compare à celle d'un Shack-Hartmann classique.

4.4 Propagation du bruit dans le Shack-Hartmann LIFTé

Dans cette partie, j'évalue la propagation du bruit dans un Shack-Hartmann LIFTé afin de pouvoir la comparer à celle d'un Shack-Hartmann classique. Je repars de la propagation du bruit au sein d'une sous-pupille puis, à partir d'une reconstruction de type maximum de vraisemblance, je déduis des coefficients de propagation de bruit pour l'estimation de modes de Zernike sur la pupille entière.

4.4.1 Propagation du bruit dans une sous-pupille

L'article au paragraphe 4.3 p.136 donne un aperçu de la propagation du bruit dans une sous-pupille pour l'estimation de 5 hyperpentes. Je détaille ici la matrice de covariance d'erreur d'estimation à l'origine des résultats de l'article, ainsi que la matrice de covariance d'erreur

pour l'estimation de 9 hyperpentes. Ces matrices nous serviront dans la suite pour déduire la propagation sur l'estimation globale du front d'onde. La méthode pour les calculer est montrée dans l'annexe C p.193.

On fait ici le calcul de chaque matrice de covariance pour un flux sur la sous-pupille de 1 photon, d'une part avec du bruit de photons uniquement, d'autre part sans bruit de photons et avec une variance de bruit de détecteur de 1 électron. Les éléments diagonaux de ces matrices sont alors respectivement égaux aux coefficients de propagation du bruit de photons α_i et aux coefficients de propagation du bruit de détecteur β_i sur chaque hyperpente a_i (voir §1.4.5 p.60). On applique ce calcul pour 5 et 9 hyperpentes estimées. Les matrices de covariances obtenues sont décrites par les équations 4.4 et 4.5.

$$\mathbf{C}_b^{5,\text{ph}} = \begin{pmatrix} 0,57 & 0 & 0 & 0 & 0 \\ 0 & 0,57 & 0 & 0 & 0 \\ 0 & 0 & 0,71 & 0 & 0 \\ 0 & 0 & 0 & 0,55 & 0 \\ 0 & 0 & 0 & 0 & 1,0 \end{pmatrix} ; \quad \mathbf{C}_b^{5,\text{det}} = \begin{pmatrix} 13 & 0 & 0 & 0 & 0 \\ 0 & 13 & 0 & 0 & 0 \\ 0 & 0 & 35 & 0 & 0 \\ 0 & 0 & 0 & 17 & 0 \\ 0 & 0 & 0 & 0 & 201 \end{pmatrix} \quad (4.4)$$

et

$$\mathbf{C}_b^{9,\text{ph}} = \begin{pmatrix} 0,61 & 0 & 0 & 0 & 0 & 0 & 0,17 & 0 & -0,036 \\ 0 & 0,61 & 0 & 0 & 0 & 0,17 & 0 & 0,036 & 0 \\ 0 & 0 & 0,71 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,55 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1,0 & 0 & 0 & 0 & 0 \\ 0 & 0,17 & 0 & 0 & 0 & 0,72 & 0 & 0,17 & 0 \\ 0,17 & 0 & 0 & 0 & 0 & 0 & 0,72 & 0 & -0,17 \\ 0 & 0,036 & 0 & 0 & 0 & 0,17 & 0 & 0,70 & 0 \\ -0,036 & 0 & 0 & 0 & 0 & 0 & -0,17 & 0 & 0,70 \end{pmatrix} \quad (4.5)$$

$$\mathbf{C}_b^{9,\text{det}} = \begin{pmatrix} 34 & 0 & 0 & 0 & 0 & 0 & 57 & 0 & -4 \\ 0 & 34 & 0 & 0 & 0 & 57 & 0 & 4 & 0 \\ 0 & 0 & 35 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 17 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 202 & 0 & 0 & 0 & 0 \\ 0 & 57 & 0 & 0 & 0 & 160 & 0 & -14 & 0 \\ 57 & 0 & 0 & 0 & 0 & 0 & 160 & 0 & 14 \\ 0 & 4 & 0 & 0 & 0 & -14 & 0 & 123 & 0 \\ -4 & 0 & 0 & 0 & 0 & 0 & 14 & 0 & 123 \end{pmatrix}$$

On peut noter que la matrice de covariance pour 5 hyperpentes estimées est diagonale dans les deux régimes de bruit. Nous avons vérifié que c'était le cas aussi lorsqu'on estime moins d'hyperpentes, et que les coefficients de propagation pour chaque hyperpente restaient les mêmes. Pour 9 hyperpentes estimées, on perd la diagonalité à cause des corrélations entre les hyperpentes, notamment entre le tip/tilt et la coma. En effet, ces deux modes s'estiment d'une manière similaire, comme on peut le voir sur les cartes de pondérations qui sont appliquées aux données (Fig. 4.1).

Maintenant que la matrice de covariance d'erreur des hyperpentes dans une sous-pupille est quantifiée, nous pouvons en déduire la propagation du bruit sur l'estimation globale.

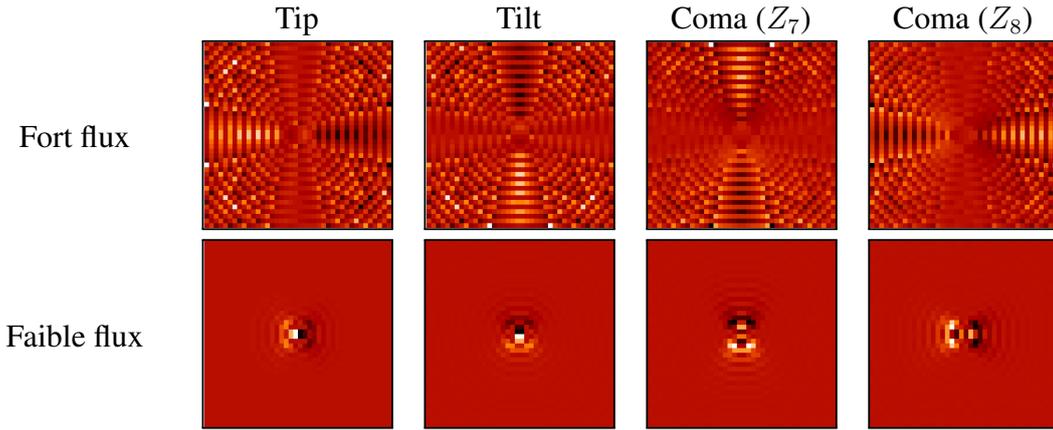


FIGURE 4.1 – Cartes de pondération de LIFT pour l’estimation du tip/tilt et de la coma en fort flux et faible flux.

4.4.2 Propagation du bruit sur l’estimation globale

Dans la suite, \mathbf{S} désigne le vecteur concaténant les hyperpentes de toutes les sous-pupilles et \mathbf{B} le bruit sur ces hyperpentes. On a le modèle suivant :

$$\mathbf{S} = \mathbf{D}\mathbf{A} + \mathbf{B} \quad (4.6)$$

avec \mathbf{D} la matrice d’interaction et \mathbf{A} le vecteur des coefficients de Zernike à estimer sur la pupille entière.

Sous l’hypothèse d’un bruit additif gaussien, l’estimation par Maximum de Vraisemblance donne la solution suivante :

$$\hat{\mathbf{A}} = (\mathbf{D}^t \mathbf{C}_B^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_B^{-1} \mathbf{S} \quad (4.7)$$

avec $\mathbf{C}_B = \langle \mathbf{B}\mathbf{B}^t \rangle$ la matrice de covariance du bruit. En considérant que le bruit est décorrélé d’une sous-pupille à l’autre, celle-ci s’écrit :

$$\mathbf{C}_B = \begin{pmatrix} \mathbf{C}_{b1} & 0 & \dots & 0 \\ 0 & \mathbf{C}_{b2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{C}_{bp} \end{pmatrix} \quad (4.8)$$

avec \mathbf{C}_{b_i} la matrice de covariance d’erreur d’estimation de la i -ème sous-pupille. \mathbf{C}_B est de taille $N_{ab}N_{pup} \times N_{ab}N_{pup}$ avec N_{ab} le nombre d’hyperpentes estimées par sous-pupille et N_{pup} le nombre de sous-pupilles. Pour simplifier les calculs, nous considérons que la propagation du bruit est la même dans chaque sous-pupille. Ainsi, pour tout i , $\mathbf{C}_{b_i} = \mathbf{C}_b$, avec \mathbf{C}_b la matrice de covariance donnée au paragraphe 4.4.1 p.141.

La matrice de covariance de l’erreur d’estimation $\mathbf{E} = \hat{\mathbf{A}} - \mathbf{A}$ est égale à :

$$\langle \mathbf{E}\mathbf{E}^t \rangle = (\mathbf{D}^t \mathbf{C}_B^{-1} \mathbf{D})^{-1} \quad (4.9)$$

Les éléments diagonaux de cette matrice donnent la variance d’erreur d’estimation sur chaque mode.

Cette approche est similaire à celle de Rigaut et Gendron lorsqu'ils ont montré la loi de propagation du Shack-Hartmann classique [Rigaut et Gendron, 1992]. Cependant, contrairement à eux, nous ne faisons pas ici de raisonnement asymptotique pour un nombre de sous-pupille qui tend vers l'infini. Dans un Shack-Hartmann classique, on n'a que deux pentes par sous-pupilles qui ont la même variance d'erreur et qui, pour une phase nulle, ne sont pas corrélées. Ainsi, C_B est l'identité multipliée par la variance d'erreur sur une pente, ce qui permet de simplifier l'expression de $\langle \mathbf{E}\mathbf{E}^t \rangle$. Finalement, on peut déduire directement la propagation du bruit de la diagonale de la matrice $(\mathbf{D}^t\mathbf{D})^{-1}$ [Rigaut et Gendron, 1992].

Dans notre cas, la variance d'erreur sur les hyperpentes n'est pas homogène et, lorsqu'on estime 9 hyperpentes, la matrice C_B n'est pas diagonale (cf §4.4.1 p.141). Il paraît donc difficile de déduire directement la propagation à partir de $(\mathbf{D}^t\mathbf{D})^{-1}$. Nous avons toutefois les outils nécessaires pour calculer la propagation du bruit sur l'estimation globale. Dans la suite, j'évalue cette propagation et la compare à celle du Shack-Hartmann classique.

4.5 Comparaison de la propagation de bruit du Shack-Hartmann classique et du Shack-Hartmann LIFTé

Deux analyseurs sont comparables s'ils sont dans une configuration qui leur permet d'estimer le même nombre de modes. Afin d'avoir une comparaison équitable, je propose donc une règle de dimensionnement du Shack-Hartmann LIFTé. Je donne ensuite un exemple de comparaison où l'on divise le nombre de sous-pupilles du Shack-Hartmann classique par 2.

4.5.1 Proposition de dimensionnement

Considérons un Shack-Hartmann classique avec $N_{pup} \times N_{pup}$ sous-pupilles. Ce Shack-Hartmann a environ $\pi(N_{pup}/2)^2$ sous-pupilles valides et mesure $P \sim \pi N_{pup}^2/2$ pentes. Un Shack-Hartmann LIFTé $N_{pup}/k \times N_{pup}/k$ a environ $\pi(N_{pup}/2k)^2$ sous-pupilles valides et mesure $M \sim N_{ab}\pi(N_{pup}/2k)^2$, N_{ab} étant le nombre d'hyperpentes estimées par sous-pupille. Pour que les deux Shack-Hartmann mesurent le même nombre de modes sur la pupille, il faut vérifier la relation $M \sim P$, ce qui revient à :

$$\pi\left(\frac{N_{pup}}{2k}\right)^2 \times N_{ab} \sim \pi N_{pup}^2/2 \quad (4.10)$$

soit

$$N_{ab} \geq 2k^2 \quad (4.11)$$

Par exemple, si on divise par 2 le nombre de sous-pupilles, il faut estimer environ 8 hyperpentes par sous-pupille.

Dans le paragraphe suivant, j'applique cette règle et je compare un Shack-Hartmann classique 20×20 et d'un Shack-Hartmann LIFTé 10×10 .

4.5.2 Comparaison d'un Shack-Hartmann 20×20 et d'un Shack-Hartmann LIFTé 10×10

D'après le paragraphe précédent, il faut que l'on estime environ 8 hyperpentes par sous-pupilles pour comparer un Shack-Hartmann LIFTé 10×10 et un Shack-Hartmann classique

20×20. Nous avons choisi ici d'estimer 9 hyperpentes, afin d'estimer l'ordre radial 3 en entier. A partir des paragraphes 4.4.1 p.141 et 4.4.2 p.143, on peut alors évaluer la propagation du bruit dans le Shack-Hartmann LIFTé 10×10.

Dans cette partie, je présente les modèles utilisés, puis la comparaison des propagations de bruit. Enfin, je discute des résultats obtenus et des perspectives de cette étude.

4.5.2.1 Modèles

Les deux Shack-Hartmann, classique et LIFTé, sont simulés avec 8×8 pixels par sous-pupille avec un échantillonnage à Shannon. Un astigmatisme de 0,5 rad est ajouté dans chaque sous-pupille pour le Shack-Hartmann LIFTé. Une matrice d'interaction D est ensuite construite pour chaque analyseur. On considère que l'on est à la limite de diffraction et les phénomènes d'interférences entre les sous-pupilles ne sont pas pris en compte. Le champ considéré est suffisant pour que LIFT estime des hyperpentes jusqu'à l'ordre radial 3 (cf §3.3.1 p.110).

Dans le cas du Shack-Hartmann classique, on calcule des coefficients de Rigaut-Gendron non-asymptotiques obtenus à partir de la diagonale de $(\mathbf{D}^t \mathbf{D})^{-1}$. En multipliant cette diagonale par la variance d'erreur d'estimation des pentes, on obtient la propagation du bruit mode à mode. Deux méthodes d'estimation des pentes sont considérées : le centre de gravité et le centre de gravité pondéré. La première méthode est plus efficace en régime de bruit de photons, tandis que la seconde est plus efficace en régime de bruit de détecteur. Nous allons donc, dans la suite, comparer le Shack-Hartmann LIFTé à un Shack-Hartmann classique utilisant le centre de gravité en bruit de photons et le centre de gravité pondéré en bruit de détecteur. Nous utilisons les variances d'erreur décrites au paragraphe 1.4.1.1 p.46 avec $N_t = N_d = N_w = 2$ et $N_s = 4$.

4.5.2.2 Comparaison

La figure 4.2 montre les coefficients de propagation du bruit mode à mode (cf §1.4.5 p.60) pour un Shack-Hartmann classique 20×20 et un Shack-Hartmann LIFTé 10×10 estimant 9 hyperpentes par sous-pupille.

On peut tout d'abord remarquer que la tendance de la propagation du Shack-Hartmann LIFTé est très similaire à celle du Shack-Hartmann classique, que ce soit en bruit de photons ou en bruit de détecteur. Il doit donc être possible de définir des coefficients de type Rigaut-Gendron pour le Shack-Hartmann LIFTé. Cependant, nous n'avons pas réussi à retrouver une relation simple entre la propagation du bruit globale et la propagation du bruit dans une sous-pupille à partir de la matrice d'interaction D . Cela reste donc une perspective.

D'autre part, le gain apporté par le Shack-Hartmann LIFTé 10×10 sur le Shack-Hartmann classique 20×20 est important : la variance d'erreur est divisée par environ 3,2 en bruit de photons et par environ 2,6 en bruit de détecteur, de manière quasi-homogène sur tous les modes.

Il est donc possible de diminuer fortement la sensibilité au bruit de l'analyseur de front d'onde en remplaçant un Shack-Hartmann classique par un Shack-Hartmann LIFTé ayant moins de sous-pupilles.

4.5.2.3 Discussion

Gain pleine pupille L'amélioration apportée par le Shack-Hartmann LIFTé est lié au gain pleine pupille (cf 1.4.8 p.61). Chaque sous-pupille du Shack-Hartmann LIFTé est divisée en k^2 sous-pupilles dans le Shack-Hartmann classique. D'après le paragraphe 1.4.8 p.61, cela

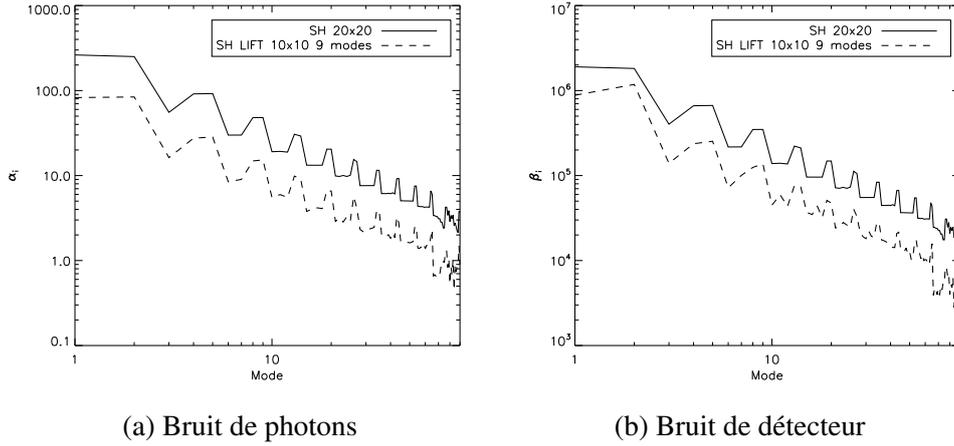


FIGURE 4.2 – Propagation du bruit dans un SH classique 20x20 par rapport à un SH LIFTé 10x10 estimant 9 hyperpentes par sous-pupille. Le coefficient α_i est la variance d’erreur sur le coefficient de Zernike a_i pour un flux total de 1 photo-électron en présence de bruit de photons uniquement. Le coefficient β_i est la variance d’erreur sur le coefficient de Zernike a_i pour un flux total de 1 photo-électron et un écart-type de bruit de détecteur égal à 1 électron en présence de bruit de détecteur uniquement.

implique un gain de k^2 en bruit de photons et k^4 en bruit de détecteur sur l’estimation de la pente locale du front d’onde. Comme on estime plusieurs hyperpentes par sous-pupille, ces gains sont diminués. En effet, nous verrons dans le chapitre 5 p.149 qu’un analyseur de type diversité de phase, comme LIFT, est plus sensible au bruit de détecteur lorsque plus de modes sont estimés, et n’est pas également efficace sur tous les modes en bruit de photons. Dans notre exemple de comparaison, le gain est de $k^2 \times 0,8$ en bruit de photons et $k^4 \times 0,16$. On est proche du gain attendu en bruit de photons, mais le gain pleine pupille est majoritairement perdu en bruit de détecteur.

Amplitude de l’astigmatisme L’amplitude d’astigmatisme utilisée dans ces simulations (0,5 rad) est optimisée pour l’estimation de 3 modes (cf §2.3 p.70). Lorsqu’on estime plus de modes, il faut regarder des pixels plus éloignés du centre de la tache. Afin d’augmenter le rapport signal-à-bruit sur les pixels périphériques, on peut étaler la tache en augmentant l’amplitude de diversité. Ainsi, pour un nombre d’hyperpentes estimés par sous-pupille, il existe une amplitude de diversité optimale (voir chapitre 5 p.149). Dans notre exemple de comparaison, le gain aurait donc sûrement pu être encore amélioré. Par ailleurs, l’astigmatisme pourrait être optimisé pour l’estimation d’un nombre d’hyperpentes supérieur à 9, ce qui permettrait une division du nombre de sous-pupilles plus importante, donc un gain en termes de propagation de bruit plus élevé. Cela est bien sûr limité par l’augmentation de la sensibilité au bruit de détecteur avec le nombre d’hyperpentes estimées, ainsi que par le diamètre r_1 à ne pas dépasser pour les sous-pupilles.

Contrôle d’un miroir Le diamètre des sous-pupilles d’un Shack-Hartmann classique correspond généralement au pitch (pas inter-actionneur) du miroir. Ainsi, pour contrôler le même miroir qu’un Shack-Hartmann classique $N_{pup} \times N_{pup}$, un Shack-Hartmann LIFTé $N_{pup}/k \times N_{pup}/k$

doit être dimensionné de manière à respecter la relation $k \leq r_1/pitch$.

Estimation sur l'image complète Une estimation sur l'image complète du détecteur, comme proposée par Cannon [Cannon, 1995] pour le Shack-Hartmann classique, peut évidemment être envisagée pour le Shack-Hartmann LIFTé afin d'améliorer sa performance. Cependant, les temps de calcul seraient très élevés.

Nous avons maintenant un bon aperçu des caractéristiques du Shack-Hartmann LIFTé. Nous pouvons alors nous demander dans quels cas cet analyseur pourrait être utilisé.

4.6 Cas d'application du Shack-Hartmann LIFTé

Les cas d'application du Shack-Hartmann LIFTé sont quasiment aussi variés que ceux du Shack-Hartmann classique. Il pourra par exemple apporter une meilleure précision sur le front d'onde pour la métrologie optique ou la focalisation d'un laser de puissance, avec une économie sur le nombre de sous-pupilles et la taille du capteur.

Le Shack-Hartmann LIFTé peut aussi être utilisé dans les systèmes d'optique adaptative classiques et extrêmes. Dans une optique adaptative classique, on cherchera à gagner en propagation de bruit pour augmenter la couverture de ciel. Dans une optique adaptative extrême, si le flux de l'étoile le permet, on pourra s'intéresser plutôt à l'atténuation du repliement apporté par le Shack-Hartmann LIFTé pour contrôler un miroir déformable avec plus d'actionneurs.

Il faut tout de même souligner que LIFT n'est pas en mesure d'estimer un objet étendu. Ce n'est pas réellement une limitation puisqu'en présence d'un objet étendu, le gain pleine pupille est fortement diminué, voire perdu. Les applications se limitent donc aux cas où l'on observe une source ponctuelle.

4.7 Conclusion

Dans ce chapitre, nous avons proposé une variation de LIFT permettant d'estimer des hauts ordres sur une étoile naturelle. Le Shack-Hartmann LIFTé est un analyseur simple à utiliser et moins sensible au bruit que le Shack-Hartmann classique.

Une première étude a permis de comparer un Shack-Hartmann classique avec un Shack-Hartmann LIFTé estimant 5 hyperpentes par sous-pupille. Au sein d'une sous-pupille, la variance d'erreur d'estimation des deux premières hyperpentes est semblable à celle des pentes du Shack-Hartmann classique. La variance d'erreur sur les hyperpentes suivantes est similaire en bruit de photons, et ne dépasse pas celle d'un centre de gravité classique en bruit de détecteur. D'autre part, en calculant les valeurs propres de la matrice d'interaction, nous avons montré qu'un Shack-Hartmann LIFTé 10×10 était capable de contrôler un miroir déformable avec 21×21 actionneurs beaucoup mieux qu'un Shack-Hartmann classique 10×10 , et presque aussi bien qu'un Shack-Hartmann classique 20×20 . Cela confirme le gain du Shack-Hartmann LIFTé par rapport au Shack-Hartmann classique en termes de repliement.

J'ai ensuite approfondi l'analyse de propagation de bruit en calculant des coefficients de propagation de bruit pour l'estimation de modes de Zernike. Cela m'a permis de caractériser la performance d'un Shack-Hartmann LIFTé de manière simple et de la mettre en relation avec celle d'un Shack-Hartmann classique. J'ai ainsi démontré que l'on pouvait gagner un facteur

d'environ 3,2 en bruit de photons et d'environ 2,6 en bruit de détecteur sur la performance d'un Shack-Hartmann classique en divisant le nombre de sous-pupilles par 2 et en estimant 9 hyperpentes par sous-pupille.

Enfin, nous avons proposé quelques cas d'application du Shack-Hartmann LIFTé, comme la métrologie optique ou l'optique adaptative extrême.

Nous avons donc maintenant un nouvel analyseur qui vient s'ajouter à ceux existant. Pour une application en optique adaptative classique ou extrême, de nombreux analyseurs peuvent alors être envisagés. L'analyseur choisi devra répondre à un certain nombre de critères qui sont déterminés par la qualité de correction visée. Dans le chapitre suivant, nous proposons de guider ce choix en comparant les analyseurs en termes de propagation du bruit.

Chapitre 5

Une comparaison équitable entre analyseurs pour l'estimation de hauts ordres ?

5.1 Introduction

La performance d'un analyseur de front d'onde en optique adaptative résulte d'un compromis entre la dynamique, le nombre de modes estimés et la sensibilité au bruit. Il faut non seulement choisir un analyseur, mais aussi sa configuration et la méthode de traitement des données pour faire au mieux ce compromis. Les possibilités lors du dimensionnement d'un système sont donc nombreuses et il est utile d'effectuer des comparaisons entre analyseurs afin d'affiner le choix.

Dans la littérature, la figure de mérite généralement utilisée pour comparer les analyseurs est la matrice de covariance d'erreur d'estimation. Celle-ci donne une évaluation de la propagation du bruit. C'est en effet cette caractéristique qui permet de discriminer deux analyseurs dans des configurations similaires en termes de linéarité et de nombre de modes estimés. Cependant, la propagation du bruit ne dépend pas seulement de la configuration de l'analyseur, mais aussi de la méthode de traitement des données.

Dans ce chapitre, nous proposons une comparaison équitable en sensibilité au bruit. Pour cela, nous nous appuyons sur la matrice d'information de Fisher, présentée dans la partie 1.4.6 p.60. La matrice d'information de Fisher permet d'évaluer une matrice de covariance d'erreur directement liée à la nature des données d'un analyseur de front d'onde. On peut montrer, dans le cas d'un bruit gaussien, que cette approche est équivalente à calculer la matrice de covariance d'erreur pour une estimation de maximum de vraisemblance (cf §1.4.7 p.61). Nous privilégions malgré cela la matrice d'information de Fisher, qui est un concept plus général.

Nous analysons tout d'abord les études de propagation de bruit et comparaisons présentes dans la littérature (§5.2). Cela donnera une idée précise des comportements respectifs de chaque analyseur et servira de référence pour les résultats de ce chapitre.

Ensuite, afin de bien appréhender la méthode de la matrice d'information de Fisher, nous l'appliquons au Shack-Hartmann, dont le comportement est bien connu (§5.3). Nous étudions notamment la propagation du bruit lorsque les données sont les pentes locales, estimées par diverses méthodes, ou l'image complète sur le détecteur.

Dans un troisième temps, nous présentons une comparaison entre le Shack-Hartmann, la

pyramide et le Shack-Hartmann LIFTé (§5.4). Le but de cette étude est de revisiter la comparaison entre le Shack-Hartmann et la pyramide, et de voir comment le Shack-Hartmann LIFTé se place par rapport à ces analyseurs.

Enfin, j'étends la comparaison aux autres analyseurs qui peuvent être envisagés en optique adaptative pour l'astronomie (§5.5).

5.2 Analyse de l'état de l'art

L'analyseur le plus souvent pris comme référence dans les comparaisons est le Shack-Hartmann. Rigaut et Gendron se sont intéressés à la matrice de covariance d'erreur donnée par l'estimation de modes de Zernike avec cet analyseur [Rigaut et Gendron, 1992]. En considérant le cas limite où le nombre de sous-pupilles tend vers l'infini, ils ont défini des coefficients de propagation de bruit qui permettent de déduire la variance d'erreur d'estimation sur chaque mode à partir de la variance d'erreur d'estimation de la pente dans chaque sous-pupille. Ces coefficients sont proportionnels à $(n + 1)^{-2}$, n étant l'ordre radial du mode de Zernike considéré.

La diversité de phase a aussi une place importante dans la littérature. Une étude basée sur l'information de Fisher a été effectuée par Lee *et al.*, afin de déterminer les paramètres optimaux de réalisation d'une diversité de phase [Lee *et al.*, 1997]. Les auteurs calculent la matrice de Fisher pour l'estimation de 8 modes de Zernike à partir du focus (Z_4 à Z_{11}). Ils semblent montrer que :

- dans leur cas d'application, il existe une amplitude de diversité optimale.
- la propagation du bruit augmente lorsqu'on utilise plus de deux images.
- l'estimation est meilleure lorsque les deux images sont placées de manière symétrique par rapport au foyer.
- une division du flux équitable entre les deux images est optimale lorsque l'on a une image focale et une image extra-focale.

D'autre part, la propagation du bruit pour l'estimation des modes de Zernike a été évaluée par Meynadier *et al.* [Meynadier *et al.*, 1999], sous l'hypothèse d'un bruit uniforme. Cette étude a confirmé l'existence d'une amplitude de diversité optimale, qui dépend du nombre de modes que l'on veut estimer. Elle a aussi montré que :

- plus le mode à estimer est d'ordre élevé, plus le coefficient de propagation pour ce mode est élevé.
- lorsque le nombre de modes à estimer augmente, tous les coefficients de propagation mode à mode augmentent.

Schulz *et al.* ont démontré, via l'information de Fisher, que la variance d'erreur donnée par la diversité de phase pouvait être environ 3 à 6 fois plus faible que celle donnée par un Shack-Hartmann selon l'amplitude des aberrations à mesurer [Schulz *et al.*, 1999].

A la frontière entre le Shack-Hartmann et la diversité de phase, on trouve l'analyseur de courbure. Il dispose de sous-pupilles, modélisées par des détecteurs individuels, mais il utilise deux plans de mesure extra-pupillaires. La propagation du bruit dans l'analyseur à courbure pour l'estimation de modes de Zernike est proportionnelle à $(n + 1)^{-4}$ [Rousset, 1999]. Cet analyseur a fait l'objet d'une comparaison avec la diversité de phase, effectuée par Fienup *et al.*, en s'appuyant sur l'information de Fisher [Fienup *et al.*, 1998]. Les auteurs considèrent différentes amplitudes de diversité, qui vont du régime de la diversité de phase jusqu'au régime

de l'analyseur de courbure. L'erreur d'estimation diminue jusqu'à environ 1λ de phase de diversité (régime de la diversité de phase) et augmente avec l'amplitude ensuite. Fienup *et al.* ont ainsi montré que la diversité de phase était beaucoup plus efficace. Il faut cependant noter que le raisonnement a été fait pour un analyseur de courbure utilisant une caméra CCD, alors que les détecteurs des analyseurs de courbure actuels sont des assemblages de photodiodes à avalanches, qui n'ont pas de bruit de détecteur.

L'analyseur à pyramide s'est imposé comme un concurrent sévère du Shack-Hartmann. Une première analyse de sa sensibilité au bruit a été faite par son inventeur, Ragazzoni, et Farinato [Ragazzoni et Farinato, 1999]. Celle-ci a montré que, pour l'estimation d'un mode de Zernike donné, la variance d'erreur de la pyramide non modulée est $(D/nd)^2$ moins élevée que la variance d'erreur du Shack-Hartmann, D étant la taille de la pupille, d la taille d'une sous-pupille du Shack-Hartmann et n l'ordre radial du mode de Zernike. En optique adaptative, on a généralement $d \sim r_0$. Cette différence est une conséquence directe du gain pleine pupille (cf §1.4.8 p.61). Ainsi la pyramide non modulée est beaucoup plus efficace que le Shack-Hartmann pour l'estimation des bas ordres et devient équivalente pour l'estimation des hauts ordres ($n \sim D/r_0$). Une analyse plus poussée de la propagation du bruit dans la pyramide a été faite par Vérinaud [Vérinaud, 2004]. Vérinaud s'est notamment intéressé à la nature des mesures issues de la pyramide. A partir d'un raisonnement 1D, il a montré que la pyramide était un senseur de pentes, comme le Shack-Hartmann, pour les basses fréquences spatiales, et un senseur de phase pour les hautes fréquences spatiales. La limite entre ces deux régimes se décale vers les hautes fréquences spatiales lorsqu'on augmente l'amplitude de la modulation. La propagation de la pyramide ressemble finalement de plus en plus à celle du Shack-Hartmann au fur et à mesure qu'on augmente l'amplitude de modulation. Vérinaud semble ensuite montrer que, pour un rayon de modulation de $\lambda/2d$, la propagation de la pyramide est équivalente à celle d'un Shack-Hartmann utilisant des quad-cells pour estimer les pentes. Au contraire, sans modulation, la pyramide est également sensible à toutes les fréquences spatiales, à l'exception d'une remontée légère de la variance d'erreur dans les hautes fréquences, due à un effet de filtrage par intégration du flux dans chacune des sous-pupilles, qui sont les pixels du détecteur dans le cas de la pyramide. Ces calculs de propagation ont été faits dans l'hypothèse d'un bruit uniforme sur toutes les sous-pupilles.

Enfin, on peut citer un dernier analyseur, le masque de Zernike. Paterson a montré que l'image donnée par un masque de Zernike permettait de maximiser l'information de Fisher, et donc d'avoir une variance d'erreur minimale [Paterson, 2008, Dohlen *et al.*, 2006]. Cet analyseur serait donc idéal. Il présente toutefois l'inconvénient de ne pouvoir estimer que des phases très faibles.

Une comparaison théorique de la propagation du bruit de tous ces analyseurs peut être trouvée dans un article de Guyon [Guyon, 2005]. Bien qu'il ne dise pas explicitement s'appuyer sur l'information de Fisher, on peut montrer que son approche est équivalente. Il calcule en effet la dérivée du signal donné par les analyseurs par rapport à des coordonnées spatiales, et les pondère par la variance de bruit. Cela revient à calculer les éléments de la matrice d'information de Fisher pour l'estimation de fréquences spatiales. Pour comparer les analyseurs dans des conditions équivalentes, il optimise la configuration de chacun pour l'estimation d'une fréquence spatiale donnée. Voici les principaux résultats de son article :

- La propagation du bruit du Shack-Hartmann est en f^{-2} , avec f la fréquence. C'est effectivement équivalent à une propagation en $(n+1)^{-2}$ sur les modes de Zernike (voir §3.3.1 p.110).

- La propagation de l'analyseur de courbure semble bien être en f^{-4} jusqu'à sa fréquence optimale, pour laquelle sa propagation atteint la limite inférieure de tous les analyseurs. L'analyseur n'est ensuite sensible qu'aux harmoniques de cette fréquence.
- La pyramide fixe et l'analyseur plan focal qu'il définit (assimilable à la diversité de phase) ont la même propagation et sont également sensibles à toutes les fréquences. Ceci n'est pas tout à fait exact pour la pyramide, à cause de l'effet de filtrage par l'intégration du flux dans chacune des sous-pupilles.
- Le masque de Zernike a une propagation optimale à toutes les fréquences, qui définit la limite inférieure de tous les analyseurs.

Cette bibliographie n'est bien sûr pas exhaustive, mais résume les principaux résultats connus. Il est important de souligner qu'à l'exception du Shack-Hartmann et de l'étude de la diversité par Meynadier, ces analyses ne considèrent que le bruit de photons et n'abordent pas le cas du bruit de détecteur. L'étude présentée dans ce chapitre prend en compte ces deux bruits.

Dans la suite, nous appliquons la méthode de calcul de propagation du bruit avec l'information de Fisher présentée dans la partie 1.4.6 p.60. Notamment, nous revisitons la plupart des résultats décrits précédemment, et nous revenons sur la propagation du bruit du Shack-Hartmann LIFTé. On considérera que l'on est proche de la limite de diffraction, notre point de fonctionnement sera donc la phase nulle et les relations phase-signal seront supposées linéaires.

5.3 Information de Fisher : application au Shack-Hartmann

Dans cette partie, nous nous intéressons uniquement au Shack-Hartmann. Son principe, ainsi que la méthode de simulation et le modèle de bruit utilisés, ont été expliqués au paragraphe 1.4.1.1 p.46. Nous proposons de calculer la matrice d'information de Fisher pour deux types de données : les pentes habituellement utilisées, ou l'image globale. L'image globale correspond à des données non traitées, ce sont les "données brutes", tandis que les pentes constituent déjà un traitement des données brutes, ce sont des "pseudo-données". Nous utiliserons cette appellation pour tous les analyseurs dans la suite. Le modèle de bruit pour les données brutes sera à chaque fois le modèle gaussien décrit au paragraphe 1.2.1.1 p.22.

Nous présentons d'abord les paramètres de simulation, puis nous montrons la propagation obtenue pour différents types de données.

5.3.1 Paramètres de simulation

Les simulations sont faites pour un Shack-Hartmann 40×40 sous-pupilles. L'échantillonnage est à Shannon, avec 16×16 pixels par sous-pupille. On suppose qu'on est à la limite de diffraction, les variances sont donc calculées pour $N_t = N_d = N_w = 2$ pixels. Nous prenons $N_w = N_t$ car c'est le réglage optimal pour le centre de gravité pondéré [Thomas *et al.*, 2006, Nicolle, 2006]. N_s est arbitrairement fixé à 4 pixels, qui est un compromis entre la sensibilité au bruit de détecteur et une dynamique correcte pour le centre de gravité.

5.3.2 Propagation du bruit dans le Shack-Hartmann

Nous évaluons maintenant les coefficients de propagation du bruit d'un Shack-Hartmann 40×40 pour l'estimation de 1000 modes de Karhunen-Loève sur source ponctuelle en lumière monochromatique. Les modes de Karhunen-Loève sont calculés à partir de la méthode décrite par Cannon [Cannon, 1996]. Nous comparons d'abord la propagation à partir des pentes, puis de l'image complète.

5.3.2.1 Pentès

La figure 5.1 montre les coefficients de propagation du bruit mode à mode pour une estimation à partir des pentes.

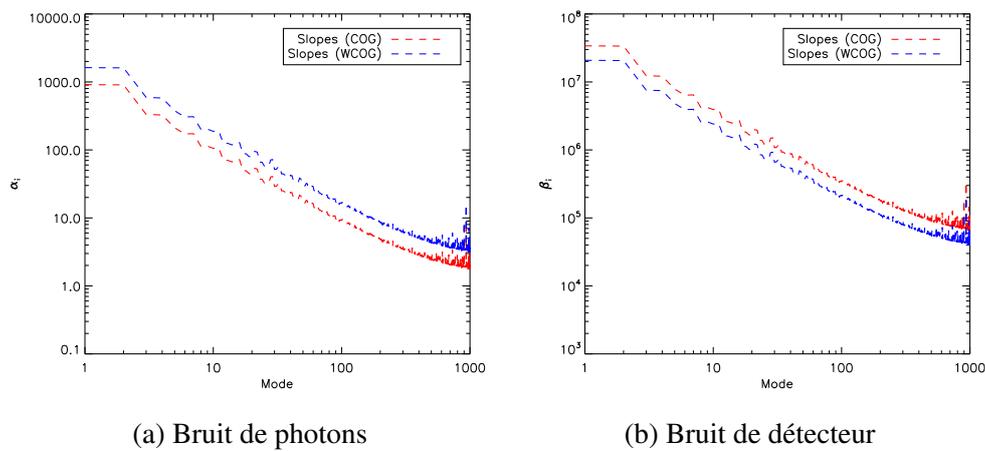


FIGURE 5.1 – Propagation du bruit dans le Shack-Hartmann lorsqu'on mesure des pentes avec un centre de gravité (COG) ou un centre de gravité pondéré (WCOG).

La propagation est en j^{-1} , avec j le mode considéré. Sachant que $j \sim (n + 1)^2$, n étant l'ordre radial, la propagation est donc en $(n + 1)^{-2}$, comme attendu. L'amplitude des coefficients dépend simplement de la variance d'erreur d'estimation d'une pente. Le centre de gravité pondéré est la moins bonne technique en bruit de photons car seul le centre de la tache est exploité. C'est d'ailleurs pour cette raison qu'il est meilleur que le centre de gravité en bruit de détecteur.

Le quad-cell doit être vu comme une technique à part, puisqu'il ne se base pas sur les mêmes images. En effet, le centre de gravité et le centre de gravité pondéré sont calculés à partir d'images échantillonnées à Shannon, tandis que le quad-cell forme une image très sous-échantillonnée. Ainsi, il y a plus d'informations dans les images qu'utilise le centre de gravité, la propagation en bruit de photons est donc meilleure. A l'inverse, comme on se limite à 4 pixels, le quad-cell est la technique qui permet un minimum de propagation du bruit de détecteur.

5.3.2.2 Image complète

La figure 5.2 montre les coefficients de propagation du bruit mode à mode pour une estimation à partir de l'image complète au plan focal des micro-lentilles. Les propagations pour le

centre de gravité et le centre de gravité pondéré sont rappelées pour comparaison.

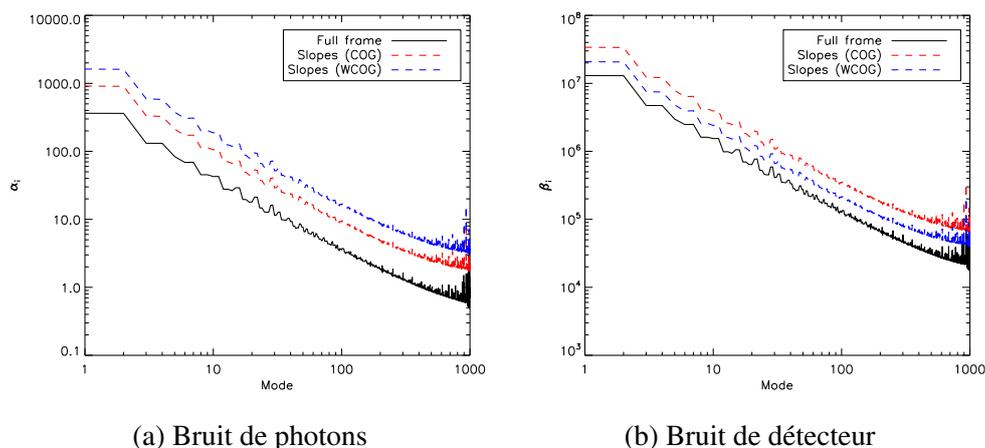


FIGURE 5.2 – Propagation du bruit dans le Shack-Hartmann lorsqu’on utilise les données brutes. Les propagations pour le centre de gravité et le centre de gravité pondéré sont tracées pour comparaison.

La propagation du bruit pour une estimation à partir de l’image complète est meilleure que celles obtenues avec les pentes. L’image complète contient en effet plus d’informations sur les aberrations, comme l’avait déjà montré Cannon [Cannon, 1995]. Le gain est supérieur à 2 par rapport à une estimation à partir des pentes, que ce soit en bruit de photons ou en bruit de détecteur. On peut noter que l’allure de la courbe reste identique, bien que l’on n’estime pas de pentes locales. La loi en $(n + 1)^{-2}$ semble donc être due à la nature même du Shack-Hartmann.

5.4 Comparaison du Shack-Hartmann, de la pyramide et du Shack-Hartmann LIFTé

5.4.1 Guide de lecture

La comparaison du Shack-Hartmann, de la pyramide et du Shack-Hartmann LIFTé a fait l’objet de l’article reproduit dans les pages suivantes (en cours de soumission). La majorité de la partie 3 de l’article se retrouve dans la thèse, dans ce chapitre et dans le chapitre 1 p.21.

Une première partie (§3.1) décrit les méthodes de simulation et les modèles de bruit utilisés pour les trois analyseurs. Pour la pyramide et le Shack-Hartmann, ces informations se retrouvent dans la partie 1.4 p.46. Les paramètres de simulation sont aussi indiqués dans cette partie. Dans le cas du Shack-Hartmann, il s’agit des mêmes paramètres que dans le paragraphe 5.3 p.152.

La partie 3.2 montre ensuite le lien entre la matrice d’information de Fisher et la matrice de covariance d’erreur d’un estimateur de maximum de vraisemblance. Ce lien est déjà évoqué dans le paragraphe 1.4.7 p.61 et démontré dans l’annexe D.2 p.196. La partie 3.3 reprend le paragraphe 1.4.5 p.60.

Nous faisons ensuite une première comparaison entre le Shack-Hartmann et la pyramide, pour l’estimation de 1000 modes de Karhunen-Loève en lumière monochromatique, avec 40×40

sous-pupilles (§4). Enfin, un Shack-Hartmann LIFTé avec 20×20 sous-pupilles est ajouté dans la comparaison (§5).

5.4.2 Article

L'article est reproduit dans les pages suivantes.

Revisiting the comparison between the Shack-Hartmann and the pyramid wavefront sensor via the Fisher information matrix

C. Plantet, S. Meimon, J.-M. Conan, T. Fusco

July 29, 2014

1 Abstract

The choice of a wavefront sensor for an adaptive optics system is the result of a compromise between many parameters such as linearity range and sensitivity to noise. To guide this choice, we propose a comparison of noise propagations in wavefront sensors based on the Fisher information matrix, which quantifies the information present in the data. We compare the Shack-Hartmann, widely used in adaptive optics, and the pyramid wavefront sensor, which should lead to a better performance. We quantify their respective noise propagations for the estimation of Karhunen-Loève polynomials and show the gain provided by the pyramid in photon noise and read-out noise. We then extend the comparison to a recently proposed variation of the Shack-Hartmann, the LIFTed Shack-Hartmann. The LIFTed Shack-Hartmann leads to a performance similar to a pyramid with a low modulation in photon noise, while keeping the practical advantages of the Shack-Hartmann.

2 Introduction

Adaptive optics rely on wavefront sensing to correct the effects of atmospheric turbulences. The choice of an appropriate wavefront sensor is therefore decisive for the system performance. Depending on the application, one has to compromise between linearity range, number of estimated modes, sensitivity to noise, cost, size... A lot of different setups and data processing methods can be used for each wavefront sensor in order to make this compromise, rendering the choice difficult.

When looking for high performance in closed loop, we can expect very small wavefronts, and sensitivity to noise becomes the limit of the wavefront sensor's accuracy. The sensitivity to noise is quantified by the covariance matrix of the wavefront estimation error, from which noise propagation coefficients can be deduced. This metric is commonly used to evaluate the performance of a wavefront sensor [1–8]. The error covariance matrix depends on the data and the type of estimator used. Two types of data can be considered : *raw data* and *pseudo-data*. Raw data correspond to the intensity distribution measured on the wavefront sensor's detector, while pseudo-data are the result of a first processing on the raw data. The classical approach is to assume that a *maximum likelihood estimation* is performed with a Gaussian noise model

on pseudo-data [1,2,4,5], if pseudo-data can be processed. The maximum likelihood estimator is unbiased, meaning that for a great number of noise occurrences, the mean of the estimations converges to the real value.

In order to get a fundamental limit of a wavefront sensor's performance, other studies rely on the *Cramer-Rao lower bound* [9–13] deduced from raw data. The Cramer-Rao lower bound defines a lower limit of the error covariance matrix ¹. In the case of an unbiased estimation, it is equal to the inverse of the *Fisher information matrix*, which quantifies the information available in the data. An estimator is called *efficient* if its error covariance matrix is equal to the Cramer-Rao lower bound. It has been shown that, in the assumption of a Gaussian noise on data, the maximum likelihood estimator is efficient [14]. Under this assumption, evaluating noise propagation with the Cramer-Rao lower bound or with the maximum likelihood estimator's error covariance matrix is thus equivalent and only the considered data have an influence on the result.

In this paper, we assume a Gaussian noise on data, but we choose to compute noise propagation from the Fisher information matrix. Although it is equivalent to evaluating the performance of a maximum likelihood estimation in this context, the Fisher information matrix corresponds to a more general concept and is therefore more attractive.

We apply this method to the comparison of the Shack-Hartmann and the pyramid. The Shack-Hartmann is well known and widely spread in adaptive optics. The pyramid was introduced in 1996 by Ragazzoni [15]. It has been shown that it offers a gain in limiting star magnitude over the Shack-Hartmann, as long as the modulation radius is lower than the radius of a subaperture spot in the Shack-Hartmann [4,5]. We would like to compare them for a fine wavefront estimation on a natural guide star. This is the typical goal of extreme adaptive optics systems, such as SAXO (SPHERE's Adaptive Optics for eXoplanet Observation) [16] on the Very Large Telescope or FLAO (First Light Adaptive Optics) on the Large Binocular Telescope. These systems respectively sample the wavefront on 40×40 subapertures and 30×30 subapertures. For our comparison, we use SAXO's sampling, 40×40 subapertures, in order to evaluate the noise propagation on orders as high as possible, but still relevant for current systems. We assume the sensing wavelength is the same as the imaging wavelength, so that the phase to sense is almost null.

Former comparisons of the Shack-Hartmann and the pyramid rely mostly on theoretical reasonings, and only take photon noise into account [3–5]. In this study, the noise propagations are computed from diffractive simulations of Shack-Hartmann and pyramid images. Furthermore, both photon noise and detector noise are considered.

We also add to the comparison a recently proposed derivation of the Shack-Hartmann, called the LIFTed Shack-Hartmann [17]. It uses LIFT [18,19] to extract more information from the pixels on the detector. This concept should offer a gain in performance over the classical Shack-Hartmann.

We first present in section 3 the hypotheses and method for computing noise propagation coefficients from the Fisher information matrix. Using this method, we compare the performances of the Shack-Hartmann and the pyramid in section 4. Finally, we add the LIFTed Shack-Hartmann to the comparison in section 5.

¹For two matrices A and B, $A \geq B$ means that $A - B$ is positive semidefinite.

3 Hypotheses and method

In this section, we present the method used to compute noise propagation in the different sensors.

We first describe our models for the simulation of the sensors (§3.1). We then show the link between the error covariance matrix of a maximum likelihood estimation and the Fisher information matrix (§3.2) for a classical wavefront estimation, and finally explain how to compute noise propagation coefficients from the Fisher information matrix (§3.3).

3.1 Wavefront sensors models

A wavefront sensor uses optical elements and/or a method to turn the wavefront deformations into an interpretable intensity distribution on a detector. The pixels values, which form the raw data, are degraded by photon noise and the detector’s read-out noise.

In this paper, we assume that the noise on raw data is a zero-mean additive Gaussian noise. We approximate the noise variance on each pixel by the addition of the mean flux on the pixel and the variance of read-out noise [20].

All the wavefront sensors considered here sample the pupil into subapertures. The number of subapertures is a key parameter which determines the number of estimated modes and noise propagation. As we said in the introduction, we consider here a pupil sampling on 40×40 subapertures for the Shack-hartmann and the pyramid.

We present hereafter the method and parameters used to simulate the different wavefront sensors. All simulations are made for a diffraction limited point source, which corresponds to the ideal case in an adaptive optics closed loop when observing a natural guide star.

3.1.1 Shack-Hartmann

Principle The Shack-Hartmann samples the incoming electromagnetic field thanks to a lenslet array placed in a pupil plane. Each lenslet focuses the part of the field it receives to create a sub-aperture image. The position of each subaperture spot is measured while the Shack-Hartmann is illuminated with a reference beam. A local tilt of the incoming wavefront changes the position of the spot. One can assume that there exists, in the absence of noise, a linear relation between the spot’s displacement with respect to the reference and the slope difference between the incoming wavefront and the reference wavefront in the subaperture. The incoming wavefront can therefore be estimated from the measurement of all spots’ displacements, which correspond to local slopes of the wavefront.

Simulation method The images of each subaperture are computed one by one, then gathered to form the global image in the focal plane of the lenslet array. The interferences normally produced by the lenslet array are neglected. To compute the Shack-Hartmann slopes, we use either a Center of Gravity (CoG) or a Weighted Center of Gravity (WCoG) corrected to have a unitary linearity coefficient [21]. CoG is more efficient in photon noise regime, while WCoG is more efficient in read-out noise regime.

Noise model The noise variance on slopes can be computed with Nicolle’s and Thomas’s theoretical formulas for the center of gravity and the weighted center of gravity [21, 22]:

$$\text{Center of gravity : } \sigma_{CoG_{ph}}^2 = \frac{\pi^2}{2 \ln(2) n_{ph}} \left(\frac{N_t}{N_d} \right)^2 \quad (1)$$

$$\sigma_{CoG_{ron}}^2 = \frac{\pi^2}{3} \frac{\sigma_e^2}{n_{ph}^2} \left(\frac{N_s}{N_d} \right)^2 \quad (2)$$

$$\text{Weighted center of gravity : } \sigma_{WC_{CoG_{ph}}}^2 = \frac{\pi^2}{2 \ln(2) n_{ph}} \left(\frac{N_t}{N_d} \right)^2 \frac{(N_t^2 + N_w^2)^4}{(2N_t^2 + N_w^2)^2 N_w^4} \quad (3)$$

$$\sigma_{WC_{CoG_{ron}}}^2 = \frac{\pi^3}{32(\ln 2)^2} \frac{\sigma_e^2}{n_{ph}^2} \frac{(N_t^2 + N_w^2)^4}{N_d^2 N_w^4} \quad (4)$$

$$(5)$$

with n_{ph} the flux and σ_e^2 the read out noise variance. The index ph stands for photon noise and ron for read out noise. N_t is the Full Width at Half Maximum (FWHM) of the subaperture spot, N_d the FWHM of the theoretical subaperture diffraction spot, N_s the number of pixels used to compute the center of gravity and N_w the FWHM of the weighting function in the weighted center of gravity. These variances are given for the measurement of a subaperture edge-to-edge phase difference and are expressed in radians².

Simulation parameters The simulations are made at Shannon sampling, with 16×16 pixels per subaperture. We assume we are at the diffraction limit, the noise variances on slopes are thus computed for $N_t = N_d = N_w = 2$ pixels. It has been shown that having $N_t = N_w$ is the optimal tuning for the weighted center of gravity [21]. We choose $N_s = 4$ pixels to be less sensitive to read-out noise, but still have enough pixels to estimate a center of gravity.

3.1.2 Pyramid

Principle The pyramid was inspired from Foucault's knife edge test. The light from the telescope focuses on the apex of a pyramidal prism. The beam is splitted into 4 and re-imaged in the pupil plane. The 4 resulting pupil images are similar to what one would obtain with a Foucault knife edge used in 2 perpendicular directions. Each pixel on the pupil images defines a subaperture.

In order to get a linear signal, a modulation can be performed by tilting the wavefront so that the image spot follows a circle around the pyramid's apex. By doing this modulation very fast compared to the exposure time, one can deduce the local wavefront slopes from the four pupil images, as in the Shack-Hartmann [15].

Simulation method The pyramid phase mask is applied to the complex amplitude in the focal plane, as in [4], creating a new complex amplitude from which the intensity in the detector plane is deduced. With this method, the image includes the interferences between the 4 beams leaving the pyramid. These interferences are stronger if the beams, and consequently the pupil images, are angularly close. In our case, the centers of the pupil images are separated by 2 pupil diameters. The pyramid is simulated for different radii of circular modulation. To simulate the modulation, we sample the circle into several points and an image is computed for each points.

The sum of all the resulting images forms the final image. To have a well-sampled modulation, the space between two points of integration is lesser or equal to the FWHM of the Airy pattern.

The pseudo-data of the pyramid are computed as follows for the pixel k :

$$\begin{aligned} S_x[k] &= \frac{(P_1[k] + P_3[k]) - (P_2[k] + P_4[k])}{N} \\ S_y[k] &= \frac{(P_1[k] + P_2[k]) - (P_3[k] + P_4[k])}{N} \\ \text{with } N &= \frac{1}{n_{subap}} \sum_k P_1[k] + P_2[k] + P_3[k] + P_4[k] \end{aligned} \quad (6)$$

with P_i the pupil image from face i and n_{subap} the number of subapertures. S_x is the signal linked to local slopes in the x direction and S_y is the signal linked to local slopes in the y direction. N is the flux per subaperture averaged over all the subapertures.

Noise model We consider there are enough subapertures to neglect the noise on N . The noise variance on S_x and S_y is thus equal to $1/N^2$ times the noise variance of their numerators.

In photon noise, the variance on the pixel $P_i[k]$ is equal to its mean flux. The numerator error variance is thus the sum of the mean fluxes of $P_1[k]$, $P_2[k]$, $P_3[k]$ and $P_4[k]$, which is equal to the mean flux of $P_1[k] + P_2[k] + P_3[k] + P_4[k]$.

In read-out noise, the variance on the pixel $P_i[k]$ is equal to σ_e^2 , with σ_e^2 the read-out noise variance. The numerator noise variance is thus $4\sigma_e^2$ in read-out noise.

Hence, the noise variance on $S_x[k]$ and $S_y[k]$ is $(P_1[k] + P_2[k] + P_3[k] + P_4[k])/N^2$ in photon noise and $4\sigma_e^2/N^2$ in read-out noise.

Simulation parameters We simulate the pyramid with and without modulation. The considered radii of modulation are $2 \lambda/D$, $3 \lambda/D$ and $6 \lambda/D$, λ being the sensing wavelength and D the pupil diameter. They correspond to modulations performed on the LBT [23].

3.1.3 LIFTed Shack-Hartmann

Principle The LIFTed Shack-Hartmann consists in using LIFT on the subapertures of a Shack-Hartmann. LIFT is a focal plane wavefront sensor which performs a maximum likelihood estimation of the phase on a single image, with a small-phase approximation [18, 19]. To remove the even modes indetermination, an astigmatism offset is added to the incoming phase. It is therefore possible to implement LIFT in a Shack-Hartmann by using astigmatic lenslets. A classical Shack-Hartmann measures 2 slopes per subaperture, whereas a LIFTed Shack-Hartmann can estimate more than 2 modes per subaperture. For the same number of estimated modes on the full aperture, a LIFTed Shack-Hartmann thus needs less subapertures and has a lower noise propagation.

Simulation method The simulation method is the same as for the Shack-Hartmann, but we add an astigmatism on each subaperture. The added astigmatism is the equivalent of the Zernike mode Z_5 , taken from a Zernike basis orthonormalized on a square subaperture.

In this paper, we do not use pseudo-data for the LIFTed Shack-Hartmann. We assume that the propagation derived from the raw data is equivalent to the propagation obtained with a well

tuned number of estimated modes per subaperture. In effect, by estimating several modes in each subaperture, we take most of the information contained in the pixels.

Noise model We use the raw data noise model described at the beginning of section 3.1.

Simulation parameters Since LIFT uses a small-phase approximation, its estimation becomes too degraded if the wavefront to estimate has a variance greater than 1 rad^2 (tip/tilt removed). Hence, the subapertures of the LIFTed Shack-Hartmann cannot be larger than a maximum diameter r_1 , for which the phase variance at the sensing wavelength is always lesser or equal than 1 rad^2 (tip/tilt removed). In the following, we assume this condition is verified.

It has been shown that a LIFTed Shack-Hartmann 10×10 estimating 5 modes per subaperture can control a deformable mirror with 21×21 actuators almost as well as a classical Shack-Hartmann 20×20 [17], for a pitch equal to the subaperture size of the classical Shack-Hartmann. We can thus assume that, by estimating a few more modes in each subaperture, the LIFTed Shack-Hartmann can control the same deformable mirror as a classical Shack-hartmann with subapertures twice as small. Of course, the mirror pitch has to be lesser or equal to $r_1/2$.

We thus chose to simulate a LIFTed Shack-Hartmann with 20×20 subapertures. The astigmatism offset on each subaperture is set to 0.5 rad rms .

3.1.4 General model

A wavefront sensor generally reconstructs a wavefront using a polynomial basis. We consider here a reconstruction on the Karhunen-Loève polynomials, which constitute the optimal basis for sensing turbulent wavefronts with a high spatial resolution [24], since they are statistically uncorrelated and they maximize the energy in low order modes. There are no analytical expressions of the Karhunen-Loeve polynomials. In this paper, we calculate them with a computationally effective method proposed by Cannon [25].

The vector of unknowns is $\mathbf{A} = \{a_2, a_3, \dots, a_n\}$, with a_i the coefficient of the i -th polynomial. We assume that the relation between the aberrations and the data is linear around the operating point. We can thus write the data formation model for any wavefront sensor:

$$\mathbf{y} = \mathbf{D}\mathbf{A} + \mathbf{n} \quad (7)$$

with \mathbf{y} the vector of data, \mathbf{D} the interaction matrix and \mathbf{n} the noise.

Wavefront sensors have similar behaviours in terms of photon noise and detector noise [1]. For any wavefront sensor and estimator, we can write the variance of estimation as:

$$\sigma^2(\hat{\mathbf{A}} - \mathbf{A}) = \overbrace{\left(\sum_i \alpha_i\right)}^{\alpha} \frac{1}{n_{ph}} + \overbrace{\left(\sum_i \beta_i\right)}^{\beta} \left(\frac{\sigma_e}{n_{ph}}\right)^2 \quad (8)$$

with $\hat{\mathbf{A}}$ the vector of estimated coefficients, n_{ph} the incoming flux in photo-electrons and σ_e the standard deviation of the read-out noise. α_i and β_i are the noise propagation coefficients on the i -th mode for photon noise and read-out noise respectively.

In the following, we focus on the evaluation of noise propagation in wavefront sensors. We first relate the Fisher information matrix to the classical approach based on the maximum likelihood estimation and then show how the coefficients α_i and β_i are computed from the Fisher information matrix. The considered operating point is a null phase, i. e. $\mathbf{A} = \mathbf{0}$.

3.2 Maximum likelihood estimation and Fisher information matrix

The Fisher information matrix is often seen as a complicated mathematical object which cannot be easily related to physical concepts. The goal of this paragraph is to link the Fisher information matrix with a more familiar figure of merit: the covariance matrix of estimation error for a maximum likelihood estimator.

3.2.1 Noise propagation in a maximum likelihood estimation

The solution given by the maximum likelihood estimator for the linear model described by Eq. 7 is:

$$\hat{\mathbf{A}} = (\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{y} \quad (9)$$

with $\mathbf{C}_n \triangleq \langle \mathbf{n} \mathbf{n}^t \rangle$ the noise covariance matrix. The estimation error is then:

$$\mathbf{E} \triangleq \hat{\mathbf{A}} - \mathbf{A} = (\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D})^{-1} \mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{n} \quad (10)$$

and the covariance matrix of estimation error is:

$$\langle \mathbf{E} \mathbf{E}^t \rangle = (\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D})^{-1} \quad (11)$$

The variance of estimation error for each mode is given by the diagonal elements of this matrix.

For Shack-Hartmann slopes, assuming the noise is homogeneous and uncorrelated from one slope to another, \mathbf{C}_n is diagonal. The noise propagation coefficients on each mode are then proportionnal to the diagonal elements of $(\mathbf{D}^t \mathbf{D})^{-1}$. Rigaut and Gendron used this result to find an analytical formulation of the noise propagation in the Shack-Hartmann [2]. In the asymptotic case of an infinite number of subapertures, they demonstrated that the noise propagation coefficient for each mode was proportionnal to $(n + 1)^{-2}$, with n the radial order of the considered mode. This result is typical for slope sensors [1].

3.2.2 Link with the Fisher information matrix

The Fisher information matrix can be expressed as a function of the derivatives of the data with respect to each unknown. One can show that the elements of the Fisher information matrix \mathbf{F} are (see appendix A):

$$\mathbf{F}_{i,j} = \sum_k \frac{1}{\sigma_k^2} \frac{\partial y_k}{\partial a_i} \frac{\partial y_k}{\partial a_j} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial y_l}{\partial a_i} \frac{\partial y_k}{\partial a_j} \quad (12)$$

with σ_k^2 the error variance on the k -th data element and σ_{kl}^2 the error covariance between the data elements k and l . This equation is valid under the assumption of a Gaussian noise. Let us see if we can compare $\langle \mathbf{E} \mathbf{E}^t \rangle = (\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D})^{-1}$ and \mathbf{F} .

We first need to find the expression of the interaction matrix \mathbf{D} . From Eq. 7, we can write:

$$\mathbf{y}(\mathbf{A} + \delta \mathbf{A}) = \mathbf{D} \mathbf{A} + \mathbf{D} \delta \mathbf{A} + \mathbf{n} = \mathbf{y}(\mathbf{A}) + \mathbf{D} \delta \mathbf{A} \quad (13)$$

with $\delta \mathbf{A}$ a small perturbation of wavefront coefficients. We thus have:

$$\mathbf{D} = \frac{\mathbf{y}(\mathbf{A} + \delta \mathbf{A}) - \mathbf{y}(\mathbf{A})}{\delta \mathbf{A}} \triangleq \frac{\partial \mathbf{y}}{\partial \mathbf{A}} \quad (14)$$

For m modes and p data elements, the expression of \mathbf{D} is then:

$$\mathbf{D} = \begin{pmatrix} \frac{\partial y_1}{\partial a_1} & \frac{\partial y_1}{\partial a_2} & \cdots & \frac{\partial y_1}{\partial a_m} \\ \frac{\partial y_2}{\partial a_1} & \frac{\partial y_2}{\partial a_2} & \cdots & \frac{\partial y_2}{\partial a_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_p}{\partial a_1} & \frac{\partial y_p}{\partial a_2} & \cdots & \frac{\partial y_p}{\partial a_m} \end{pmatrix} \quad (15)$$

Knowing that the noise covariance matrix is:

$$\mathbf{C}_n = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{p1}^2 & \cdots & \cdots & \sigma_p^2 \end{pmatrix} \quad (16)$$

We find:

$$[\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D}]_{i,j} = \sum_k \frac{1}{\sigma_k^2} \frac{\partial y_k}{\partial a_i} \frac{\partial y_k}{\partial a_j} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial y_l}{\partial a_i} \frac{\partial y_k}{\partial a_j} \quad (17)$$

Hence $\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D} = \mathbf{F}$, meaning that $\langle \mathbf{E} \mathbf{E}^t \rangle = \mathbf{F}^{-1}$. The covariance matrix of error estimation of the maximum likelihood estimator is thus equal to the Cramer-Rao lower bound, as expected for a Gaussian noise model. In the case of the Shack-Hartmann, this means that the Cramer-Rao lower bound for pseudo-data (i. e. local slopes) can be directly deduced from Rigaut-Gendron coefficients.

However, the Fisher information matrix is still a more general concept than the covariance matrix of the maximum likelihood estimator, and seems more fair for comparisons. For this reason, we compute noise propagations from the Fisher information matrix in the following.

3.3 Computing the noise propagation coefficients for a given wavefront sensor

Assuming the noise is uncorrelated from one data element to another, the elements of the Fisher information matrix are:

$$F_{i,j}(\mathbf{A}_p) = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial y_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial y_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \quad (18)$$

As we have seen in section 3.2, the inverse of the Fisher information matrix is equivalent to an error covariance matrix. Its diagonal elements are the error variances for each mode. We can thus deduce noise propagation coefficients for each sensor from Eq. 8.

When computing the Fisher information matrix with $n_{ph} = 1$ photo-electron and $\sigma_e = 0$ electron, we directly get $\alpha_i = \mathbf{F}^{-1}[i, i]$. Similarly, with $n_{ph} = 1$ photo-electron and $\sigma_e = 1$ electron, without photon noise, we have $\beta_i = \mathbf{F}^{-1}[i, i]$. The coefficients calculated this way verify Eq. 8 in good approximation.

In the next section, we compute the noise propagation coefficients of the Shack-Hartmann and the pyramid.

4 Comparison of noise propagation in the Shack-Hartmann and pyramid sensors

The modal noise propagation coefficients α_i and β_i , for the estimation of 1000 Karhunen-Loève polynomials in monochromatic light, are plotted in Fig. 1. The data considered are the pseudo-data. The change in these propagations with polychromatic light is small, at least in photon noise (see appendix B).

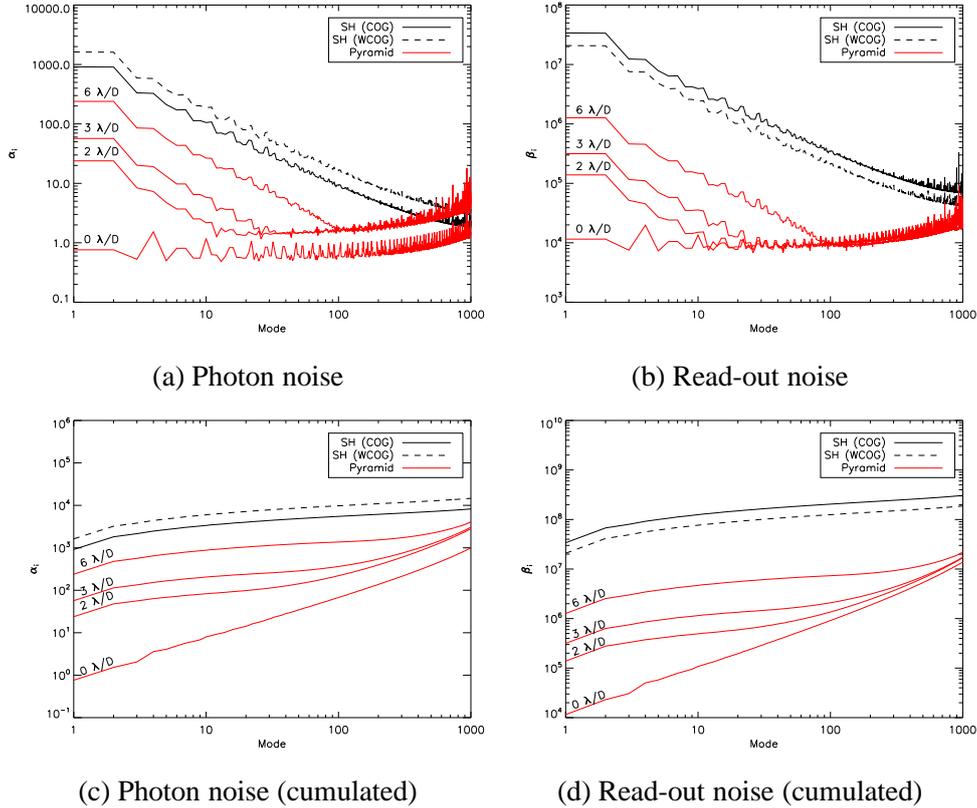


Figure 1: Noise propagation coefficients of the Shack-Hartmann and the pyramid in monochromatic light for the estimation of 1000 Karhunen-Loève modes, with pseudo-data. The pyramid propagation goes up in low-order modes as the modulation radius gets larger. The modulation radii are $0 \lambda/D$, $2 \lambda/D$, $3 \lambda/D$ and $6 \lambda/D$.

The Shack-Hartmann has a propagation in j^{-1} with j the polynomial number. Since $j \sim (n + 1)^2$, n being the radial order, this is equivalent to the propagation found by Rigaut and Gendron [2]. As expected, the best estimator for local slopes is the center of gravity in photon noise and the weighted center of gravity in read-out noise.

In photon noise, the fixed pyramid has a flat propagation which goes up for high frequencies because of the filtering effect of the subaperture size, as shown by V erinaud [4]. Also, the modulation makes the sensor act as a slope sensor in low orders and the propagation follows the same trend as the propagation of the Shack-Hartmann. The associated loss of performance in low orders is proportionnal to the square of the modulation radius. All this is consistent with V erinaud's theoretical analysis. However, the gap between the pyramid with and without modulation in high orders does not appear in his study. This difference may come from V erinaud's

assumption of a uniform noise over the pupil images. When the pyramid is not modulated, there is an important contrast between the flux received by the subapertures in the center and the flux received by the subapertures on the edge. The assumption of uniform noise is thus false for photon noise. Read-out noise is always uniform whatever the regime. This is why the gap disappears in the read-out noise propagation.

The pyramid is clearly more performant than the Shack-Hartmann in both noise regimes. In order to quantify the gain given by the pyramid over the Shack-Hartmann for the estimation of 1000 Karhunen-Loève, we compute the sum of all modal coefficients for each noise (Tab. 1).

	$\alpha = \sum \alpha_i$	$\beta = \sum \beta_i$
Shack-Hartmann (CoG)	8242	3.06×10^8
Shack-Hartmann (WCoG)	14652	1.88×10^8
Pyramid ($0\lambda/D$)	991	1.38×10^7
Pyramid ($2\lambda/D$)	2808	1.70×10^7
Pyramid ($3\lambda/D$)	3058	1.75×10^7
Pyramid ($6\lambda/D$)	4091	2.21×10^7

Table 1: Global noise propagation coefficients for the estimation of 1000 Karhunen-Loève by a Shack-Hartmann or a pyramid. α is the photon noise coefficient and β the read-out noise coefficient.

In photon noise, the gain in performance given by the pyramid goes from 2 (for a modulation of $6\lambda/D$) to more than 8 (no modulation) compared to a Shack-Hartmann using center of gravity. This means a gain up to 2.25 in limiting star magnitude in a photon noise limited case. Vérinaud and Ragazzoni have already found a similar gain [4, 5]. Besides, in read-out noise, the gain reaches approximately 10 for all considered modulations compared to a Shack-Hartmann using weighted center of gravity. Hence, the pyramid is much more performant than the Shack-Hartmann whatever the noise conditions.

A better performance could be reached with the Shack-Hartmann by using raw data [26]. However, the computational complexity of the estimation is not compatible with an adaptive optics application. In order to get closer to the performance of a Shack-Hartmann using raw data, we recently proposed a variation, the LIFTed Shack-Hartmann, which extracts more information from the pixels. Its principle is described in section 3.1. It has the potential to give a performance similar to a modulated pyramid, while keeping the advantages of the Shack-Hartmann: simple setup, easy to debug, real-time estimation. . .

In the next section, we compare the noise propagation of the Shack-Hartmann and the pyramid to the LIFTed Shack-Hartmann.

5 Comparison of the Shack-Hartmann and the pyramid with the LIFTed Shack-Hartmann

We verify here our assumptions on the LIFTed Shack-Hartmann’s performance. We plot in Fig. 2 the noise propagation coefficients of this sensor, as well as the classical Shack-Hartmann and the pyramid, in monochromatic light. The LIFTed Shack-Hartmann relies on LIFT, which

is still efficient in polychromatic light [19]. We can thus expect a negligible change of the LIFTed Shack-Hartmann's noise propagation in polychromatic light.

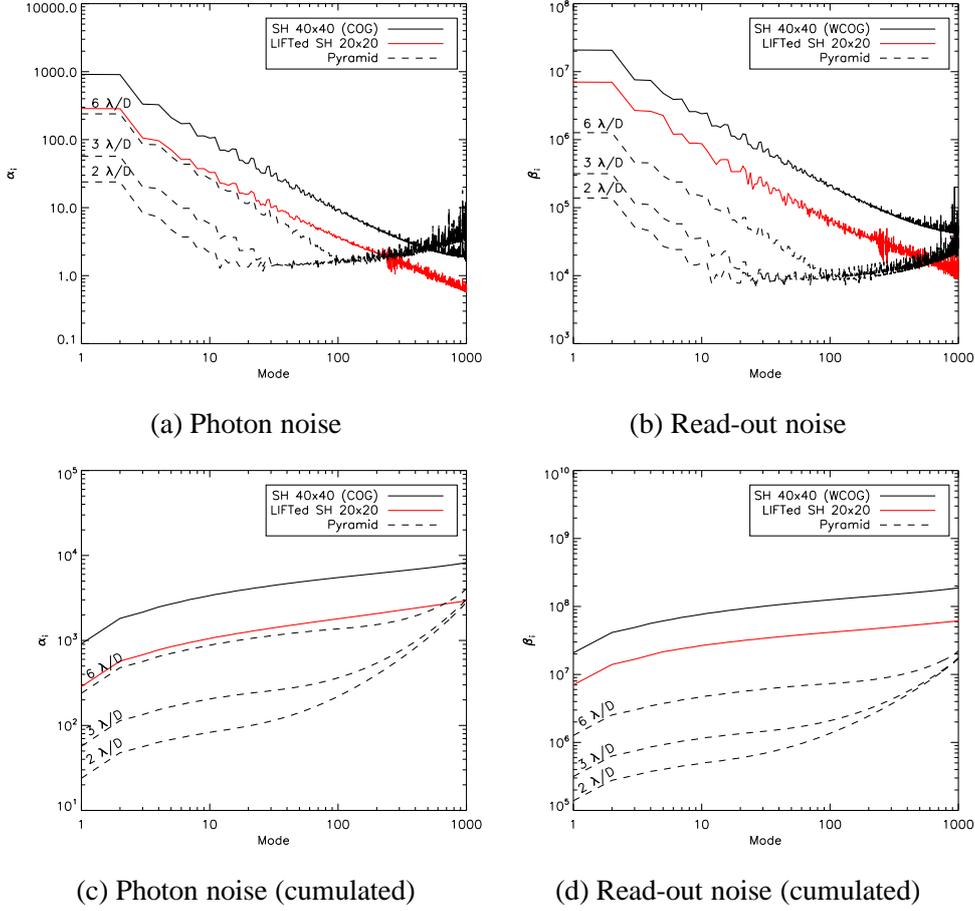


Figure 2: Noise propagation coefficients of the classical Shack-Hartmann, the LIFTed Shack-Hartmann and the pyramid in monochromatic light for the estimation of 1000 Karhunen-Loeve modes. The pyramid propagation goes up in low frequencies as the modulation radius gets larger. The modulation radii are $2 \lambda/D$, $3 \lambda/D$ and $6 \lambda/D$.

The LIFTed Shack-Hartmann, as expected, gives a better performance than the classical Shack-Hartmann. Its performance is similar to a pyramid with a $6 \lambda/D$ modulation in low orders, and outperforms the modulated pyramid in high orders. Its global noise coefficients are $\alpha = 2922$ in photon noise and $\beta = 6.14 \times 10^7$ in read-out noise. The gain over the Shack-Hartmann is thus approximately 3 in both noise regime. In photon noise, it is as performant as a pyramid with a modulation between $2 \lambda/D$ and $3 \lambda/D$. However, its noise propagation in read-out noise is approximately 3 times as high as the pyramid's.

In conclusion, the LIFTed Shack-Hartmann is an important improvement the classical Shack-Hartmann. Though, to be competitive with the pyramid, it has to be used in photon noise limited applications.

6 Conclusion

We have used a method based on the Fisher information matrix to compare wavefront sensors, and related it to the classical approach which relies on a maximum likelihood estimation. It allows a fair comparison as it evaluates directly the information available in the data.

We have applied this method to evaluate the noise propagations of the Shack-Hartmann and pyramid sensors for the estimation of 1000 karhunen-Loève polynomials on 40×40 subapertures, in photon noise and read-out noise. Our study relies on diffractive simulations, and less approximations than previous comparisons. It confirms most of the known results, such as the noise propagation trends, or the ~ 2 magnitudes gain given by the non-modulated pyramid over the Shack-Hartmann in a photon noise limited case. Besides, we have demonstrated that the pyramid also offers an important gain in performance (~ 10) in read-out noise regime for all considered modulation radii (from $2 \lambda/D$ to $6 \lambda/D$). Therefore, whatever the noise regime, an important gain in limiting magnitude gain can be expected when using a pyramid rather than a Shack-hartmann.

Finally, we have extended the comparison to a concept we proposed recently, the LIFTed Shack-Hartmann, which is able to extract more information from the pixels than a classical Shack-Hartmann, without a significative loss of computational time. By using subapertures twice as large (20×20 configuration), it can lower the propagation of the Shack-Hartmann by a factor ~ 3 in both noise regimes. Its global performance would be comparable to a pyramid with a low modulation ($2-3 \lambda/D$) in photon noise limited applications.

A Fisher information matrix of data with additive gaussian noise

Let $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ be a set of data, depending on a set of wavefront mode coefficients $\mathbf{A} = \{a_1, a_2, \dots, a_m\}$ and a noise $\mathbf{n} = \{n_1, n_2, \dots, n_n\}$. The expression of the Fisher matrix is:

$$F_{i,j}(\mathbf{A}_p) = E\left\{\left[\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p} \left[\frac{\partial}{\partial a_j} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p}\right\} \quad (19)$$

with $p(\mathbf{y}|\mathbf{A})$ the likelihood probability density function and \mathbf{A}_p the operating point.

For an additive Gaussian noise on data, the likelihood probability density function is:

$$p(\mathbf{y}|\mathbf{A}) = p_b(\mathbf{y}) = \prod_{k=1}^N \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{[y_k - \bar{y}_k(\mathbf{A})]^2}{2\sigma_k^2}\right\} \quad (20)$$

To find $\mathbf{F}(\mathbf{A}_p)$, we need to compute $\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p}$. We can write:

$$\ln p(\mathbf{y}|\mathbf{A}) = \sum_{k=1}^N -\frac{1}{2} \ln(2\pi\sigma_k^2) - \frac{[y_k - \bar{y}_k(\mathbf{A})]^2}{2\sigma_k^2} \quad (21)$$

$$\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})] \quad (22)$$

Hence:

$$\begin{aligned} \frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} \times \frac{\partial}{\partial a_j} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} &= \sum_{k=1}^N \frac{1}{\sigma_k^4} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})]^2 \\ &+ \sum_{l \neq k} \frac{1}{\sigma_l^2 \sigma_k^2} \frac{\partial \bar{y}_l(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})] \times [y_l - \bar{y}_l(\mathbf{A})] \end{aligned} \quad (23)$$

Knowing that $E\{[y_k - \bar{y}_k(\mathbf{A})]^2\} = \sigma_k^2$ and $E\{[y_k - \bar{y}_k(\mathbf{A})][y_l - \bar{y}_l(\mathbf{A})]\} = \sigma_{kl}^2$, we finally have:

$$F_{i,j}(\mathbf{A}_p) = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial \bar{y}_l(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \quad (24)$$

If we consider that the noise is uncorrelated from one data element y_i to another, the expression is simplified into:

$$F_{i,j}(\mathbf{A}_p) = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \quad (25)$$

B Noise propagation with polychromatic data

We evaluate here the consequences of having polychromatic data on the photon noise propagations of the Shack-Hartmann and the pyramid.

B.1 Variance of slope estimation error due to photon noise in the Shack-Hartmann

The estimation error variances of local slopes seen in Eq. 5 are expressed for the measurement of an edge-to-edge phase difference $\Delta\varphi$ in radians. However, to be able to apply a correction with a deformable mirror, the wavefront has to be estimated in optical path difference. We assume that the aberrations are achromatic in optical path difference. For an estimation at wavelength λ , the slope measurement becomes:

$$\Delta\delta_\lambda = \frac{\lambda}{2\pi} \Delta\varphi \quad (26)$$

with $\Delta\delta$ the edge-to-edge optical path difference.

The slope estimation in polychromatic light is the mean of the slopes over the spectral bandwidth $\Delta\lambda$:

$$\Delta\delta_{\Delta\lambda} = \frac{1}{n_{ph}} \int_\lambda \gamma_\lambda \Delta\delta_\lambda d\lambda \quad (27)$$

with n_{ph} the total flux and γ_λ the flux at wavelength λ . The error variance is then:

$$\sigma_{\Delta\lambda}^2 = \frac{1}{n_{ph}^2} \int_\lambda \gamma_\lambda^2 \sigma_\lambda^2 d\lambda \quad (28)$$

Let us take the central wavelength λ_c as a reference for the measurements. We have:

$$\Delta\delta_\lambda = \frac{\lambda}{2\pi} \frac{2\pi}{\lambda_c} \Delta\delta_{\lambda_c} = \frac{\lambda}{\lambda_c} \Delta\delta_{\lambda_c} \quad (29)$$

The error variance on a slope is dependant on the inverse of the flux in photon noise (see Eq. 5). We thus have:

$$\sigma_{\lambda, ph}^2 = \frac{\gamma_{\lambda_c}}{\gamma_\lambda} \left(\frac{\lambda}{\lambda_c} \right)^2 \sigma_{\lambda_c, ph}^2 \quad (30)$$

The error of slope estimation becomes:

$$\sigma_{\Delta\lambda, ph}^2 = \frac{1}{n_{ph}^2} \int_\lambda \gamma_\lambda^2 \frac{\gamma_{\lambda_c}}{\gamma_\lambda} \left(\frac{\lambda}{\lambda_c} \right)^2 \sigma_{\lambda_c, ph}^2 d\lambda = \frac{\gamma_{\lambda_c}}{n_{ph}^2} \sigma_{\lambda_c, ph}^2 \int_\lambda \gamma_\lambda \left(\frac{\lambda}{\lambda_c} \right)^2 d\lambda \quad (31)$$

If we consider a uniform flux over the spectral band, we have $\gamma_\lambda = n_{ph}/\Delta\lambda$. In this case, the variance is:

$$\sigma_{\Delta\lambda, ph}^2 = \frac{\gamma_{\lambda_c}}{n_{ph}} \frac{\sigma_{\lambda_c, ph}^2}{\Delta\lambda} \int_\lambda \left(\frac{\lambda}{\lambda_c} \right)^2 d\lambda \quad (32)$$

The integral over the the squared wavelengths is equal to:

$$\int_{\lambda_c - \Delta\lambda/2}^{\lambda_c + \Delta\lambda/2} \left(\frac{\lambda}{\lambda_c} \right)^2 d\lambda = \Delta\lambda + \frac{\Delta\lambda}{12R^2} \quad (33)$$

with $R = \lambda_c/\Delta\lambda$ the spectral resolution.

Finally, we have:

$$\sigma_{\Delta\lambda, ph}^2 = \frac{\gamma_{\lambda_c}}{n_{ph}} \sigma_{\lambda_c, ph}^2 \left(1 + \frac{1}{12R^2} \right) \quad (34)$$

In monochromatic light, all the flux is in the central wavelength, so that $\gamma_{\lambda_c} = n_{ph}$. Finally, we get:

$$\sigma_{\Delta\lambda, ph}^2 = \left(1 + \frac{1}{12R^2}\right)\sigma_{mono, ph}^2 \quad (35)$$

For $R \leq 2$, which is the case of most filters, the error variance is increased by less than 2% in polychromatic light. We can thus consider that there is no change in the propagation of photon noise.

B.2 Variance of slope estimation error due to photon noise in the pyramid

The notations are the same as precedently. The pyramid signal S_x in polychromatic light for the pixel k is equal to (see Eq. 6):

$$S_{x_{\Delta\lambda}}[k] = \frac{\sum_{\lambda}(P_{1\lambda}[k] + P_{3\lambda}[k]) - (P_{2\lambda}[k] + P_{4\lambda}[k])}{N_{\Delta\lambda}} \quad (36)$$

with $N_{\Delta\lambda} = \sum_{\lambda} N_{\lambda}$. Hence:

$$N_{\Delta\lambda}S_{x_{\Delta\lambda}}[k] = \sum_{\lambda} N_{\lambda}S_{x_{\lambda}} \quad (37)$$

Knowing that $N_{\lambda}/N_{\Delta\lambda} = \gamma_{\lambda}/n_{ph}$, the signal in polychromatic is thus the mean of signals over the spectral band:

$$S_{x_{\Delta\lambda}}[k] = \frac{1}{n_{ph}} \sum_{\lambda} \gamma_{\lambda}S_{x_{\lambda}} \quad (38)$$

The signal is related to the wavefront slopes by the equation [4]:

$$S_x = \frac{\lambda}{r_m\pi^2} \frac{\partial\varphi}{\partial x} \quad (39)$$

with r_m the modulation radius. We can deduce:

$$S_{x_{\lambda}} = \frac{\lambda}{\lambda_c} S_{x_{\lambda_c}} \quad (40)$$

The error variance on the signal depends on the inverse of the flux in photon noise, as in the Shack-Hartmann. We thus have:

$$\sigma_{\lambda, ph}^2 = \frac{\gamma_{\lambda_c}}{\gamma_{\lambda}} \left(\frac{\lambda}{\lambda_c}\right)^2 \sigma_{\lambda_c, ph}^2 \quad (41)$$

We retrieve a situation similar to the previous paragraph, and we finally get:

$$\sigma_{\Delta\lambda, ph}^2 = \left(1 + \frac{1}{12R^2}\right)\sigma_{mono, ph}^2 \quad (42)$$

As for the Shack-Hartmann, we can assume that the change in photon noise propagation due to polychromatic light is negligible if $R \leq 2$.

References

- [1] G. Rousset. Wave-front sensors. In F. Roddier, editor, *Adaptive Optics in Astronomy*, chapter 5, pages 91–130. 1999.
- [2] F Rigaut and E Gendron. Laser guide star in adaptive optics-the tilt determination problem. *Astronomy and Astrophysics*, 261:677–684, 1992.
- [3] Olivier Guyon. Limits of adaptive optics for high-contrast imaging. *The Astrophysical Journal*, 629(1):592, 2005.
- [4] Christophe Vérinaud. On the nature of the measurements provided by a pyramid wave-front sensor. *Optics Communications*, 233(1):27–38, 2004.
- [5] Roberto Ragazzoni and J Farinato. Sensitivity of a pyramidal wave front sensor in closed loop adaptive optics. *Astronomy and Astrophysics*, 350:L23–L26, 1999.
- [6] Brent L Ellerbroek, Brian J Thelen, David J Lee, David A Carrara, and Richard G Paxman. Comparison of shack-hartmann wavefront sensing and phase-diverse phase retrieval. In *Optical Science, Engineering and Instrumentation'97*, pages 307–320. International Society for Optics and Photonics, 1997.
- [7] Byron M Welsh, Brent L Ellerbroek, Michael C Roggemann, and Timothy L Pennington. Fundamental performance comparison of a hartmann and a shearing interferometer wave-front sensor. *Applied optics*, 34(21):4186–4195, 1995.
- [8] Ludovic Meynadier, Vincent Michau, Marie-Thérèse Velluet, Jean-Marc Conan, Laurent M Mugnier, and Gérard Rousset. Noise propagation in wave-front sensing with phase diversity. *Applied optics*, 38(23):4967–4979, 1999.
- [9] James R Fienup, Brian J Thelen, Richard G Paxman, and David A Carrara. Comparison of phase diversity and curvature wavefront sensing. In *Astronomical Telescopes & Instrumentation*, pages 930–940. International Society for Optics and Photonics, 1998.
- [10] David J Lee, Byron M Welsh, and Michael C Roggemann. Cramer-rao analysis of phase diversity imaging. In *Optical Science, Engineering and Instrumentation'97*, pages 161–172. International Society for Optics and Photonics, 1997.
- [11] Timothy J Schulz, Wei Sun, and Michael C Roggemann. Cramer-rao bounds for estimation of turbulence-induced wavefront aberrations. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 23–28. International Society for Optics and Photonics, 1999.
- [12] C Paterson. Towards practical wavefront sensing at the fundamental information limit. In *Journal of Physics: Conference Series*, volume 139, page 012021. IOP Publishing, 2008.
- [13] Shane Barwick. Performance comparison between shack-hartmann and astigmatic hybrid wavefront sensors. *Applied optics*, 48(36):6967–6972, 2009.

- [14] Harrison H Barrett, JL Denny, Robert F Wagner, and Kyle J Myers. Objective assessment of image quality. ii. fisher information, fourier crosstalk, and figures of merit for task performance. *JOSA A*, 12(5):834–852, 1995.
- [15] Roberto Ragazzoni. Pupil plane wavefront sensing with an oscillating prism. *Journal of modern optics*, 43(2):289–293, 1996.
- [16] J-F Sauvage, T Fusco, C Petit, S Meimon, E Fedrigo, M Suarez Valles, M Kasper, N Hubin, J-L Beuzit, J Charton, et al. Saxo, the extreme adaptive optics system of sphere: overview and calibration procedure. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 77360F–77360F. International Society for Optics and Photonics, 2010.
- [17] Serge Meimon, Thierry Fusco, Vincent Michau, and Cédric Plantet. Sensing more modes with fewer sub-apertures: the lifted shack–hartmann wavefront sensor. *Optics Letters*, 39(10):2835–2837, 2014.
- [18] Serge Meimon, Thierry Fusco, and Laurent M Mugnier. Lift: a focal-plane wavefront sensor for real-time low-order sensing on faint sources. *Optics letters*, 35(18):3036–3038, 2010.
- [19] C Plantet, S Meimon, J-M Conan, and T Fusco. Experimental validation of lift for estimation of low-order modes in low-flux wavefront sensing. *Optics Express*, 21(14):16337–16352, 2013.
- [20] Laurent M. Mugnier, Thierry Fusco, and Jean-Marc Conan. Mistral: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *J. Opt. Soc. Am. A*, 21(10):1841–1854, Oct 2004.
- [21] S Thomas, T Fusco, A Tokovinin, M Nicolle, V Michau, and G Rousset. Comparison of centroid computation algorithms in a shack–hartmann sensor. *Monthly Notices of the Royal Astronomical Society*, 371(1):323–336, 2006.
- [22] M Nicolle, T Fusco, G Rousset, and V Michau. Improvement of shack-hartmann wavefront sensor measurement for extreme adaptive optics. *Optics letters*, 29(23):2743–2745, 2004.
- [23] Simone Esposito, Armando Riccardi, Enrico Pinna, Alfio Puglisi, Fernando Quirós-Pacheco, Carmelo Arcidiacono, Marco Xompero, Runa Briguglio, Guido Agapito, Lorenzo Busoni, et al. Large binocular telescope adaptive optics system: new achievements and perspectives in adaptive optics. In *SPIE Optical Engineering+ Applications*, pages 814902–814902. International Society for Optics and Photonics, 2011.
- [24] RG Lane and M Tallon. Wave-front reconstruction using a shackhartmann sensor. *Applied optics*, 31(32):6902–6908, 1992.
- [25] Robert C Cannon. Optimal bases for wave-front simulation and reconstruction on annular apertures. *JOSA A*, 13(4):862–867, 1996.
- [26] Robert C Cannon. Global wave-front reconstruction using shack–hartmann sensors. *JOSA A*, 12(9):2031–2039, 1995.

5.4.3 Bilan et discussion

Bilan Dans cet article, nous avons comparé les propagations du bruit dans le Shack-Hartmann, la pyramide et le Shack-Hartmann LIFTé à l'aide de l'information de Fisher. Nous avons considéré leur application dans un système d'optique adaptative extrême, avec 40×40 sous-pupilles pour le Shack-Hartmann classique et la pyramide, et 20×20 sous-pupilles pour le Shack-Hartmann LIFTé.

Dans un premier temps, nous nous sommes intéressés au Shack-Hartmann classique et à la pyramide. Nous avons calculé leurs coefficients de propagation du bruit pour l'estimation de 1000 modes de Karhunen-Loève. Nous avons ainsi pu vérifier la majorité des résultats de la bibliographie, et quantifier le gain apporté par la pyramide par rapport au Shack-Hartmann pour l'estimation de ces modes et pour les deux types de bruit. Le gain est entre 2 (rayon de modulation = $6 \lambda/D$) et 8 (pyramide non modulée) en bruit de photons et de l'ordre de 10 en bruit de détecteur. Quel que soit le régime de bruit, la pyramide est donc bien plus performante que le Shack-Hartmann.

Nous avons ensuite ajouté le Shack-Hartmann LIFTé dans la comparaison. Nous avons ainsi montré qu'en utilisant un Shack-Hartmann LIFTé avec deux fois moins de sous-pupilles que le Shack-Hartmann classique, la propagation du bruit est divisée par 3 environ pour les deux bruits, avec un même nombre de modes estimés sur la pupille globale. Ce résultat est cohérent avec les résultats du chapitre 4 p.133). Pour des applications limitées par le bruit de photon, la performance globale atteinte est alors similaire à celle d'une pyramide avec une faible modulation ($2-3 \lambda/D$). Cependant, en bruit de détecteur, le Shack-Hartmann LIFTé est 3 fois moins efficace que la pyramide pour ces modulations.

Discussion De même que dans le chapitre 4 p.133, l'amplitude d'astigmatisme utilisée dans l'article (0.5 rad) pour le Shack-Hartmann LIFTé n'est a priori pas optimale pour le cas considéré. En effet, cette amplitude est optimisée pour l'estimation de 3 modes dans la sous-pupille (cf §2.3 p.70), et une division du nombre de sous-pupilles par 2 nécessite l'estimation d'environ 8 modes par sous-pupille (cf §4.5.1 p.144). Il est donc sûrement possible d'obtenir une propagation du Shack-Hartmann LIFTé un peu meilleure, notamment en bruit de détecteur.

Dans la suite, nous ajoutons quelques éléments de comparaison et nous l'étendons à d'autres analyseurs.

5.5 Autres comparaisons : premiers résultats et perspectives

Dans cette partie, je présente quelques résultats complémentaires et j'étends la comparaison à plus d'analyseurs. Je décris tout d'abord les paramètres de simulation utilisés pour chaque analyseur. Leurs principes, ainsi que les méthodes pour les simuler, ont été présentés dans la partie 1.4 p.46. Je montre ensuite les coefficients de propagation obtenus en lumière monochromatique, puis je donne des pistes pour une comparaison en polychromatique. Le calcul de la matrice d'information de Fisher pour chacun des analyseurs est illustré dans l'annexe E p.199.

5.5.1 Paramètres de simulation

Je présente ici les paramètres de simulation pour chaque analyseur. Tous les analyseurs qui échantillonnent la pupille sont simulés avec 40×40 sous-pupilles, afin d'être en accord avec le

dimensionnement du Shack-Hartmann décrit au paragraphe 5.3 p.152. Le modèle de bruit pour les données brutes sera toujours celui d'une image classique (cf §1.2.1.1 p.22).

Pyramide Les coefficients de propagation sont calculés pour des rayons de modulation allant de 0 à 40 rayons de tache d'Airy.

Diversité de phase L'échantillonnage est à Shannon. On simule les images pour des amplitudes de diversité et des répartitions de flux diverses.

Analyseur de courbure Toutes les sous-pupilles du détecteur ont une taille égale¹. Pour choisir les autres paramètres, nous nous sommes inspirés du système AO188 du télescope Subaru [Guyon *et al.*, 2008]. La pupille a un diamètre de 8 m, et la distance focale est de 528 m. La distance extra-focale l vaut entre 0,2 et 0,5 m sur ce système. Pour avoir une propagation de bruit optimale, nous avons choisi la distance la plus proche du foyer, 0,2 m, en accord avec l'analyse de Fienup *et al.* [Fienup *et al.*, 1998]. Enfin, la lumière est à une longueur d'onde de 700 nm.

Masque de Zernike Il n'y a pas de paramètre particulier à ajuster.

5.5.2 Propagation du bruit en lumière monochromatique

Je présente maintenant les coefficients de propagation de bruit calculés pour l'estimation de 1000 modes de Karhunen-Loève sur source ponctuelle en lumière monochromatique.

5.5.2.1 Pyramide

Nous comparons ici les coefficients de propagation du bruit mode à mode pour une estimation à partir des données brutes et pour une estimation à partir des pseudo-données (Fig. 5.3).

Lorsque la pyramide est modulée, l'estimation par les pseudo-données exploite parfaitement les informations de l'image totale. Pour la pyramide non modulée, les pseudo-données donnent une performance légèrement plus faible. En effet, ces pseudo-données ont été définies pour retrouver les pentes de la phase [Ragazzoni, 1996] lorsque la pyramide est modulée. La pyramide non modulée étant un senseur de phase, ce type de données n'est a priori pas adapté. La différence est cependant négligeable, et on peut considérer que l'information est bien conservée par les pseudo-données dans tous les cas.

5.5.2.2 Diversité de phase

Cas classique La figure 5.4 montre les coefficients de propagation du bruit mode à mode dans la diversité de phase, pour une amplitude de diversité classique de 1,8 rad.

La tendance de la propagation est plate en bruit de photons, comme montré par Guyon [Guyon, 2005], et augmente avec le numéro du mode en bruit de détecteur. En effet, l'estimation de hauts ordres nécessite de prendre en compte des pixels éloignés du centre de la tache, qui

1. Sur la figure 1.21 du paragraphe 1.4.1.3, les photodiodes au bord du détecteur sont plus grandes, mais ne sont éclairées que sur une zone de dimensions égales aux autres photodiodes.

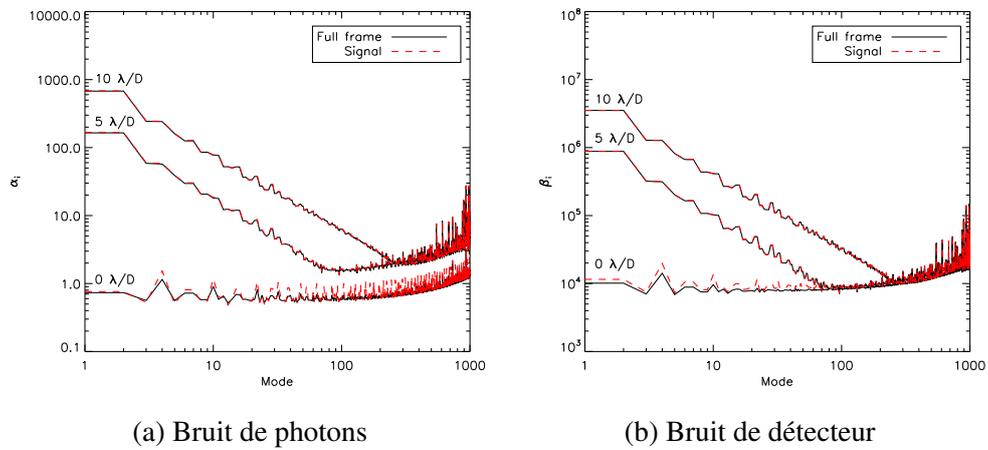


FIGURE 5.3 – Propagation du bruit dans la pyramide en fonction de la modulation, en considérant les données brutes ou les pseudo-données. Plus la modulation augmente, plus la propagation du bruit augmente. Rayons de modulation (de bas en haut) : $0 \lambda/D$, $5 \lambda/2D$, $10 \lambda/D$.

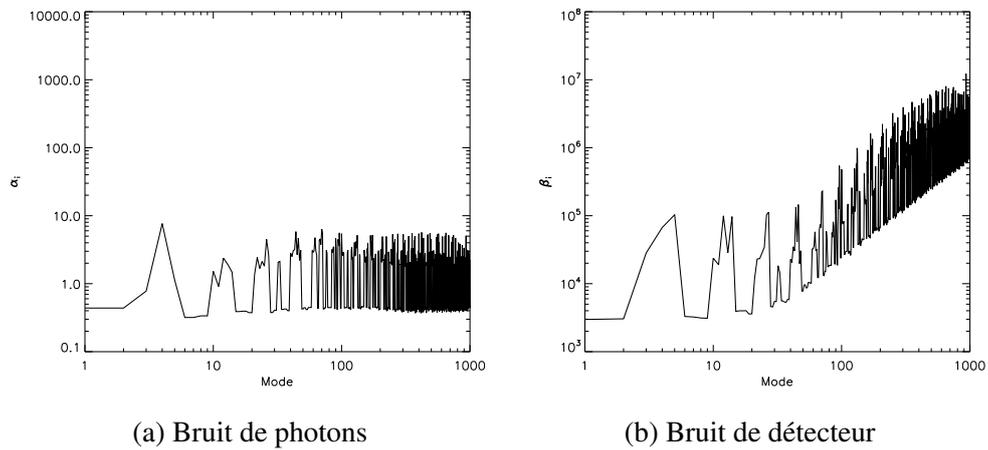


FIGURE 5.4 – Propagation du bruit dans la diversité de phase, pour une image focale et une image défocalisée de 1,8 rad.

reçoivent peu de flux. Le bruit de détecteur étant uniforme, ce résultat est en accord avec celui trouvé par Meynadier [Meynadier *et al.*, 1999].

Les pics réguliers que l'on observe correspondent à des modes ayant au moins deux axes de symétrie, tels que l'astigmatisme et l'aberration sphérique. Ces modes sont indéterminés lorsqu'il n'y a pas de phase de diversité. On peut donc s'attendre à une diminution des pics lorsqu'on augmente l'amplitude de diversité et inversement, une augmentation des pics lorsqu'on diminue l'amplitude de diversité. C'est ce que nous vérifions dans le paragraphe suivant.

Influence de l'amplitude de la diversité La propagation du bruit pour plusieurs amplitudes de diversité (0,1 rad, 1,8 rad et 4 rad) est tracée dans la figure 5.5. On observe en effet une augmentation ou une diminution des pics lorsqu'on change l'amplitude de diversité. L'estimation des hauts ordres bénéficie majoritairement de l'augmentation de l'amplitude de diversité, puisque le flux sur les pixels périphériques augmente.

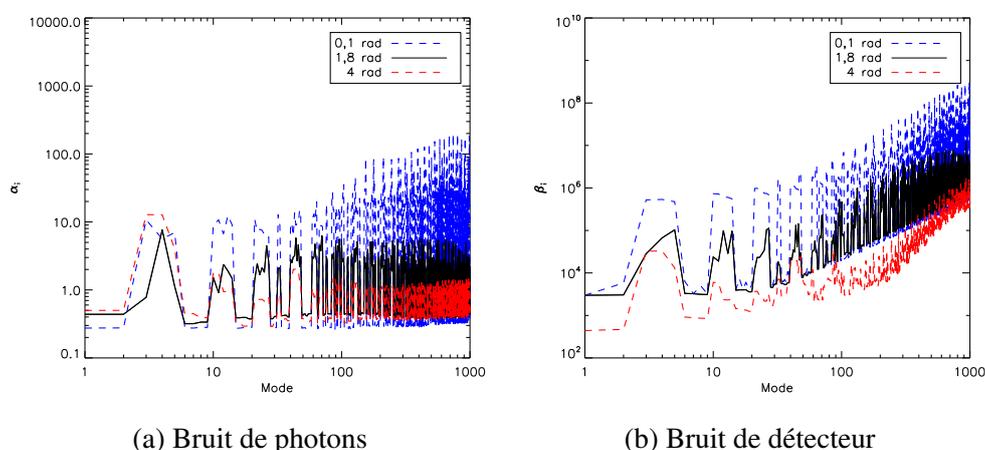


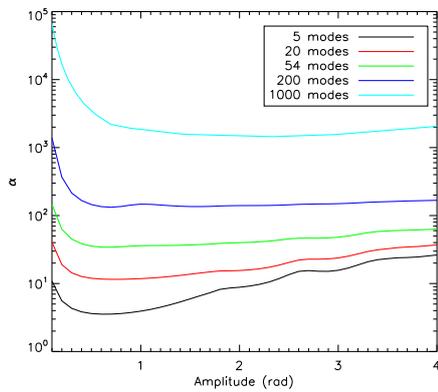
FIGURE 5.5 – Propagation du bruit dans la diversité de phase, pour différentes amplitudes de diversité.

Dans la figure 5.6, nous vérifions s'il existe une amplitude optimale pour un nombre de modes estimés donné, en traçant les coefficients $\alpha = \sum \alpha_i$ et $\beta = \sum \beta_i$.

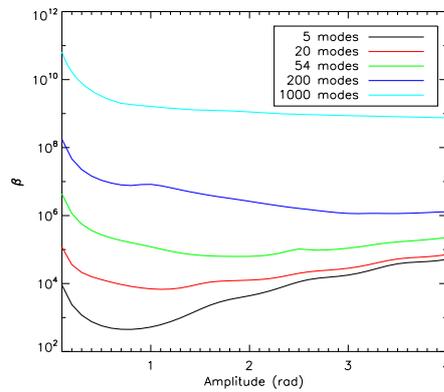
Un minimum de propagation apparaît clairement sur la majorité des courbes. L'amplitude optimale pour le bruit de photons est peu dépendante du nombre de modes estimés. La courbe a cependant tendance à s'aplatir lorsqu'on estime plus de modes. Au final, pour 1000 modes, l'amplitude optimale est à 2,4 rad au lieu de $\sim 0,7$ rad pour les autres cas. En bruit de détecteur, on a clairement intérêt à étaler la tache pour pouvoir estimer efficacement plus de modes. Pour 1000 modes, l'amplitude optimale est au-delà des points considérés.

Selon les conditions d'analyse, un compromis entre l'optimum du bruit de photons et l'optimum du bruit de détecteur permettra d'avoir une analyse efficace.

Influence de la position des images La figure 5.7 montre la propagation du bruit pour une disposition symétrique des images autour du focus. Les images sont défocalisées de $\pm 0,9$ rad. Nous comparons cela à la propagation de bruit pour la diversité classique avec une amplitude de 1,8 rad.

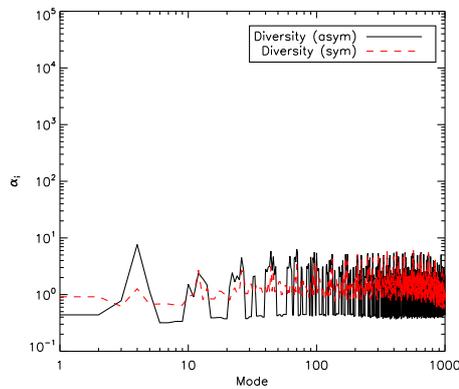


(a) Bruit de photons

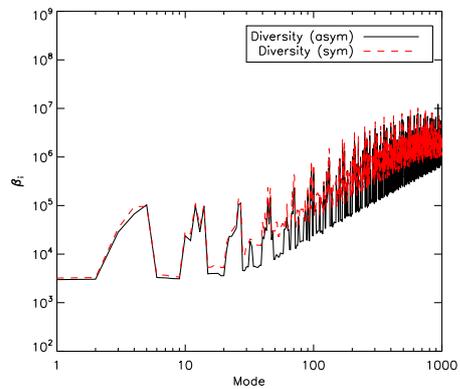


(b) Bruit de détecteur

FIGURE 5.6 – Propagation du bruit globale dans la diversité de phase en fonction de l’amplitude de défocalisation et du nombre de modes estimés, en bruit de photons et en bruit de détecteur.



(a) Bruit de photons



(b) Bruit de détecteur

FIGURE 5.7 – Propagation du bruit dans la diversité de phase, pour une image focale et une image défocalisée de 1,8 rad (configuration asymétrique), ou deux images défocalisées de $\pm 0,9$ rad (configuration symétrique).

Dans cette configuration, le flux est autant étalé sur une image que sur l'autre. Cela permet de diminuer les pics, au détriment des modes qui ne nécessitent pas de diversité pour être estimés, ce qui rehausse la propagation globale. Au final, les propagations du bruit globales sont quasiment identiques dans notre cas. Ainsi, on peut supposer que pour une amplitude de diversité totale équivalente (ici, $0,9 \text{ rad} \times 2 = 1,8 \text{ rad}$), la propagation est la même.

Lee *et al.* avaient au contraire montré que la configuration symétrique était plus performante [Lee *et al.*, 1997]. Cependant, ils avaient comparé les deux configurations pour leurs amplitudes optimales respectives. Le cas traité ici ne tient pas compte de l'amplitude optimale. Notre résultat n'est donc pas forcément incompatible avec le résultat de Lee *et al.*. D'autre part, leur matrice de Fisher est calculée pour seulement 8 modes à partir du focus. Elle est donc très restrictive, et ne tient pas compte de l'estimation du tip/tilt, qui est de manière générale bien meilleure au plan focal que dans un plan extra-focal. En effet, l'estimation du tip/tilt par un maximum de vraisemblance ressemble fortement à l'estimation d'un centre de gravité (cf §1.4.2 p.56), qui est plus efficace lorsque le flux est concentré (cf §1.4.8 p.61).

Influence de la répartition du flux La figure 5.8 montre la propagation du bruit pour une répartition du flux de 30% sur l'image focale et 70% sur l'image extra-focale. Le cas classique 50%/50% est rappelé pour comparaison.

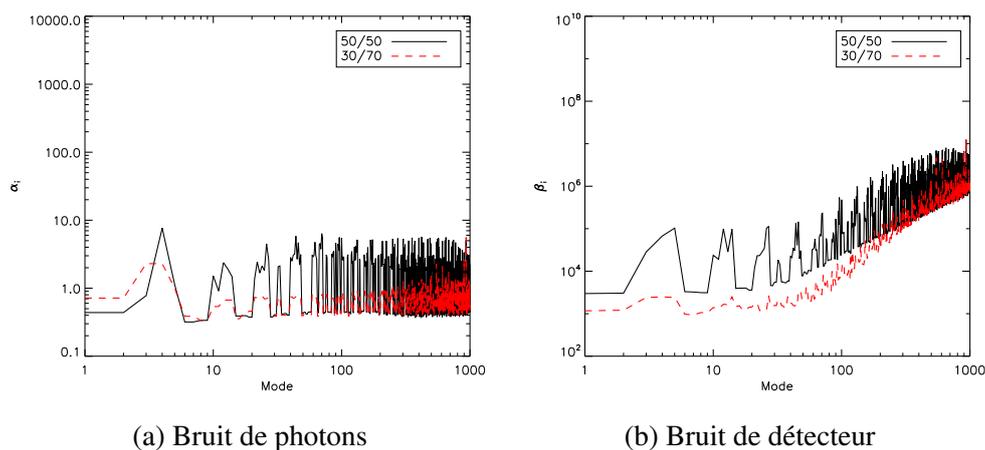


FIGURE 5.8 – Propagation du bruit dans la diversité de phase pour une répartition du flux image focale/image extra-focale de 50%/50% ou 30%/70%. Amplitude de la diversité : 1,8 rad.

On observe une nette amélioration de la propagation, que ce soit en bruit de photons ou en bruit de détecteur, lorsqu'on met plus de flux dans la tache extra-focale. Lee *et al.* avaient pressenti cela mais trouvaient finalement qu'une répartition équitable entre les images était optimale. De même que précédemment, cette différence est peut-être due aux conditions restrictives de leur étude.

5.5.2.3 Analyseur de courbure

La figure 5.9 montre les coefficients de propagation du bruit mode à mode dans l'analyseur de courbure.

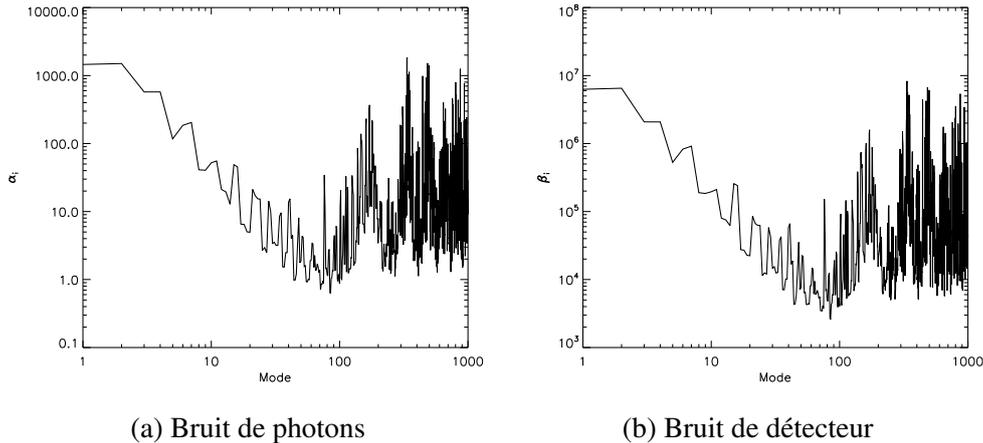


FIGURE 5.9 – Propagation du bruit dans l’analyseur de courbure pour les paramètres donnés dans la partie 5.5.1 p.174.

On a une tendance des courbes en j^{-2} , j étant le numéro du mode, ce qui correspond à $(n + 1)^{-4}$. Cela est cohérent avec la tendance indiquée par Rousset [Rousset, 1999]. La propagation dans les deux régimes de bruit est très similaire au résultat donné par Guyon pour le bruit de photons [Guyon, 2005]. Elle tend vers un optimum pour un mode donné. Les modes de Karhunen-Loève, comme les modes de Zernike, ne représentent pas des fréquences spatiales pures mais des mélanges de hautes et basses fréquences. On peut donc supposer que la fréquence spatiale pour laquelle l’analyseur est optimisé apparaît pour la première fois à l’ordre radial correspondant au mode optimal. Les modes suivants peuvent présenter des fréquences auxquelles l’analyseur n’est pas sensible, ce qui rend la propagation très chahutée.

Le mode optimal pour la propagation du bruit dépend de la distance extra-focale choisie. On peut donc, comme pour la diversité de phase, définir une distance idéale en fonction du nombre de modes estimés. La figure 5.10 trace la somme des coefficients de propagation des 1000 modes de Karhunen-Loève en fonction de la distance extra-focale.

Que ce soit en bruit de photons ou en bruit de détecteur, la longueur extra-focale optimale est 0,2 rad dans notre cas.

5.5.2.4 Masque de Zernike

La figure 5.9 montre les coefficients de propagation du bruit mode à mode dans le masque de Zernike. Nous montrons la propagation pour une image théorique, d’après les formules de la partie 1.4.1.4 p.54, et pour une image simulée.

La propagation reste quasiment la même pour tous les modes, comme indiqué par Guyon [Guyon, 2005]. Une remontée dans les hauts ordres, similaire à ce qui arrive sur la pyramide, est visible. Cela est sûrement dû au même effet de filtrage par l’intégration du flux dans chaque sous-pupille.

La différence entre la propagation donnée par une image théorique et celle donnée par une image calculée en simulation est importante. On a un facteur ~ 4 en bruit de photons et un facteur ~ 7 en bruit de détecteur. Cela est dû aux approximations faites dans les formules théoriques, ainsi qu’au flux perdu par les effets de diffraction ($\sim 50\%$). Cette perte de flux induit

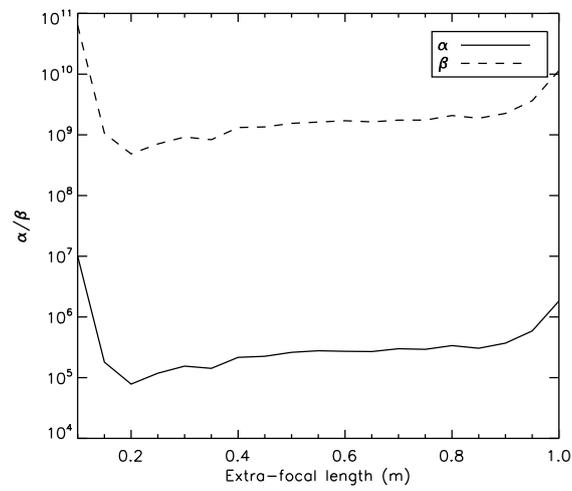
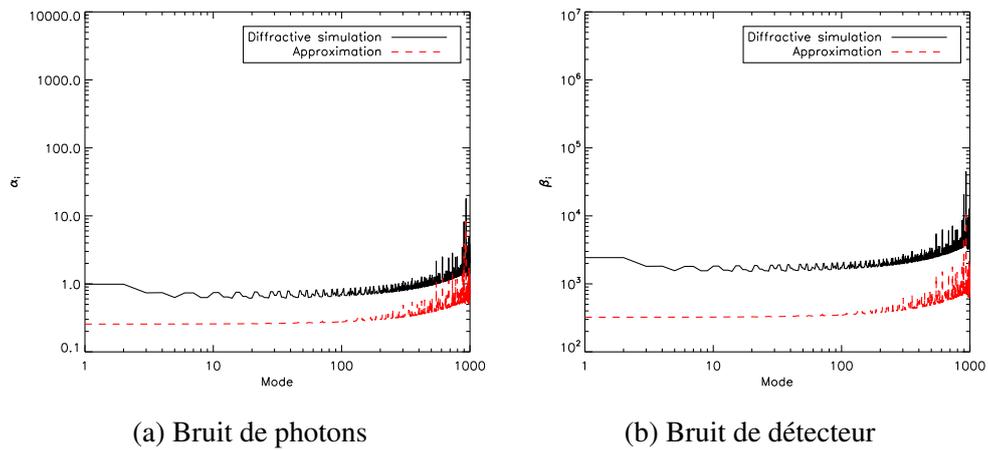


FIGURE 5.10 – Propagation globale dans l’analyseur à courbure en fonction de la distance extra-focale l pour le bruit de photons (coefficient de propagation α) et le bruit de détecteur (coefficient de propagation β).



(a) Bruit de photons

(b) Bruit de détecteur

FIGURE 5.11 – Propagation du bruit dans le masque de Zernike, pour une image théorique (approximation) et une image simulée.

un facteur 2 en bruit de photons et 4 en bruit de détecteur. La perte due aux approximations est donc de l'ordre de 2 en bruit de photons et de détecteur.

5.5.2.5 Bilan comparatif

Nous regroupons maintenant les principaux résultats afin de comparer les analyseurs. La figure 5.12 montre les propagations de bruit pour le Shack-Hartmann, une pyramide non modulée, la diversité de phase dans sa configuration classique, l'analyseur de courbure et le masque de Zernike (sans approximation). L'amplitude de diversité utilisée est 2,4 rad, avec une répartition égale du flux sur les deux images. Toutes les propagations sont calculées pour les données brutes. Pour des raisons de clarté, le Shack-Hartmann LIFTé n'apparaît pas dans cette figure, mais il a été comparé au Shack-Hartmann et à la pyramide dans la partie 5.4 p.154.

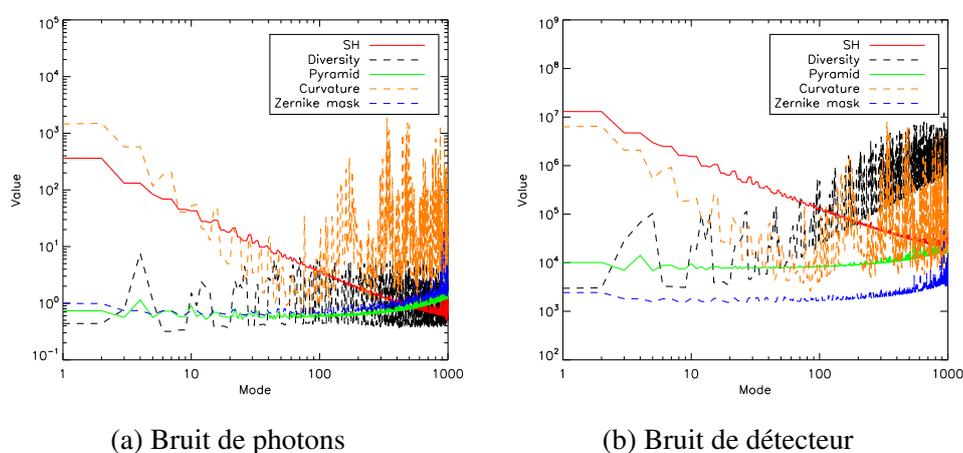


FIGURE 5.12 – Comparaison des propagations de bruit dans divers analyseurs pour l'estimation de 1000 modes de Karhunen-Loève.

En bruit de photons, la diversité de phase, la pyramide et le masque de Zernike offrent tous les trois une propagation faible vis-à-vis du Shack-Hartmann et de l'analyseur de courbure. Ils bénéficient en effet tous les trois du gain pleine pupille. L'analyseur de courbure peut atteindre une propagation équivalente à ces analyseurs, mais seulement pour un ordre radial donné. Comme la littérature le montrait, l'analyseur de courbure est moins performant que le Shack-Hartmann pour l'estimation des bas ordres. Il devient plus performant dans les hauts ordres, jusqu'à ce qu'on dépasse l'ordre radial optimal, après lequel il n'est plus sensible à certaines hautes fréquences spatiales.

En bruit de détecteur, à part pour les très hauts ordres, l'analyseur de courbure est meilleur que le Shack-Hartmann, car il nécessite beaucoup moins de pixels. Les performances de ces analyseurs deviennent comparables à celle de la diversité de phase, qui perd beaucoup en efficacité dans les hauts ordres. La pyramide et le masque de Zernike gardent une faible propagation égale pour tous les modes. Le masque de Zernike est tout de même nettement supérieur à la pyramide puisqu'il répartit le flux sur 4 fois moins de pixels. D'après la partie 1.4.8 p.61, sa

propagation doit donc être 4 fois inférieure à celle de la pyramide (cas d'une division simple du flux), ce qui est vérifié ici.

Dans le tableau 5.1, nous donnons les coefficients de propagation globale de ces analyseurs pour l'estimation de 1000 modes de Karhunen-Loève. La propagation globale du Shack-Hartmann LIFTé 20×20, donnée dans l'article (§5.4 p.154), est rappelée ici.

	α	β
Shack-Hartmann	3 199	$1,15 \cdot 10^8$
Shack-Hartmann LIFTé 20×20	2922	$6,14 \cdot 10^7$
Diversité	1 517	$1,21 \cdot 10^9$
Pyramide non modulée	885	$1,29 \cdot 10^7$
Courbure	58 354	$2,61 \cdot 10^8$
Masque de Zernike	1 178	$2,9 \cdot 10^6$

TABLE 5.1 – Coefficients de propagation du bruit globale des différents analyseurs considérés pour l'estimation de 1000 modes de Karhunen-Loève. Le coefficient de propagation globale est égale à la somme des coefficients mode à mode.

Le masque de Zernike est globalement l'analyseur le plus performant parmi ceux considérés ici. Sa performance en bruit de photons est légèrement moins bonne que celle de la pyramide, mais il offre un gain égal à 4 en bruit de détecteur. Il faut cependant rappeler que cet analyseur souvent considéré comme "idéal" [Guyon, 2005, Paterson, 2008] nécessite aussi des conditions idéales : une phase très faible (amplitude $< 0,1$ rad) et une lumière quasi-monochromatique.

La pyramide non modulée arrive en deuxième position. Elle est quasiment deux fois meilleure que la diversité de phase et offre un gain supérieur à 3 par rapport au Shack-Hartmann en bruit de photons. En bruit de détecteur, elle offre aussi une bien meilleure performance, avec un gain d'environ 10 sur le Shack-Hartmann et 100 sur la diversité de phase. Il faut aussi noter que la pyramide non modulée a une dynamique semblable au masque de Zernike, mais peut fonctionner en lumière polychromatique.

La diversité de phase est deux fois meilleure que le Shack-Hartmann en bruit de photons, mais elle est 10 fois plus sensible au bruit de détecteur. Pour l'estimation d'autant de modes, elle peut donc être avantageuse dans des applications limitées par le bruit de photons, mais elle perd son intérêt si le détecteur présente du bruit.

Le Shack-Hartmann LIFTé 20×20 est juste un peu plus performant que le Shack-Hartmann classique en bruit de photons. Nous rappelons en effet que les données considérées ici pour le Shack-Hartmann classique sont les données brutes, contrairement au cas de l'article présenté au paragraphe 5.4 p.154, où le Shack-Hartmann LIFTé présentait un gain supérieur. Il offre tout de même un gain de 2 par rapport au Shack-Hartmann classique en bruit de détecteur. Même lorsque l'estimation du front d'onde est faite sur les données brutes, le Shack-Hartmann LIFTé reste donc une meilleure solution que le Shack-Hartmann classique.

Enfin, l'analyseur de courbure présente une très mauvaise propagation en bruit de photons à cause des hautes fréquences spatiales non vues. Il pourrait ainsi être compétitif vis-à-vis des autres analyseurs pour l'estimation de moins de modes, mais il est clairement beaucoup moins adapté à une estimation très fine du front d'onde.

En conclusion, le masque de Zernike et la pyramide non modulée sont les meilleurs analyseurs envisageables pour estimer de manière fine un front d'onde très faible. Ils pourraient tout

à fait servir de complément à une optique adaptative extrême. La diversité de phase, ainsi que l'analyseur de courbure sont plutôt adaptés aux applications requérant l'estimation d'un nombre moyen de modes, comme l'optique adaptative classique. Nous rappelons toutefois qu'une estimation en temps réel par la diversité de phase nécessite encore des développements (algorithmiques ou technologiques) pour avoir la dynamique requise dans une optique adaptative classique. Enfin, le Shack-Hartmann LIFTé garde la polyvalence du Shack-Hartmann, en étant plus performant.

Nous avons ici donné un aperçu des performances respectives de divers analyseurs en lumière monochromatique. En pratique, la lumière à analyser est souvent polychromatique, en particulier en optique adaptative, où l'on a besoin de récupérer un maximum de photons. Il paraît donc important de prendre cet aspect en compte. Dans la suite, nous discutons des perspectives pour une comparaison en polychromatique.

5.5.3 Perspectives pour une comparaison en lumière polychromatique

Dans la littérature, l'aspect polychromatique est la plupart du temps négligé ou discuté qualitativement. En effet, bien que le modèle classique d'imagerie polychromatique soit simple (image totale = somme des images à chaque longueur d'onde), les calculs de propagation se complexifient lorsqu'on considère une lumière polychromatique. Pour le bruit de photons, on peut raisonner en moyennant les estimations sur la bande spectrale, car le bruit peut être décomposé sur chaque longueur d'onde, mais un calcul analytique n'est pas forcément évident, même dans un cas simple (voir annexe B de l'article §5.4 p.154). Le bruit de détecteur est quant à lui indépendant de la longueur d'onde, et s'ajoute au signal une fois que les contributions de toutes les longueurs d'onde sont mélangées. Il est donc très difficile d'évaluer sa propagation en lumière polychromatique.

La matrice d'information de Fisher peut alors se révéler comme un outil pratique pour évaluer la propagation du bruit en lumière polychromatique. Elle ne requiert en effet que le modèle simple de formation d'images polychromatiques pour être calculée. La suite présente quelques axes de recherche.

Bruit de détecteur sur les pseudo-données du Shack-Hartmann et de la pyramide Nous avons montré analytiquement l'effet de la polychromaticité sur la propagation du bruit de photons dans les pseudo-données du Shack-Hartmann et de la pyramide, pour un spectre uniforme sur la bande spectrale (voir annexe B de l'article §5.4 p.154). Nous n'avons pas considéré le bruit de détecteur à cause de la complexité des calculs que l'on a évoquée plus haut. L'évaluation de la propagation avec la matrice de Fisher permettrait de vérifier l'analyse en bruit de photons et d'obtenir la propagation en bruit de détecteur. De même, la propagation de bruit sur les données brutes pourra être évaluée.

Sensibilité d'un analyseur plan focal au polychromatisme Un analyseur plan focal, tel que la diversité de phase ou LIFT, estime le front d'onde à partir des effets des aberrations sur l'image dans le plan focal et/ou dans un plan proche du foyer. Ces aberrations évoluent avec la longueur d'onde. La tache image en lumière polychromatique est donc une somme de taches de formes variées. Une fois le flux intégré sur le détecteur, il est difficile de discerner les contributions des différentes longueurs d'onde. On peut donc s'attendre à une augmentation de la propagation du bruit lorsque l'on a une lumière polychromatique. Dans le cas de LIFT, nous

avons montré que l'estimation restait précise en large bande (cf §2.3 p.70), mais nous n'avons jamais quantifié finement la perte en termes de propagation de bruit. Cette grandeur pourra être obtenue via l'information de Fisher, et pourra de plus être utilisée pour étudier la propagation du bruit du Shack-Hartmann LIFTé en lumière polychromatique.

Sensibilité du masque de Zernike au polychromatisme Nous avons évoqué le besoin d'avoir une lumière quasi-monochromatique pour mettre en oeuvre le masque de Zernike. La taille de la pastille et le déphasage qu'elle introduit dépend en effet de la longueur d'onde d'analyse. Il serait intéressant de déterminer la largeur de bande maximale pour laquelle le masque de Zernike est efficace.

5.6 Conclusion

Dans ce chapitre, je me suis appuyé sur la matrice de Fisher pour comparer les principaux analyseurs envisagés en optique adaptative. La comparaison a été faite pour l'estimation de 1000 modes de Karhunen-Loève sur 40×40 sous-pupilles (pour les analyseurs plan pupille), en lumière monochromatique et proche de la limite de diffraction.

J'ai tout d'abord appliqué cette méthode au Shack-Hartmann et j'ai retrouvé les lois de propagation attendues. Par ailleurs, j'ai démontré que l'on gagnait un facteur supérieur à 2 en performance en utilisant toute l'image du Shack-Hartmann plutôt que les pentes dans les sous-pupilles, en bruit de photons comme en bruit de détecteur.

J'ai ensuite revisité la comparaison entre le Shack-Hartmann et la pyramide avec cette méthode, en ajoutant un Shack-Hartmann LIFTé avec deux fois moins de sous-pupilles (20×20). J'ai vérifié le gain apporté par la pyramide vis-à-vis du Shack-Hartmann en bruit de photons, et j'ai montré que la pyramide était aussi beaucoup plus performante en bruit de détecteur (facteur ~ 10 pour toutes les modulations considérées). Le Shack-Hartmann LIFTé, quant à lui, permet de gagner un facteur 3 par rapport au Shack-Hartmann classique 40×40 dans les deux régimes de bruit, et il atteint la performance d'une pyramide faiblement modulée ($2-3 \lambda/D$) en bruit de photons. En bruit de détecteur, il perd un facteur 3 par rapport à la pyramide, mais cet écart pourra être réduit, voire comblé en optimisant la phase de diversité introduite dans les sous-pupilles.

Dans un troisième temps, j'ai étudié la propagation de bruit dans les principaux analyseurs envisagés en optique adaptative. Cette étude a permis de compléter les analyses présentes dans la littérature, notamment en ajoutant la propagation du bruit de détecteur. J'ai ainsi pu comparer ces analyseurs pour les deux termes de bruit, photons et détecteur. J'ai montré que pour une estimation très fine du front d'onde, une pyramide non modulée ou un masque de Zernike offrent la meilleure performance en monochromatique. Ces deux analyseurs ont toutefois une dynamique faible, et doivent être utilisés derrière un premier système correcteur.

Enfin, j'ai évoqué les perspectives pour une étude de la propagation de bruit en polychromatique, qui serait plus représentative des cas réels.

Ces premiers résultats ouvrent la voie vers une comparaison exhaustive, qui permettra un choix plus facile de l'analyseur de front d'onde dans un système d'optique adaptative.

Conclusion et perspectives

Conclusion

L'observation depuis le sol en astronomie avec de grands télescopes est rendue possible grâce à l'optique adaptative, mais reste limitée à un certain nombre de directions d'observation privilégiées. En effet, sur un système d'optique adaptative classique, il faut une étoile suffisamment brillante à proximité de l'objet à observer pour que l'analyseur de front d'onde puisse estimer correctement les déformations induites par la turbulence atmosphérique. Sur un système assisté par laser, les bas ordres doivent être estimés sur une étoile naturelle, et on retrouve un problème similaire. Finalement, dans les deux cas de figure, la couverture de ciel est limitée par la sensibilité au bruit de l'analyseur de front d'onde sur étoile naturelle.

Le travail de cette thèse accorde une place importante à l'étude de la propagation du bruit des analyseurs sur étoile naturelle, que ce soit pour les hauts ordres (optique adaptative classique ou extrême) ou pour les bas ordres (systèmes assistés par laser). L'accent est porté sur un analyseur de bas ordres pour les systèmes assistés par laser, LIFT, proposé par Meimon *et al.* en 2010 [Meimon *et al.*, 2010b]. Le premier objectif est d'optimiser LIFT et de le valider expérimentalement. Le second est de comparer les analyseurs existants pour l'estimation sur étoile naturelle du tip/tilt et focus d'une part, et une estimation des hauts ordres d'autre part, afin de justifier l'utilisation de LIFT et de son dérivé, le Shack-Hartmann LIFTé, dans une optique adaptative.

Dans le chapitre 2 p.65, j'ai abordé l'optimisation de LIFT et sa validation en laboratoire. Pour rendre LIFT plus performant, je me suis intéressé à la phase de diversité qui est introduite pour l'estimation du focus. J'ai d'abord montré que l'astigmatisme, choisi initialement comme mode de diversité, permettait une bonne convergence de l'algorithme. J'ai ensuite déterminé une amplitude d'astigmatisme optimale en m'appuyant sur une étude de la propagation du bruit dans LIFT. Cette étude a aussi permis une première comparaison de LIFT avec d'autres analyseurs. Le gain de LIFT par rapport au Shack-Hartmann 2×2 est supérieur à 1 magnitude et sa performance est comparable à celle de la pyramide, en ayant l'avantage d'une installation et d'une utilisation simples.

Une première étape de validation en laboratoire a permis de vérifier la bonne linéarité de LIFT, la pertinence des modèles de propagation de bruit, ainsi que la possibilité d'utiliser le senseur en large bande spectrale ($\lambda/\Delta\lambda = 3,4$).

Enfin, j'ai comparé LIFT aux principaux analyseurs existants pour l'estimation du tip/tilt et du focus sur une étoile naturelle, en m'appuyant sur la matrice d'information de Fisher. J'ai ainsi confirmé les résultats de la première comparaison, et montré que LIFT était un analyseur très compétitif pour cette application.

Dans un système d'optique adaptative, LIFT bénéficiera de la correction des hauts ordres analysés sur les étoiles laser. La correction n'étant pas parfaite, des hauts ordres résiduels apparaîtront sur les images. Les expériences en laboratoire ne tenaient pas compte de cet aspect. Dans le chapitre 3 p.95, nous avons testé LIFT en boucle ouverte sur le ciel grâce au système d'optique adaptative du télescope Gemini Sud, GeMS. Nous avons démontré la robustesse de LIFT aux hauts ordres en confirmant la linéarité de l'estimation du focus. Ces tests, effectués sur des données du 31 janvier 2013, ont constitué la première validation de LIFT sur le ciel.

En vue d'une validation en boucle fermée, je me suis ensuite intéressé aux effets des hauts ordres sur la performance ultime de LIFT. J'ai d'abord étudié les effets de repliement, ainsi que les solutions pour les minimiser. La première solution, qui consiste à filtrer les hauts ordres par troncature de la fenêtre de calcul, ne s'est pas montrée totalement satisfaisante. La deuxième solution, qui consiste à prendre en compte la statistique *a priori* de la phase résiduelle en utilisant une estimation de type MAP, a permis de minimiser les effets de repliement à des niveaux compatibles avec le budget d'erreur envisagé. La combinaison des deux solutions a été retenue pour la suite de l'étude.

A partir de simulations de phases résiduelles typiques de GeMS, j'ai évalué l'effet de l'étalement de la tache, dû aux hauts ordres, sur la propagation du bruit de LIFT et d'un Shack-Hartmann 2×2 . J'ai alors démontré que le gain apporté par LIFT en termes de propagation de bruit reste le même quel que soit le niveau de correction des hauts ordres. Ensuite, après la simulation d'une première boucle fermée, j'ai souligné l'intérêt de faire des itérations numériques dans l'estimation par LIFT, en mettant en évidence un gain de performance de l'ordre de 10 après 2 itérations seulement.

En progressant sur la validation de LIFT, la question suivante est apparue : LIFT peut-il estimer des hauts ordres sur étoile naturelle, en tant que "truth sensor" ? Comme LIFT utilise une approximation de faible phase, et que les déformations du front d'onde induites par la turbulence sont fortes, la réponse est "non" dans la majorité des cas. Cependant, cela a donné des idées pour concevoir un nouvel analyseur, combinant le Shack-Hartmann et LIFT, justement nommé le Shack-Hartmann LIFTé.

Dans le chapitre 4 p.133, nous avons évalué l'intérêt d'utiliser un Shack-Hartmann LIFTé plutôt qu'un Shack-Hartmann classique. Une première étude a montré que le Shack-Hartmann LIFTé permettait de réduire fortement les effets de repliement et ainsi de contrôler un miroir avec un pas inter-actionneur plus petit que ses sous-pupilles. J'ai approfondi cette étude en évaluant la propagation du bruit dans un Shack-Hartmann LIFTé. J'ai alors démontré qu'un Shack-Hartmann LIFTé 10×10 offrait un gain en performance de l'ordre de 3 en bruit de photons et en bruit de détecteur par rapport à un Shack-Hartmann classique 20×20 .

Afin de placer ce nouveau concept dans la zoologie des analyseurs de front d'onde, j'ai effectué dans le chapitre 5 p.149 une comparaison en termes de propagation de bruit via l'information de Fisher pour une estimation très fine du front d'onde (40×40 sous-pupilles). J'ai ainsi retrouvé, de manière rigoureuse et sans approximation, la plupart des résultats d'études précédentes, tout en incluant la propagation du bruit de détecteur, qui est très rarement considéré dans la littérature.

J'ai notamment comparé le Shack-Hartmann, classique et LIFTé, avec la pyramide. J'ai vérifié le gain apporté par la pyramide par rapport au Shack-Hartmann en bruit de photons (\sim

2 magnitudes), et j'ai montré qu'elle offrait aussi un gain de l'ordre de 10 en bruit de détecteur pour des modulations jusqu'à $6 \lambda/D$, λ étant la longueur d'onde et D le diamètre de la pupille. D'autre part, en bruit de photons, un Shack-Hartmann LIFTé avec deux fois moins de sous-pupilles (20×20 ici) atteint la performance d'une pyramide pour un rayon de modulation entre $2 \lambda/D$ et $3 \lambda/D$, soit une modulation très faible. En bruit de détecteur, il perd un facteur 3, mais cet écart pourra éventuellement être comblé en changeant la phase de diversité dans chaque sous-pupille, qui n'est pour le moment optimisée que pour l'estimation du tip/tilt et du focus.

Perspectives

Les travaux présentés ici peuvent encore être approfondis. Nous décrivons ci-après les différents axes de recherche envisagés.

Court terme

Validation de LIFT sur ciel en boucle fermée La perspective la plus évidente, à court terme, est bien sûr celle-ci. Avec les dernières simulations présentées au chapitre 3 p.95, LIFT est en effet prêt à être testé sur le ciel en boucle fermée et à faible flux. Cette validation finale de LIFT est envisagée sur le télescope Gemini Sud. Le senseur sera de plus testé sur le banc ODISSEE de l'Onera à l'Observatoire de la Côte d'Azur.

Propagation du bruit du Shack-Hartmann LIFTé Dans le chapitre 4 p.133, nous avons montré que le Shack-Hartmann LIFTé avait une loi de propagation similaire au Shack-Hartmann classique. Ainsi, il doit être possible de retrouver cette propagation à partir de la matrice d'interaction via une relation simple, telle que les coefficients de Rigaut et Gendron pour le Shack-Hartmann classique.

Optimisation du Shack-Hartmann LIFTé Comme nous l'avons évoqué précédemment, la phase de diversité introduite dans les sous-pupilles du Shack-Hartmann LIFTé n'est pas optimisée pour l'estimation de plus de 3 modes par sous-pupille. Or, pour gagner un facteur important en performance vis-à-vis du Shack-Hartmann classique, nous avons eu besoin d'estimer 9 modes par sous-pupille. Il faudra donc revoir l'optimisation de la phase de diversité.

Information de Fisher en lumière polychromatique Les propagations de bruit présentées dans la thèse ont été obtenues pour une lumière monochromatique. En pratique, les analyseurs sont plutôt éclairés en lumière polychromatique, pour collecter un maximum de flux. Il sera donc intéressant de calculer la matrice d'information de Fisher polychromatique, afin d'en déduire la propagation de bruit dans un cas plus réaliste.

Moyen terme

Information de Fisher pour un objet étendu La possibilité de calculer la matrice d'information de Fisher pour un objet étendu permettrait d'évaluer la propagation de bruit obtenue lorsque l'on analyse le front d'onde sur une étoile laser, ou sur une pose moyenne. L'extension

de la source dans le champ se modélise par un produit de convolution, qui est une opération linéaire. De ce point de vue, les modifications à apporter au calcul de la matrice de Fisher sont simples. Cependant, les étoiles laser sont aussi étendues sur l'axe d'observation. Les images résultantes seront alors plus difficiles à modéliser. D'autre part, dans le cas de la pyramide, les calculs sont aussi plus complexes, car un objet étendu ne peut pas être modélisé par un produit de convolution. Toutefois, lorsqu'on module la pyramide, on a un objet étendu fictif, et cela a bien été simulé dans la thèse (cf §1.4.1.2 p.50). L'adaptation du calcul de la matrice de Fisher pour le cas des étoiles laser n'est donc pas facile, mais cela permettra de mettre en évidence l'analyseur le mieux adapté à une estimation du front d'onde sur étoile laser.

Validation expérimentale du Shack-Hartmann LIFTé Pour l'instant, le Shack-Hartmann LIFTé n'a fait l'objet que de simulations. Une validation en laboratoire est prévue à l'Onera dans l'année à venir.

Long terme

Implémentation de LIFT dans un système assisté par laser A plus long terme, LIFT pourra être utilisé comme analyseur dédié aux étoiles naturelles dans un système de LTAO ou MCAO. En LTAO, il nécessitera un miroir déformable dédié afin de corriger les hauts ordres dans la direction de l'étoile naturelle et d'introduire l'astigmatisme de diversité, tandis qu'en MCAO, il pourra directement bénéficier de la bonne correction dans tout le champ, et n'aura besoin que d'une simple optique pour l'astigmatisme. D'autre part, une utilisation de LIFT directement sur le directeur scientifique, dans le but de s'affranchir des aberrations différentielles, a été évoquée. Cette application permettrait une estimation du tip/tilt à haute cadence, et focus à basse cadence, sans installation de matériel et sans dégradation significative sur une image longue pose (perte en rapport de Strehl inférieure à 0,2%).

Implémentation du Shack-Hartmann LIFTé dans une optique adaptative extrême Dans la thèse, nous avons démontré que le Shack-Hartmann LIFTé pouvait être une alternative intéressante à la pyramide pour l'estimation de hauts ordres. Il est en effet aussi précis, au moins dans un cas limité par le bruit de photons, qu'une pyramide très faiblement modulée. Une fois qu'il sera optimisé, il sera d'autant plus compétitif. Cet analyseur permettra sans doute d'améliorer encore les performances des optiques adaptatives extrêmes, et par conséquent l'imagerie des exoplanètes.

Annexe A

Démonstrations mathématiques des formules utilisées dans LIFT

A.1 Expression de $I'_k(0)$

L'image $I(\mathbf{A})$ s'exprime :

$$I(\mathbf{A}) = |TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j}\}|^2 = TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j}\} \times TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j}\}^* \quad (\text{A.1})$$

Sa dérivée par rapport au coefficient de Zernike a_k est alors :

$$\frac{\partial I(\mathbf{A})}{\partial a_k} = TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j} i \mathbf{Z}_k\} \times TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j}\}^* + TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j}\} \times TF\{\mathbf{P}_d e^{i\sum a_j \mathbf{Z}_j} i \mathbf{Z}_k\}^* \quad (\text{A.2})$$

Pour une phase nulle ($\mathbf{A} = \mathbf{0}$), on a donc :

$$\left. \frac{\partial I(\mathbf{A})}{\partial a_k} \right|_{\mathbf{A}=\mathbf{0}} = TF\{\mathbf{P}_d i \mathbf{Z}_k\} \times TF\{\mathbf{P}_d\}^* + TF\{\mathbf{P}_d\} \times TF\{\mathbf{P}_d i \mathbf{Z}_k\}^* \quad (\text{A.3})$$

On obtient enfin :

$$I'_k(0) = 2\Re\{\widetilde{\mathbf{P}}_d^* \cdot i TF\{\mathbf{P}_d \cdot \mathbf{Z}_k\}\} \quad (\text{A.4})$$

A.2 Expression de P_{MV}

Dans l'équation 1.49 p.57, \mathbf{n} est considéré comme un bruit gaussien. On peut écrire sa loi de probabilité :

$$p(\mathbf{X}) \propto \exp\left(-\frac{1}{2}\mathbf{X}^t \mathbf{R}_n^{-1} \mathbf{X}\right) \quad (\text{A.5})$$

avec \mathbf{R}_n^{-1} l'inverse de la matrice de covariance de \mathbf{n} . L'estimation par un maximum de vraisemblance consiste à chercher \mathbf{A} tel que la probabilité d'obtenir $\Delta \mathbf{I}$ devienne maximum. Cela revient à maximiser la probabilité d'avoir \mathbf{n} tel que la relation 1.49 p.57 soit vérifiée, la variable étant \mathbf{A} :

$$p(\Delta \mathbf{I} | \mathbf{A}) \propto \exp\left(-\frac{1}{2}(\Delta \mathbf{I} - \mathbf{H}\mathbf{A})^t \mathbf{R}_n^{-1} (\Delta \mathbf{I} - \mathbf{H}\mathbf{A})\right) \quad (\text{A.6})$$

avec $p(\Delta\mathbf{I}|\mathbf{A})$ la probabilité d'avoir $\Delta\mathbf{I}$ sachant \mathbf{A} . Elle est maximum lorsque $(\Delta\mathbf{I} - \mathbf{H}\mathbf{A})^t \mathbf{R}_n^{-1} (\Delta\mathbf{I} - \mathbf{H}\mathbf{A})$ est minimum. On note :

$$\mathbf{F} = (\Delta\mathbf{I} - \mathbf{H}\mathbf{A})^t \mathbf{R}_n^{-1} (\Delta\mathbf{I} - \mathbf{H}\mathbf{A}) \quad (\text{A.7})$$

Alors :

$$\mathbf{F} = \Delta\mathbf{I}^t \mathbf{R}_n^{-1} \Delta\mathbf{I} - \Delta\mathbf{I}^t \mathbf{R}_n^{-1} \mathbf{H}\mathbf{A} - \mathbf{A}^t \mathbf{H}^t \mathbf{R}_n^{-1} \Delta\mathbf{I} + \mathbf{A}^t \mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H}\mathbf{A} \quad (\text{A.8})$$

Le minimum de \mathbf{F} est tel que la dérivée de \mathbf{F} par rapport au vecteur \mathbf{A} est nulle, soit $\frac{\partial \mathbf{F}}{\partial \mathbf{A}} = 0$.

$$\frac{\partial \mathbf{F}}{\partial \mathbf{A}} = \frac{\partial}{\partial \mathbf{A}} (\cancel{\Delta\mathbf{I}^t \mathbf{R}_n^{-1} \Delta\mathbf{I}}) - 2\mathbf{H}^t \mathbf{R}_n^{-1} \Delta\mathbf{I} + 2\mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H}\mathbf{A} \quad (\text{A.9})$$

Dans ce calcul nous avons utilisé les propriétés suivantes :

$$\frac{\partial \mathbf{A}^t \mathbf{X}}{\partial \mathbf{A}} = \frac{\partial \mathbf{X}^t \mathbf{A}}{\partial \mathbf{A}} = \mathbf{X} \quad (\text{A.10})$$

et

$$\frac{\partial \mathbf{A}^t \mathbf{X} \mathbf{A}}{\partial \mathbf{A}} = (\mathbf{X} + \mathbf{X}^t) \mathbf{A} \quad (\text{A.11})$$

Si $\mathbf{X} = \mathbf{X}^t$ on a $\frac{\partial \mathbf{A}^t \mathbf{X} \mathbf{A}}{\partial \mathbf{A}} = 2\mathbf{X}\mathbf{A}$. Par définition, \mathbf{R}_n^{-1} vérifie cette relation, et donc $\mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H}$ aussi.

Nous n'avons plus qu'à résoudre l'équation :

$$2\mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H}\mathbf{A} - 2\mathbf{H}^t \mathbf{R}_n^{-1} \Delta\mathbf{I} = 0 \quad (\text{A.12})$$

On trouve aisément la solution :

$$\hat{\mathbf{A}}_{MV} = (\mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H})^{-1} \mathbf{H}^t \mathbf{R}_n^{-1} \Delta\mathbf{I} \quad (\text{A.13})$$

On a donc bien $\hat{\mathbf{A}}_{MV} = P_{MV} \Delta\mathbf{I}$ avec $P_{MV} = (\mathbf{H}^t \mathbf{R}_n^{-1} \mathbf{H})^{-1} \mathbf{H}^t \mathbf{R}_n^{-1}$.

Annexe B

Fonctionnement de LIFT en pratique

On considère l'estimation de n_{modes_est} modes sur une image de taille $npix \times npix$. La phase correspondant à l'image acquise est l'addition de la phase de diversité et de la phase à estimer. Les étapes de l'estimation sont décrites ci-après.

1. **Détection des pixels défectueux (optionnel)** Les pixels défectueux sont détectés afin de ne pas perturber l'estimation. Dans la version courante, on remplace la valeur de ces pixels par la médiane des pixels environnants. Ce n'est pas rigoureux vis-à-vis de l'estimation, et la variable contenant l'image est modifiée directement. Cette option devra être améliorée.
2. **Ajustement de l'échantillonnage de la pupille** Par défaut, les transformées de Fourier pour le calcul de l'image de référence $I(A_d)$ et de la matrice d'interaction $H(A_d)$ (cf §2.2 p.66) sont effectués sur $npix \times npix$ pixels. La taille du champ $npix$ et l'échantillonnage en plan focal ech ($ech = 2$ à Shannon) sont liés à l'échantillonnage de la pupille par la relation $\Phi_{pup} = npix/ech$, avec Φ_{pup} le diamètre de la pupille. Si $npix$ est petit, la phase peut être mal échantillonnée. Pour augmenter l'échantillonnage de la pupille, il est alors possible de définir un nouveau champ $npix' > npix$ qui sera utilisé pour effectuer les transformées de Fourier (voir Fig. B.1). Dans la suite de l'explication, $npix$ et $npix'$ ne sont pas différenciés.
3. **Estimation sur une image de référence (optionnel)** Il est possible d'ajouter une image de référence en entrée, et d'estimer n_{modes_calib} modes sur cette image. On obtient alors le point de fonctionnement que l'on considérera comme phase de référence $phase_ref$.
4. **Centrage de la zone utile** L'estimation de la phase se fait par une pondération sur une zone carrée (zone "utile") dont le centre est confondu avec le centre de gravité de la tache, et dont la taille $npixutil \times npixutil$ est spécifiée par l'utilisateur (voir Fig. B.1). Ce premier centrage permet notamment d'avoir un domaine de linéarité sur le tip/tilt qui n'est limité que par la taille de l'image en entrée. Par défaut, la position de la tache est calculée par centre de gravité classique, avec un seuil égal à 5% du maximum. Elle peut aussi être évaluée par corrélation avec une image non aberrante (mais présentant la phase de diversité). Le centre de gravité obtenu est reporté à la fin sur l'estimation du tip/tilt.
5. **Sélection de la zone utile** A partir du centre précédemment calculé, on extrait la zone utile, de taille $npixutil \times npixutil$.
6. **Estimation du flux** Si le flux n'est pas connu avant la mise en oeuvre de l'algorithme (cas nominal), il est estimé par la somme des valeurs des pixels dans toute la zone utile

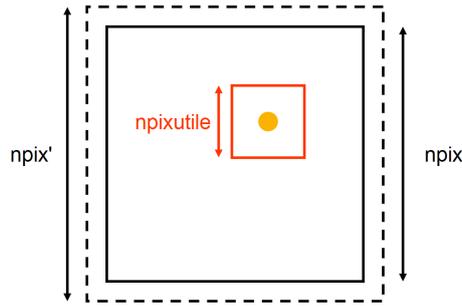


FIGURE B.1 – Représentation de l’image en entrée dans LIFT, ainsi que la zone utilisée pour l’estimation. La tache focale est représentée par le disque orange. La taille de l’image peut être artificiellement augmentée de $npix$ à $npix'$, afin de mieux échantillonner la pupille dans les calculs de transformée de Fourier (cf paragraphe "Ajustement de l’échantillonnage de la pupille").

(option par défaut). Il est aussi possible de l’estimer en même temps que les aberrations (cf étape 7.e), mais le bon fonctionnement de cette option n’a pas été confirmé, et elle ne semble pas plus efficace que l’option par défaut.

7. **Estimation de la phase** L’estimation est faite par itérations successives. A chaque itération, un nouveau point de fonctionnement est pris en compte à partir des estimations précédentes, et la différence d’aberrations entre l’image et ce point de fonctionnement est évaluée. Cette différence est ensuite ajoutée à l’estimation globale. Voici les différentes étapes de la boucle d’estimation :

- (a) **Calcul de $I(\mathbf{A}_d)$ et de $H(\mathbf{A}_d)$** L’image $I(\mathbf{A}_d)$ est calculée à partir du champ complexe en plan pupille correspondant au point de fonctionnement courant (\mathbf{A}_d , cf §2.2 p.66). Ce champ complexe, ainsi que les modes à estimer, interviennent dans le calcul de la dérivée de l’image par rapport aux coefficients des modes à estimer. Les dérivées permettent de constituer la matrice d’interaction $H(\mathbf{A}_d)$ (1 colonne = 1 dérivée). Les calculs sont faits sur $npix \times npix$, puis la zone utile, de taille $npix_{util} \times npix_{util}$, est extraite.
- (b) **Calcul de la matrice de covariance de bruit** Les variances de bruit sur chaque pixel (cf §1.2.1.1 p.22) sont intégrées dans la matrice de covariance de bruit. On considère que le bruit n’est pas corrélé d’un pixel à l’autre, cette matrice est donc diagonale. Un seuil est appliqué par défaut sur la matrice ($10^{-6} \times \max$) pour éviter d’éventuelles divergences lors de l’inversion. Ce seuil peut être modifié par l’utilisateur.
- (c) **Calcul de la pondération P_{MV}** Voir annexe A.2 p.189.
- (d) **Estimation de la phase** Estimation de la différence entre la phase vraie et le point de fonctionnement courant par application de la pondération P_{MV} sur la zone utile.
- (e) **Estimation du flux (optionnel)** Estimation du flux à partir du critère Maximum de Vraisemblance. N’a pas été validé.
- (f) **Calcul du critère** Si l’évolution relative du critère (cf 2.2 p.66) ou de l’estimation par rapport à l’itération précédente est faible, la boucle s’arrête. Sinon, on s’arrête lorsque le nombre d’itérations spécifié est atteint.

Annexe C

Propagation du bruit au sein d'une sous-pupille du Shack-Hartmann LIFTé

LIFT utilise une estimation de maximum de vraisemblance. La solution donnée par cette estimation est une pondération qui est appliquée à la différence entre l'image acquise et une image de référence (cf §1.4.2 p.56). Pour n pixels dans l'image et m modes estimés dans la sous-pupille, le vecteur d'hyperpentes \mathbf{s} s'exprime :

$$\mathbf{s} = \underbrace{\begin{pmatrix} P_{21} & P_{22} & \dots & P_{2n} \\ P_{31} & P_{32} & \dots & P_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1} & P_{m2} & \dots & P_{m_n} \end{pmatrix}}_{\mathbf{P}_{\text{LIFT}}} [\Delta\bar{\mathbf{I}} + \mathbf{w}] \quad (\text{C.1})$$

$\mathbf{P}_i = [P_{i1}, P_{i2}, \dots, P_{in}]$ est la carte de pondération pour le mode de Zernike Z_i , $\Delta\bar{\mathbf{I}} = [\Delta\bar{I}_1, \dots, \Delta\bar{I}_n]$ est la moyenne de la différence d'images et \mathbf{w} est le bruit sur l'image. Nous faisons l'approximation d'un bruit additif gaussien, de moyenne nulle et de variance $\sigma_k^2 = \bar{I}_k + \sigma_e^2$ sur le k -ième pixel, \bar{I}_k étant le flux moyen sur ce pixel et σ_e l'écart-type de bruit de détecteur.

Le bruit propagé sur l'estimation, \mathbf{b} , vaut alors :

$$\mathbf{b} = \mathbf{P}_{\text{LIFT}} \mathbf{w} \quad (\text{C.2})$$

et la matrice de covariance d'erreur d'estimation sur la sous-pupille vaut :

$$\mathbf{C}_b = \langle \mathbf{b}\mathbf{b}^t \rangle = \mathbf{P}_{\text{LIFT}} \langle \mathbf{w}\mathbf{w}^t \rangle \mathbf{P}_{\text{LIFT}}^t \quad (\text{C.3})$$

En considérant que le bruit entre les pixels est décorrélé, on a :

$$\langle \mathbf{w}\mathbf{w}^t \rangle = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n^2 \end{pmatrix} \quad (\text{C.4})$$

Les éléments de la matrice de covariance d'erreur d'estimation valent donc :

$$\mathbf{C}_{\mathbf{b}i,j} = \sum_{k=0}^n P_{ik} P_{jk} \sigma_k^2 \quad (\text{C.5})$$

Le même raisonnement peut être appliqué pour le centre de gravité (CdG) et le centre de gravité pondéré (CdGP), qui peuvent être mis sous la forme d'une pondération appliquée sur l'image. On trouve alors respectivement $\mathbf{C}_b = \sigma_{CdG}^2 \times Id$ et $\mathbf{C}_b = \sigma_{CdGP}^2 \times Id$, σ_{CdG}^2 et σ_{CdGP}^2 étant les variances de bruit du CdG et du CdGP telles que définies par Magalie Nicolle et Sandrine Thomas [Nicolle *et al.*, 2004, Nicolle, 2006, Thomas *et al.*, 2006].

Annexe D

Information de Fisher pour un modèle de bruit additif gaussien

Nous montrons ici l'expression générale de la matrice de Fisher en présence d'un bruit additif gaussien, ainsi que le lien avec la matrice de covariance d'un estimateur de maximum de vraisemblance.

D.1 Expression générale

On considère un ensemble de données $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, qui dépendent d'un ensemble de coefficients de modes (de Zernike par exemple) $\mathbf{A} = \{a_1, a_2, \dots, a_m\}$ et d'un bruit $\mathbf{b} = \{b_1, b_2, \dots, b_n\}$. L'expression de la matrice de Fisher est (cf Eq. 1.56) :

$$F_{i,j}(\mathbf{A}_p) = E\left\{\left[\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p} \left[\frac{\partial}{\partial a_j} \ln p(\mathbf{y}|\mathbf{A})\right]\Big|_{\mathbf{A}=\mathbf{A}_p}\right\} \quad (\text{D.1})$$

avec $p(\mathbf{y}|\mathbf{A})$ la densité de probabilité de vraisemblance et \mathbf{A}_p le point de fonctionnement.

Pour un bruit additif gaussien, la densité de probabilité de vraisemblance s'écrit :

$$p(\mathbf{y}|\mathbf{A}) = p_b(\mathbf{y}) = \prod_{k=1}^N \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{[y_k - \bar{y}_k(\mathbf{A})]^2}{2\sigma_k^2}\right\} \quad (\text{D.2})$$

Pour trouver $\mathbf{F}(\mathbf{A}_p)$, il faut calculer $\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p}$. On peut écrire :

$$\ln p(\mathbf{y}|\mathbf{A}) = \sum_{k=1}^N -\frac{1}{2} \ln(2\pi\sigma_k^2) - \frac{[y_k - \bar{y}_k(\mathbf{A})]^2}{2\sigma_k^2} \quad (\text{D.3})$$

$$\frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})] \quad (\text{D.4})$$

Donc :

$$\begin{aligned} \frac{\partial}{\partial a_i} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} \times \frac{\partial}{\partial a_j} \ln p(\mathbf{y}|\mathbf{A})\Big|_{\mathbf{A}=\mathbf{A}_p} &= \sum_{k=1}^N \frac{1}{\sigma_k^4} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})]^2 \\ &+ \sum_{l \neq k} \frac{1}{\sigma_l^2 \sigma_k^2} \frac{\partial \bar{y}_l(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \times [y_k - \bar{y}_k(\mathbf{A})] \times [y_l - \bar{y}_l(\mathbf{A})] \end{aligned} \quad (\text{D.5})$$

Sachant que $E\{[y_k - \bar{y}_k(\mathbf{A})]^2\} = \sigma_k^2$ et $E\{[y_k - \bar{y}_k(\mathbf{A})][y_l - \bar{y}_l(\mathbf{A})]\} = \sigma_{kl}^2$, on a finalement :

$$F_{i,j}(\mathbf{A}_p) = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial \bar{y}_l(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \quad (\text{D.6})$$

Si on considère que le bruit est décorréolé d'une donnée y_i à une autre, l'expression se simplifie et devient :

$$F_{i,j}(\mathbf{A}_p) = \sum_{k=1}^N \frac{1}{\sigma_k^2} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_i} \Big|_{\mathbf{A}=\mathbf{A}_p} \frac{\partial \bar{y}_k(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} \quad (\text{D.7})$$

D.2 Lien avec l'estimateur de maximum de vraisemblance

D'après le paragraphe précédent, on peut exprimer la matrice de Fisher de la manière suivante :

$$\mathbf{F}_{i,j} = \sum_k \frac{1}{\sigma_k^2} \frac{\partial y_k}{\partial a_i} \frac{\partial y_k}{\partial a_j} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial y_l}{\partial a_i} \frac{\partial y_k}{\partial a_j} \quad (\text{D.8})$$

avec σ_k^2 la variance d'erreur sur la k -ième donnée et σ_{kl}^2 la covariance d'erreur entre les données k et l .

Comparons maintenant la matrice de covariance d'un estimateur de vraisemblance (cf 1.4.4 p.59) $\langle \mathbf{E}\mathbf{E}^t \rangle = (\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D})^{-1}$ et \mathbf{F} .

Nous devons d'abord trouver l'expression de la matrice d'interaction \mathbf{D} . D'après l'équation 1.51 (§1.4.4 p.59), on peut écrire :

$$\mathbf{y}(\mathbf{A} + \delta \mathbf{A}) = \mathbf{D}\mathbf{A} + \mathbf{D}\delta \mathbf{A} + \mathbf{n} = \mathbf{y}(\mathbf{A}) + \mathbf{D}\delta \mathbf{A} \quad (\text{D.9})$$

avec $\delta \mathbf{A}$ une petite perturbation des modes. On a alors :

$$\mathbf{D} = \frac{\mathbf{y}(\mathbf{A} + \delta \mathbf{A}) - \mathbf{y}(\mathbf{A})}{\delta \mathbf{A}} \triangleq \frac{\partial \mathbf{y}}{\partial \mathbf{A}} \quad (\text{D.10})$$

Pour m modes et p données, l'expression de \mathbf{D} est finalement :

$$\mathbf{D} = \begin{pmatrix} \frac{\partial y_1}{\partial a_1} & \frac{\partial y_1}{\partial a_2} & \cdots & \frac{\partial y_1}{\partial a_m} \\ \frac{\partial y_2}{\partial a_1} & \frac{\partial y_2}{\partial a_2} & \cdots & \frac{\partial y_2}{\partial a_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_p}{\partial a_1} & \frac{\partial y_p}{\partial a_2} & \cdots & \frac{\partial y_p}{\partial a_m} \end{pmatrix} \quad (\text{D.11})$$

Sachant que la matrice de covariance de bruit est :

$$\mathbf{C}_n = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{p1}^2 & \cdots & \cdots & \sigma_p^2 \end{pmatrix} \quad (\text{D.12})$$

On trouve :

$$[\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D}]_{i,j} = \sum_k \frac{1}{\sigma_k^2} \frac{\partial y_k}{\partial a_i} \frac{\partial y_k}{\partial a_j} + \sum_{l \neq k} \frac{\sigma_{kl}^2}{\sigma_l^2 \sigma_k^2} \frac{\partial y_l}{\partial a_i} \frac{\partial y_k}{\partial a_j} \quad (\text{D.13})$$

Donc $\mathbf{D}^t \mathbf{C}_n^{-1} \mathbf{D} = \mathbf{F}$, soit $\langle \mathbf{E} \mathbf{E}^t \rangle = \mathbf{F}^{-1}$. La matrice de covariance d'erreur de l'estimateur de maximum de vraisemblance est donc égale à la borne de Cramér-Rao dans l'hypothèse d'un bruit additif gaussien.

Annexe E

Calcul de l'information de Fisher en pratique

D'après l'équation D.7 p.196, le calcul de la matrice de Fisher ne nécessite que la connaissance de la variance de bruit sur les données et de la dérivée des données par rapport aux coefficients des modes à estimer. Dans cette annexe, nous montrons le calcul des dérivées des données brutes, c'est-à-dire l'image sur le détecteur, de chaque analyseur considéré. Les images pourront être formées dans le plan focal ou dans le plan pupille, selon l'analyseur. En appliquant le modèle de bruit décrit au paragraphe 1.2.1.1 p.22 à partir de l'image donnée par l'analyseur au point de fonctionnement \mathbf{A}_p considéré, on peut alors déduire la matrice de Fisher. Dans toute la suite, les images en plan focal sont données pour un échantillonnage à Shannon et pour un point source à l'infini, en lumière monochromatique. Les principes des analyseurs considérés ici et les méthodes de simulation utilisées sont présentés dans la partie 1.4 p.46.

Remarque : Certains analyseurs peuvent avoir des images qui sont des regroupements de plusieurs sous-images. Dans ce cas, l'image totale sera exprimée en un vecteur de sous-images à l'aide d'accolades.

E.1 Analyseur plan focal à image unique : LIFT

Le calcul qui va suivre est valable pour LIFT, mais aussi pour n'importe quel autre analyseur plan focal. Dans le cas de LIFT, on forme une image focale d'un point à l'infini avec un astigmatisme de 0.5 rad. On a $\mathbf{A}_p = \{0, 0, 0, 0.5, 0, \dots, 0\}$. L'expression de la répartition d'intensité dans l'image est la suivante :

$$I(\mathbf{A}) = |\text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})]|^2 \quad (\text{E.1})$$

où TF désigne la transformée de Fourier, \mathbf{P} est la fonction de transmission de la pupille et \mathbf{Z} est une matrice regroupant les polynômes de Zernike. On a $\mathbf{Z}\mathbf{A} = \sum_k a_k \mathbf{Z}_k$. La figure E.1a montre l'image résultante.

On peut réécrire l'équation E.1 de la manière suivante :

$$I(\mathbf{A}) = \text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})] \times \text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})]^* \quad (\text{E.2})$$

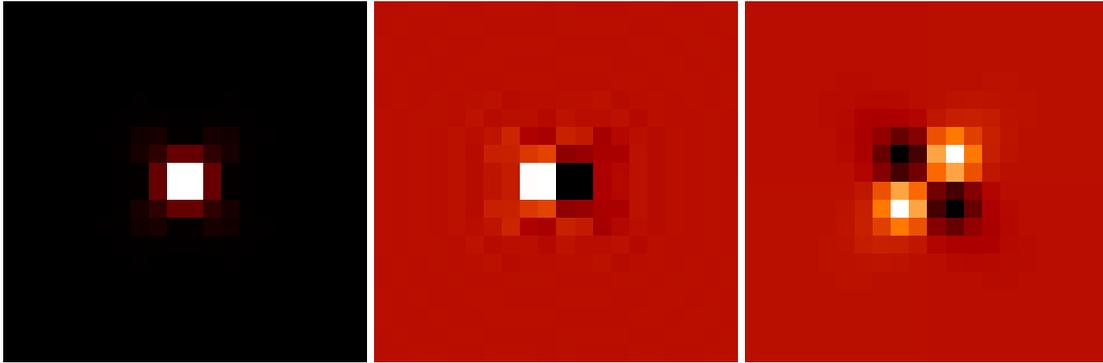
Dérivons maintenant sur a_j :

$$\frac{\partial \mathbf{I}(\mathbf{A})}{\partial a_j} = \text{TF}[i\mathbf{P}\mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A})] \times \text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})]^* + \text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})] \times \text{TF}[i\mathbf{P}\mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A})]^* \quad (\text{E.3})$$

On a donc :

$$\left. \frac{\partial \mathbf{I}(\mathbf{A})}{\partial a_j} \right|_{\mathbf{A}=\mathbf{A}_p} = 2\Re\{i\text{TF}[\mathbf{P}\mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A}_p)] \times \text{TF}[\mathbf{P} \exp(i\mathbf{Z}\mathbf{A}_p)]^*\} \quad (\text{E.4})$$

Les dérivées pour un tip et un focus sont montrées dans les figures [E.1b](#) et [E.1c](#). Ces dérivées interviennent dans le calcul de la matrice d'interaction de LIFT, pour finalement donner les pondérations vues au paragraphe [1.4.2](#) p.56.

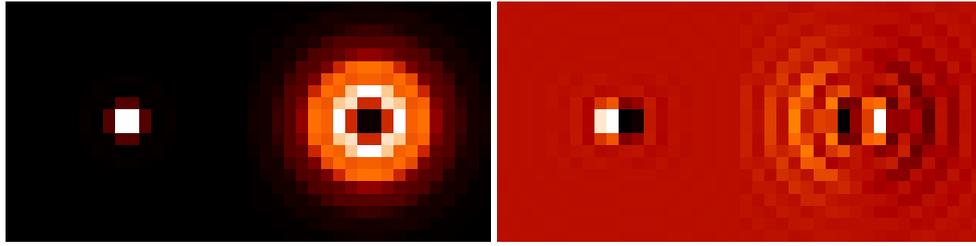


(a) Image LIFT en l'absence d'aberrations. (b) Dérivée par rapport au tip. (c) Dérivée par rapport au focus.

FIGURE E.1 – Image LIFT et exemples de dérivées.

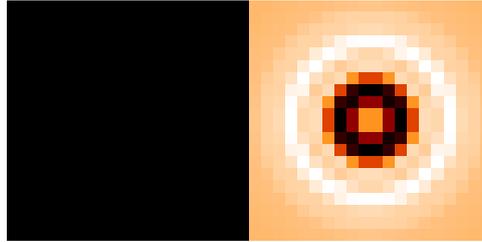
E.2 Diversité de phase

La diversité de phase est un algorithme qui utilise une image focale et une image extra-focale. On calcule la dérivée des deux images de la même manière que précédemment, à partir de l'équation [E.4](#). La figure [E.2](#) donne un exemple d'images et de dérivées.



(a) Image plan focal (gauche) et image extra-focale (droite, defocus de 1.8 rad rms) en l'absence d'aberrations.

(b) Dérivées par rapport au tip.



(c) Dérivées par rapport au focus.

FIGURE E.2 – Images en plan focal et extra-focales et leurs dérivées. Dans chaque couple, les images ne sont pas à la même échelle.

E.3 Shack-Hartmann

Le Shack-Hartmann utilise un réseau de micro-lentilles pour diviser le champ incident. Chaque micro-lentille focalise la partie du champ qu'elle reçoit pour créer au final un réseau d'imassettes. En considérant que les imassettes sont formées indépendamment, l'image finale s'exprime :

$$\mathbf{I}(\mathbf{A}) = \{|\text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A})]|\}^2 = \{\mathbf{I}_1(\mathbf{A})\}_l \quad (\text{E.5})$$

avec \mathbf{T}_1 la transmission d'une micro-lentille.

La dérivée de l'image par rapport aux coefficients de Zernike est alors simplement la concaténation des dérivées des imassettes. Comme précédemment, on peut réécrire l'expression de l'imassette :

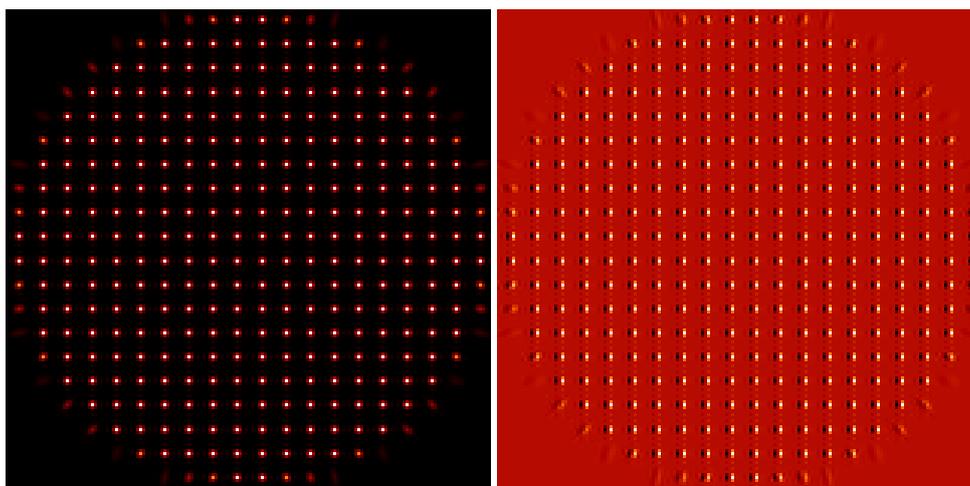
$$\mathbf{I}_1(\mathbf{A}) = \text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A})] + \text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A})]^* \quad (\text{E.6})$$

D'où

$$\begin{aligned} \frac{\partial \mathbf{I}_1(\mathbf{A})}{\partial a_j} &= \text{TF}[i\mathbf{P}\mathbf{T}_1 \mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A})] \times \text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A})]^* \\ &\quad + \text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A})] \times \text{TF}[i\mathbf{P}\mathbf{T}_1 \mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A})]^* \end{aligned} \quad (\text{E.7})$$

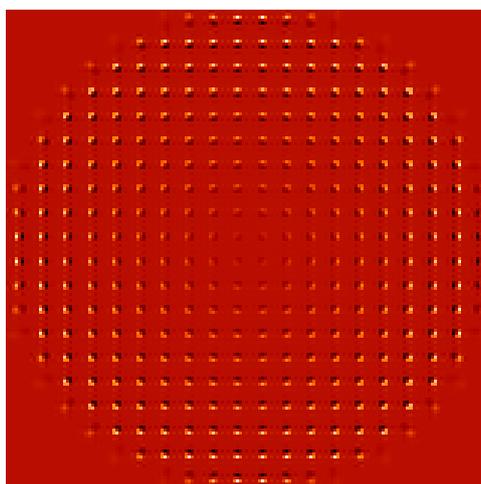
$$\left. \frac{\partial \mathbf{I}_1(\mathbf{A})}{\partial a_j} \right|_{\mathbf{A}=\mathbf{A}_p} = 2\Re\{i\text{TF}[\mathbf{P}\mathbf{T}_1 \mathbf{Z}_j \exp(i\mathbf{Z}\mathbf{A}_p)] \times \text{TF}[\mathbf{P}\mathbf{T}_1 \exp(i\mathbf{Z}\mathbf{A}_p)]^*\} \quad (\text{E.8})$$

Les dérivées pour un tip et un focus sont montrées dans les figures E.3b et E.3c.



(a) Image d'un Shack-Hartmann 20x20 en l'absence d'aberrations.

(b) Dérivée par rapport au tip.



(c) Dérivée par rapport au focus.

FIGURE E.3 – Image Shack-Hartmann et exemples de dérivées.

E.4 Pyramide

L'analyseur à pyramide s'inspire de l'expérience du couteau de Foucault. Un prisme pyramidal coupe le plan focal en 4. Les faisceaux en sortie sont ensuite détectés dans un plan pupille. L'expression de l'image est :

$$\mathbf{I}(\mathbf{A}) = \left\{ \left| \text{TF}^{-1} \{ \mathbf{H}_1 \text{TF} [\mathbf{P} \exp(i\mathbf{Z}\mathbf{A})] \} \right|^2 \right\}_l = \{ \mathbf{I}_l(\mathbf{A}) \}_l \quad (\text{E.9})$$

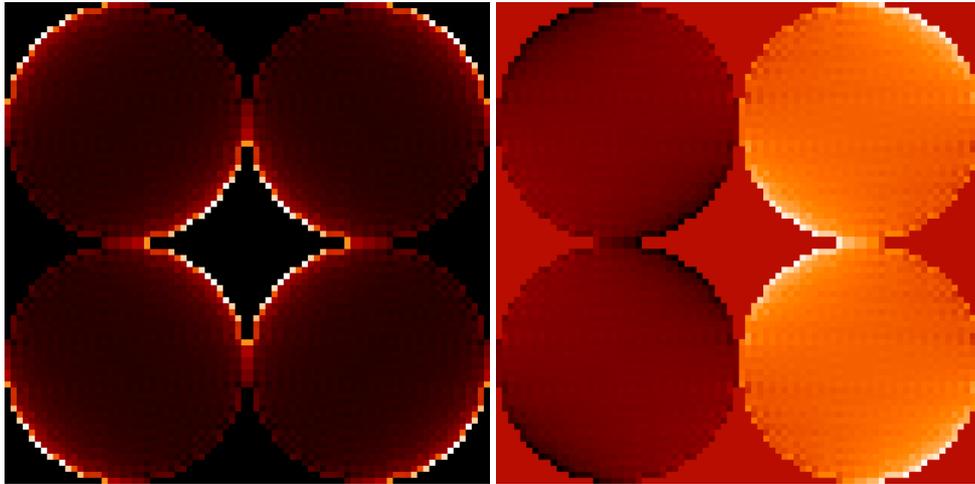
où \mathbf{H}_1 est une fonction de Heaviside 2D sélectionnant un quart du plan focal. La figure E.4a montre l'image résultante.

De la même manière que précédemment, on calcule la dérivée d'une sous-image $\mathbf{I}_l(\mathbf{A})$:

$$\begin{aligned} \frac{\partial \mathbf{I}_1(\mathbf{A})}{\partial a_j} = & \text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[i\mathbf{PZ}_j \exp(i\mathbf{ZA})]\} \times \text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[\mathbf{P} \exp(i\mathbf{ZA})]\}^* \\ & + \text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[\mathbf{P} \exp(i\mathbf{ZA})]\} \times \text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[i\mathbf{PZ}_j \exp(i\mathbf{ZA})]\}^* \end{aligned} \quad (\text{E.10})$$

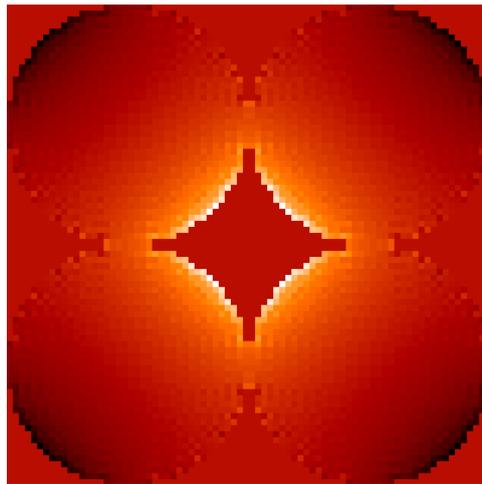
$$\left. \frac{\partial \mathbf{I}_1(\mathbf{A})}{\partial a_j} \right|_{\mathbf{A}=\mathbf{A}_p} = 2\Re\{\text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[i\mathbf{PZ}_j \exp(i\mathbf{ZA}_p)]\} \times \text{TF}^{-1}\{\mathbf{H}_1 \text{TF}[\mathbf{P} \exp(i\mathbf{ZA}_p)]\}^*\} \quad (\text{E.11})$$

Les dérivées pour un tip et un focus sont montrées dans les figures E.4b et E.4c.



(a) Image d'un analyseur à pyramide en l'absence d'aberrations.

(b) Dérivée par rapport au tip.



(c) Dérivée par rapport au focus.

FIGURE E.4 – Image d'un analyseur à pyramide et exemples de dérivées. Les 4 images de pupilles apparaissent normalement à une certaine distance les unes des autres, mais elles ont été rapprochées ici pour diminuer la taille de l'image.

E.5 Courbure

L'analyseur à courbure repose sur le fait que les variations de phase se transforment en variations d'intensité lors de la propagation d'un champ. La courbure du front d'onde est déduite de deux images prises à une même distance de part et d'autre du foyer.

En oubliant l'aspect numérique dans un premier temps, on peut exprimer l'image dans chaque plan de la manière suivante :

$$I_l(r, \phi) = \left| \int_{l=\{1,2\}} u_0 \exp(i\phi) P(\rho) s_l(\rho, r) d\rho \right|^2 \quad (\text{E.12})$$

avec r le vecteur de coordonnées dans le plan extra-pupillaire, ρ le vecteur de coordonnées dans la pupille, $\phi = \sum a_j Z_j$ la phase dans la pupille, u_0 le module du champ dans la pupille, $P(\rho)$ la fonction de transmission de la pupille, et $s_l(\rho, r)$ la fonction de propagation du champ. La figure E.5a montre les images résultantes dans chaque plan.

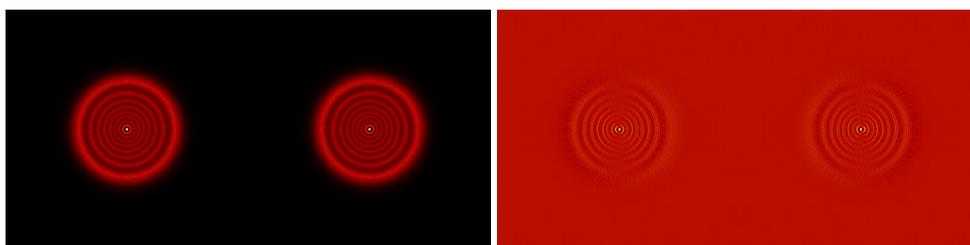
La dérivée donne, par analogie avec les parties précédentes :

$$\frac{\partial I_l(r, \phi)}{\partial a_j} = 2\Re \left\{ \underbrace{\left(\int u_0 \exp(i\phi) P(\rho) s_l(\rho, r) d\rho \right)^*}_{C_l(\phi)} \times i \underbrace{\int Z_j u_0 \exp(i\phi) P(\rho) s_l(\rho, r) d\rho}_{C'_{lj}(\phi)} \right\} \quad (\text{E.13})$$

En reprenant l'aspect numérique avec des vecteurs de données, on a alors :

$$\frac{\partial \mathbf{I}_l(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} = 2\Re \{ \mathbf{C}_l(\mathbf{A}_p)^* \times \mathbf{C}'_{lj}(\mathbf{A}_p) \} \quad (\text{E.14})$$

Les dérivées pour un tip et un focus sont montrées dans les figures E.5b et E.5c.



(a) Images d'un analyseur de courbure en l'absence d'aberrations.

(b) Dérivées par rapport au tip.

(c) Dérivées par rapport au focus.

FIGURE E.5 – Images d'un analyseur de courbure et exemples de dérivées.

E.6 Filtre de Zernike

Le filtre de Zernike est une pastille de taille λ/D , avec D le diamètre de la pupille, qui introduit un déphasage de $\pi/2$ et se place au foyer. Lorsque l'on fait ensuite une image en plan pupille, on obtient un signal proportionnel à la phase en entrée, à une constante près.

Cet effet est valable lorsque la phase est faible. L'intensité dans le plan focal, en approximation de faible phase, s'exprime ainsi (cf 1.4 p.46) :

$$\mathbf{I}(\mathbf{A}) = \mathbf{P}^2(1 + 2\varphi) = \mathbf{P}^2(1 + 2\mathbf{Z}\mathbf{A}) \quad (\text{E.15})$$

Pour calculer la dérivée d'une image par rapport à a_j , nous allons réexprimer $\mathbf{I}(\mathbf{A})$ sans approximation :

$$\mathbf{I}(\mathbf{A}) = |\text{TF}^{-1}\{\text{TF}[\mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\}|^2 \quad (\text{E.16})$$

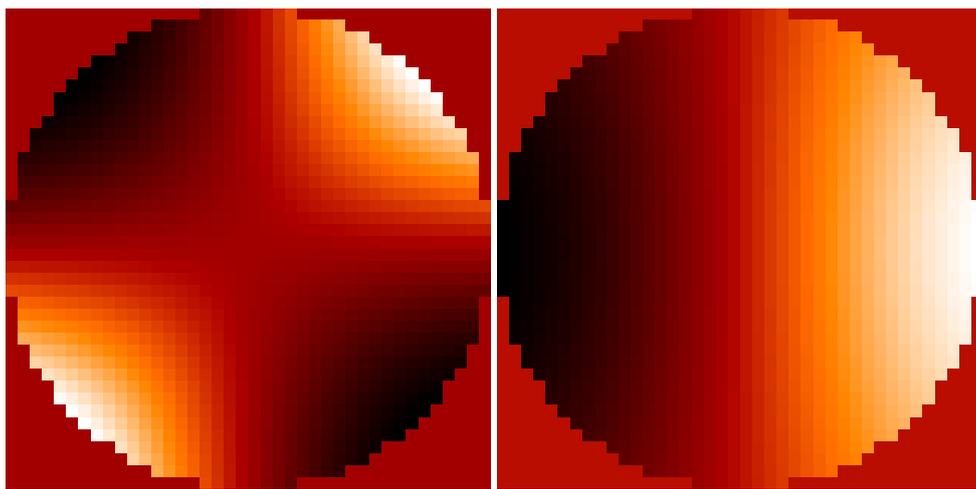
avec $\mathbf{M}_{\mathbf{Z}}$ le filtre de Zernike. On a alors :

$$\begin{aligned} \frac{\partial \mathbf{I}(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} &= \text{TF}^{-1}\{\text{TF}[i\mathbf{Z}_k \mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\} \times \text{TF}^{-1}\{\text{TF}[\mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\}^* + \\ &\quad \text{TF}^{-1}\{\text{TF}[\mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\} \times \text{TF}^{-1}\{\text{TF}[i\mathbf{Z}_k \mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\}^* \end{aligned} \quad (\text{E.17})$$

D'où :

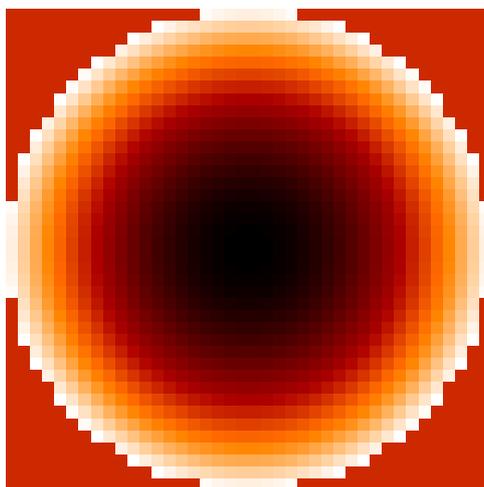
$$\frac{\partial \mathbf{I}(\mathbf{A})}{\partial a_j} \Big|_{\mathbf{A}=\mathbf{A}_p} = 2\Re \left\{ \text{TF}^{-1}\{\text{TF}[i\mathbf{Z}_k \mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\} \times \text{TF}^{-1}\{\text{TF}[\mathbf{P} \exp(i\varphi)] \times \mathbf{M}_{\mathbf{Z}}\}^* \right\} \quad (\text{E.18})$$

Les dérivées pour un tip et un focus sont montrées dans les figures E.6b et E.6c. On remarque qu'on voit directement un défaut de phase correspondant à un tip ou à un focus, ce qui est cohérent avec une dérivation de l'équation E.15.



(a) Image d'un analyseur à filtre de Zernike pour un astigmatisme de 0.1 rad. Une image de référence sans aberrations a été soustraite pour enlever la partie constante.

(b) Dérivée par rapport au tip.



(c) Dérivée par rapport au focus.

FIGURE E.6 – Image d'un analyseur à filtre de Zernike et exemples de dérivées.

Liste des publications

Revue à comité de lecture

C. Plantet, S. Meimon, J.-M. Conan, T. Fusco, *Experimental validation of LIFT for estimation of low-order modes in low-flux wavefront sensing*, Opt. Exp., 2013

S. Meimon, T. Fusco, V. Michau, **C. Plantet**, *Sensing more modes with less sub-apertures : the LIFTed Shack-Hartmann Wavefront sensor*, Opt. Lett., 2014

C. Plantet, S. Meimon, J.-M. Conan, T. Fusco, *Revisiting the comparison between the Shack-Hartmann and the pyramid wavefront sensor via the Fisher information matrix*, Opt. Exp., en cours de soumission

Actes de conférence

C. Plantet, S. Meimon, T. Fusco, J.-M. Conan, *Experimental validation of the linearized focal-plane technique (LIFT)*, Proc. AO4ELT 2, 2011

C. Plantet, B. Neichel, S. Meimon, T. Fusco, J.-M. Conan, *LIFT, a noise-effective low order focal-plane sensor : from theory to full experimental validation*, Proc. SPIE, 2012

C. Plantet, S. Meimon, J.-M. Conan, B. Neichel, T. Fusco, *On-sky validation of LIFT on GeMS*, Proc. AO4ELT 3, 2013

C. Plantet, S. Meimon, J.-M. Conan, B. Neichel, T. Fusco, *LIFT : analysis of performance in a laser assisted adaptive optics*, Proc. SPIE, 2014

Bibliographie

- [Baranec *et al.*, 2006] BARANEC, C., LLOYD-HART, M., MILTON, N. M., STALCUP, T., SNYDER, M. et ANGEL, R. (2006). Tomographic reconstruction of stellar wavefronts from multiple laser guide stars. *In Astronomical Telescopes and Instrumentation*, pages 627203–627203. International Society for Optics and Photonics.
- [Barrett *et al.*, 1995] BARRETT, H. H., DENNY, J., WAGNER, R. F. et MYERS, K. J. (1995). Objective assessment of image quality. ii. fisher information, fourier crosstalk, and figures of merit for task performance. *JOSA A*, 12(5):834–852.
- [Barwick, 2009] BARWICK, S. (2009). Performance comparison between shack-hartmann and astigmatic hybrid wavefront sensors. *Applied optics*, 48(36):6967–6972.
- [Bay *et al.*, 2008] BAY, S., HERZET, C., BROSSIER, J., BARBOT, J.-P. et GELLER, B. (2008). Analytic and asymptotic analysis of bayesian cramer–rao bound for dynamical phase offset estimation. *Signal Processing, IEEE Transactions on*, 56(1):61–70.
- [Cannon, 1995] CANNON, R. C. (1995). Global wave-front reconstruction using shack–hartmann sensors. *JOSA A*, 12(9):2031–2039.
- [Cannon, 1996] CANNON, R. C. (1996). Optimal bases for wave-front simulation and reconstruction on annular apertures. *JOSA A*, 13(4):862–867.
- [Conan, 1994] CONAN, J.-M. (1994). *Etude de la correction partielle en optique adaptative*. Thèse de doctorat, Paris XI Orsay.
- [Conan *et al.*, 1995] CONAN, J.-M., ROUSSET, G. et MADEC, P.-Y. (1995). Wave-front temporal spectra in high-resolution imaging through turbulence. *JOSA A*, 12(7):1559–1570.
- [Cramér, 1999] CRAMÉR, H. (1999). *Mathematical methods of statistics*, volume 9. Princeton university press.
- [Dauwels et Korl, 2006] DAUWELS, J. et KORL, S. (2006). A numerical method to compute cramér-rao-type bounds for challenging estimation problems. *In Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–V. IEEE.
- [Dohlen *et al.*, 2006] DOHLEN, K., LANGLOIS, M., LANZONI, P., MAZZANTI, S., VIGAN, A., MONTOYA, L., HERNANDEZ, E., REYES, M., SURDEJ, I. et YAITSKOVA, N. (2006). Zeus : a cophasing sensor based on the zernike phase contrast method. *In Astronomical Telescopes and Instrumentation*, pages 626734–626734. International Society for Optics and Photonics.
- [Ellerbroek et Rigaut, 2001] ELLERBROEK, B. L. et RIGAUT, F. (2001). Methods for correcting tilt anisoplanatism in laser-guide-star-based multiconjugate adaptive optics. *JOSA A*, 18(10):2539–2547.

- [Ellerbroek *et al.*, 1997] ELLERBROEK, B. L., THELEN, B. J., LEE, D. J., CARRARA, D. A. et PAXMAN, R. G. (1997). Comparison of shack-hartmann wavefront sensing and phase-diverse phase retrieval. *In Optical Science, Engineering and Instrumentation'97*, pages 307–320. International Society for Optics and Photonics.
- [Fienup *et al.*, 1998] FIENUP, J. R., THELEN, B. J., PAXMAN, R. G. et CARRARA, D. A. (1998). Comparison of phase diversity and curvature wavefront sensing. *In Astronomical Telescopes & Instrumentation*, pages 930–940. International Society for Optics and Photonics.
- [Foy et Labeyrie, 1985] FOY, R. et LABEYRIE, A. (1985). Letter to the editor feasibility of adaptive telescope with laser probe. *Astron. Astrophys*, 152:L29–L31.
- [Fusco *et al.*, 1999] FUSCO, T., CONAN, J.-M., MICHAU, V., MUGNIER, L. M. et ROUSSET, G. (1999). Phase estimation for large field of view : application to multiconjugate adaptive optics. *In SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 125–133. International Society for Optics and Photonics.
- [Fusco *et al.*, 2001] FUSCO, T., CONAN, J.-M., ROUSSET, G., MUGNIER, L. M. et MICHAU, V. (2001). Optimal wave-front reconstruction strategies for multiconjugate adaptive optics. *JOSA A*, 18(10):2527–2538.
- [Gonsalves, 1982] GONSALVES, R. A. (1982). Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21(5).
- [Gonsalves, 2001] GONSALVES, R. A. (2001). Small-phase solution to the phase-retrieval problem. *Optics letters*, 26(10).
- [Guyon, 2005] GUYON, O. (2005). Limits of adaptive optics for high-contrast imaging. *The Astrophysical Journal*, 629(1):592.
- [Guyon *et al.*, 2008] GUYON, O., BLAIN, C., TAKAMI, H., HAYANO, Y., HATTORI, M. et WATANABE, M. (2008). Improving the sensitivity of astronomical curvature wavefront sensor using dual-stroke curvature. *Publications of the Astronomical Society of the Pacific*, 120(868):655–664.
- [Herriot et Irvin, 2012] HERRIOT, G. et IRVIN, C. (2012). Mitigation of transient meteor events in sodium layer by tmt nfiraos. *In SPIE Astronomical Telescopes+ Instrumentation*, pages 844734–844734. International Society for Optics and Photonics.
- [Johnston et Welsh, 1994] JOHNSTON, D. C. et WELSH, B. M. (1994). Analysis of multiconjugate adaptive optics. *JOSA A*, 11(1):394–408.
- [Lane et Tallon, 1992] LANE, R. et TALLON, M. (1992). Wave-front reconstruction using a shackhartmann sensor. *Applied optics*, 31(32):6902–6908.
- [Le Louarn *et al.*, 2000] LE LOUARN, M., HUBIN, N., SARAZIN, M. et TOKOVININ, A. (2000). New challenges for adaptive optics : extremely large telescopes. *Monthly notices of the Royal Astronomical society*, 317(3):535–544.
- [Lee *et al.*, 1997] LEE, D. J., WELSH, B. M. et ROGGEMANN, M. C. (1997). Cramer-rao analysis of phase diversity imaging. *In Optical Science, Engineering and Instrumentation'97*, pages 161–172. International Society for Optics and Photonics.
- [Machado *et al.*, 2009] MACHADO, V. C., TAPIA, G., GABRIEL, D., LAFUENTE, J. et BAEZA, J. A. (2009). Systematic identifiability study based on the fisher information matrix for

- reducing the number of parameters calibration of an activated sludge model. *Environmental Modelling & Software*, 24(11):1274–1284.
- [McGregor *et al.*, 2004] MCGREGOR, P. *et al.* (2004). Gemini South Adaptive Optics Imager (GSAOI). *Proc. SPIE*, 5472:1033–1044.
- [Meimon *et al.*, 2010a] MEIMON, S., FUSCO, T. et CASSAING, F. (2010a). A focal plane sensor for low-order sensing on laser tomographic systems : Lift. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 773611–773611. International Society for Optics and Photonics.
- [Meimon *et al.*, 2010b] MEIMON, S., FUSCO, T. et MUGNIER, L. M. (2010b). Lift : a focal-plane wavefront sensor for real-time low-order sensing on faint sources. *Optics letters*, 35(18):3036–3038.
- [Meynadier *et al.*, 1999] 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. *Applied optics*, 38(23):4967–4979.
- [Mocœur *et al.*, 2009] MOCŒUR, I., MUGNIER, L. M. et CASSAING, F. (2009). Analytical solution to the phase-diversity problem for real-time wavefront sensing. *Opt. Lett.*, 34(22): 3487–3489.
- [Morris *et al.*, 2014] MORRIS, T., GENDRON, E., BASDEN, A., MARTIN, O., OSBORN, J., HENRY, D., HUBERT, Z., SIVO, G., GRATADOUR, D., CHEMLA, F. *et al.* (2014). Canary phase b : on-sky open-loop tomographic lgs ao results. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 91481I–91481I. International Society for Optics and Photonics.
- [Mugnier *et al.*, 2004] MUGNIER, L. M., FUSCO, T. et CONAN, J.-M. (2004). Mistral : a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *J. Opt. Soc. Am. A*, 21(10):1841–1854.
- [Neichel, 2008] NEICHEL, B. (2008). *Etude des galaxies lointaines et optiques adaptatives tomographiques pour les ELTs*. Thèse de doctorat, Paris VII Denis Diderot.
- [Neichel *et al.*, 2010] NEICHEL, B. *et al.* (2010). The Gemini MCAO System GeMS : nearing the end of a lab-story. *Proc. SPIE*, 7736(773606).
- [Neichel *et al.*, 2013] NEICHEL, B., D’ORGEVILLE, C., CALLINGHAM, J., RIGAUT, F., WINGE, C. et TRANCHO, G. (2013). Characterization of the sodium layer at cerro pachón, and impact on laser guide star performance. *Monthly Notices of the Royal Astronomical Society*, 429(4):3522–3532.
- [Neichel *et al.*, 2009] NEICHEL, B., FUSCO, T. et CONAN, J.-M. (2009). Tomographic reconstruction for wide-field adaptive optics systems : Fourier domain analysis and fundamental limitations. *JOSA A*, 26(1):219–235.
- [Nicolle, 2006] NICOLLE, M. (2006). *Analyse de front d’onde pour les optiques adaptatives de nouvelle génération : optiques adaptatives à large champ et optique adaptative extrême*. Thèse de doctorat, Paris XI Orsay.
- [Nicolle *et al.*, 2004] NICOLLE, M., FUSCO, T., ROUSSET, G. et MICHAU, V. (2004). Improvement of shack-hartmann wave-front sensor measurement for extreme adaptive optics. *Optics letters*, 29(23):2743–2745.

- [Paterson, 2008] PATERSON, C. (2008). Towards practical wavefront sensing at the fundamental information limit. In *Journal of Physics : Conference Series*, volume 139, page 012021. IOP Publishing.
- [Paterson et Dainty, 2000] PATERSON, C. et DAINTY, J. (2000). Hybrid curvature and gradient wave-front sensor. *Optics letters*, 25(23):1687–1689.
- [Paxman et al., 1992] PAXMAN, R. G., SCHULZ, T. J. et FIENUP, J. R. (1992). Joint estimation of object and aberrations by using phase diversity. *JOSA A*, 9(7):1072–1085.
- [Pfrommer et Hickson, 2010] PFROMMER, T. et HICKSON, P. (2010). High-resolution lidar observations of mesospheric sodium and implications for adaptive optics. *JOSA A*, 27(11): A97–A105.
- [Poyneer et Macintosh, 2004] POYNEER, L. A. et MACINTOSH, B. (2004). Spatially filtered wave-front sensor for high-order adaptive optics. *JOSA A*, 21(5):810–819.
- [Ragazzoni, 1996] RAGAZZONI, R. (1996). Pupil plane wavefront sensing with an oscillating prism. *Journal of modern optics*, 43(2):289–293.
- [Ragazzoni et Farinato, 1999] RAGAZZONI, R. et FARINATO, J. (1999). Sensitivity of a pyramidal wave front sensor in closed loop adaptive optics. *Astronomy and Astrophysics*, 350:L23–L26.
- [Retout et al., 2001] RETOUT, S., DUFFULL, S. et MENTRÉ, F. (2001). Development and implementation of the population fisher information matrix for the evaluation of population pharmacokinetic designs. *Computer methods and Programs in Biomedicine*, 65(2):141–151.
- [Rigaut et Gendron, 1992] RIGAUT, F. et GENDRON, E. (1992). Laser guide star in adaptive optics-the tilt determination problem. *Astronomy and Astrophysics*, 261:677–684.
- [Rigaut et al., 2013] RIGAUT, F., NEICHEL, B., BOCCAS, M., D'ORGEVILLE, C., VIDAL, F., van DAM, M. A., ARRIAGADA, G., FESQUET, V., GALVEZ, R. L., GAUSACHS, G. et al. (2013). Gemini multiconjugate adaptive optics system review–i. design, trade-offs and integration. *Monthly Notices of the Royal Astronomical Society*, page stt2054.
- [Roddiier, 1988] RODDIER, F. (1988). Curvature sensing and compensation : a new concept in adaptive optics. *Applied Optics*, 27(7):1223–1225.
- [Roddiier, 1999] RODDIER, F. (1999). *Adaptive optics in astronomy*. Cambridge university press.
- [Roddiier, 1990] RODDIER, N. A. (1990). Atmospheric wavefront simulation using zernike polynomials. *Optical Engineering*, 29(10):1174–1180.
- [Rousset, 1999] ROUSSET, G. (1999). Wave-front sensors. In RODDIER, F., éditeur : *Adaptive Optics in Astronomy*, chapitre 5, pages 91–130.
- [Schulz et al., 1999] SCHULZ, T. J., SUN, W. et ROGGEMANN, M. C. (1999). Cramer-rao bounds for estimation of turbulence-induced wavefront aberrations. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 23–28. International Society for Optics and Photonics.
- [Shack et Platt, 1971] SHACK, R. V. et PLATT, B. (1971). Production and use of a lenticular hartmann screen. In *Journal of the Optical Society of America*, volume 61, page 656.
- [Smith et al., 2013] SMITH, C., MARINICĂ, R., DEN DEKKER, A., VERHAEGEN, M., KOR-KIAKOSKI, V., KELLER, C. et DOELMAN, N. (2013). Iterative linear focal-plane wavefront correction. *JOSA A*, 30(10):2002–2011.

- [Tallon et Foy, 1990] TALLON, M. et FOY, R. (1990). Adaptive telescope with laser probe-isoplanatism and cone effect. *Astronomy and Astrophysics*, 235:549–557.
- [Tallon *et al.*, 1992] TALLON, M., FOY, R. et VERMIN, J. (1992). 3-d wavefront sensing for multiconjugate adaptive optics. In *European Southern Observatory Conference and Workshop Proceedings*, volume 42, page 517.
- [Thomas *et al.*, 2006] THOMAS, S., FUSCO, T., TOKOVININ, A., NICOLLE, M., MICHAU, V. et ROUSSET, G. (2006). Comparison of centroid computation algorithms in a shack-hartmann sensor. *Monthly Notices of the Royal Astronomical Society*, 371(1):323–336.
- [Tokovinin et Heathcote, 2006] TOKOVININ, A. et HEATHCOTE, S. (2006). Donut : measuring optical aberrations from a single extrafocal image. *Publications of the Astronomical Society of the Pacific*, 118(846):1165–1175.
- [Vallisneri, 2008] VALLISNERI, M. (2008). Use and abuse of the fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Physical Review D*, 77(4):042001.
- [Van Trees, 2004] VAN TREES, H. L. (2004). *Detection, estimation, and modulation theory*. John Wiley & Sons.
- [Vérinaud, 2004] VÉRINAUD, C. (2004). On the nature of the measurements provided by a pyramid wave-front sensor. *Optics Communications*, 233(1):27–38.
- [Welsh *et al.*, 1995] WELSH, B. M., ELLERBROEK, B. L., ROGGEMANN, M. C. et PENNINGTON, T. L. (1995). Fundamental performance comparison of a hartmann and a shearing interferometer wave-front sensor. *Applied optics*, 34(21):4186–4195.
- [Wynne, 1979] WYNNE, C. (1979). Extending the bandwidth of speckle interferometry. *Optics Communications*, 28(1):21–25.
- [Young *et al.*, 2012] YOUNG, P., GREGOR, P. M., van HARMELENA, J. et NEICHEL, B. (2012). Using ODGWs with GSAOI : Software and firmware implementation challenges. In ELLERBROEK, B., MARCHETTI, E. et VERAN, J.-P., éditeurs : *Adaptive Optics Systems III*. Proc. SPIE 8451.
- [Zernike, 1934] ZERNIKE, F. (1934). Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. *Monthly Notices of the Royal Astronomical Society*, 94: 377–384.

Etude d'un nouveau concept d'analyse de front d'onde en plan focal appliqué à l'optique adaptative en astronomie

L'avènement des Extremely Large Telescopes (ELT) est imminent. Cette nouvelle génération de télescopes ouvrira la voie vers des observations astronomiques toujours plus fines. Les futurs instruments nécessitent des optiques adaptatives de plus en plus sophistiquées pour corriger les effets de la turbulence et atteindre les performances requises. L'analyse de front d'onde est un élément fondamental de l'optique adaptative. Les analyseurs requièrent un flux relativement élevé pour être efficaces, notamment lors de l'estimation de hauts ordres, qui sont nécessaires pour obtenir une bonne correction. Comme le ciel est pauvre en étoiles brillantes, on utilise souvent des étoiles laser, produites par excitation d'un volume de sodium dans la haute atmosphère. Malheureusement, on ne peut pas estimer correctement un tip/tilt avec ces étoiles artificielles. De plus, le focus mesuré est biaisé à cause de l'instabilité en altitude de la couche de sodium. La mesure de ces trois aberrations doit donc pouvoir être faite sur une étoile naturelle de très faible luminosité. On peut alors se demander quelle stratégie adopter pour l'analyse des bas ordres à faible flux. De manière plus générale, les techniques d'analyse de front d'onde sont en pleine évolution pour répondre aux nouveaux besoins de l'optique adaptative. Il est donc important de comparer les stratégies d'analyse en terme de sensibilité et de dynamique. La thèse s'appuie sur un nouveau concept d'analyseur de surface d'onde : LIFT (LInearized Focal-plane Technique). Cet analyseur permettra d'estimer le tip/tilt et le focus sur une étoile naturelle faible, afin de compléter l'analyse faite sur les étoiles lasers, de façon plus efficace qu'avec les moyens actuels. Le premier but de la thèse est d'optimiser et de valider LIFT expérimentalement. Pour cela, j'ai d'abord étudié les propriétés de convergence de l'algorithme et la propagation du bruit afin de déterminer les paramètres optimaux. J'ai ensuite testé le concept de LIFT en laboratoire sur des phases statiques, puis sur le ciel grâce au système d'optique adaptative du télescope Gemini Sud, GeMS. La linéarité de l'estimation malgré la présence de hauts ordres résiduels, dus à la correction imparfaite de l'optique adaptative, a été validée en boucle ouverte, et nous pouvons maintenant envisager une validation finale en intégrant LIFT dans une boucle fermée. Dans un deuxième temps, j'ai étudié la propagation du bruit dans un analyseur dérivé de LIFT, le Shack-Hartmann LIFTé. Cet analyseur divise la pupille à l'aide de micro-lentilles astigmatiques, et une estimation par LIFT est faite dans chaque sous-pupille. Il permettra une estimation de hauts ordres plus précise qu'avec un Shack-Hartmann classique. Le dernier objectif est de comparer les analyseurs existants, en termes de propagation de bruit, pour l'estimation sur étoile naturelle des bas ordres d'une part, cas d'application de LIFT, et des hauts ordres d'autre part, dans le cas d'une optique adaptative extrême par exemple. Cette étude permet de motiver l'utilisation de LIFT et du Shack-Hartmann LIFTé dans des systèmes d'optique adaptative.

Mots-clés : OPTIQUE ADAPTATIVE ; ANALYSE FRONT ONDE

Study of a new focal plane wavefront sensing concept for adaptive optics in astronomy

The upcoming generation of telescopes, the Extremely Large Telescopes (ELT), will open the way towards finer astronomical observations. These instruments will require advanced adaptive optics systems in order to correct the effects of atmospheric turbulence and reach the performance specifications. Wavefront sensing is a fundamental element of adaptive optics. Wavefront sensors need a relatively high flux to be effective, especially when estimating high order aberrations, which are essential for a good correction. The sky being poor in bright stars, laser guide stars, created by exciting a volume of sodium atoms in the upper atmosphere, are often used. Unfortunately, the tip/tilt cannot be estimated from these artificial stars. Moreover, the focus measurement is biased because of the instability of the sodium layer's altitude. Therefore, the estimation of these three aberrations must be made on a faint natural star. A strategy for sensing low orders at low flux has thus to be defined. More generally, wavefront sensing techniques are evolving to answer the new needs of adaptive optics. It is then important to compare sensing strategies in terms of sensitivity and dynamics. This thesis has been initiated by a new wavefront sensor concept: LIFT (LInearized Focal-plane Technique). This sensor will be able to estimate tip/tilt and focus on a faint natural star, in order to complement the analysis on laser stars, more efficiently than the current means. The first goal of the thesis is the optimization of LIFT and its experimental validation. To do this, I first studied the algorithm's convergence properties and its noise propagation in order to determine the optimal parameters. I then validated LIFT in laboratory on static phases, in the absence of adaptive optics and turbulence. Once these first tests finished, we tested LIFT on sky thanks to the adaptive optics system of the Gemini South telescope, GeMS. The linearity of LIFT's estimation despite the presence of high order residuals, due to the imperfect correction of adaptive optics, has been confirmed in open loop, and we are now planning a final validation by integrating LIFT in a closed loop. In a second phase, I studied the noise propagation of a wavefront sensor derived from LIFT, the LIFTed Shack-Hartmann. This sensor uses astigmatic lenslets to divide the pupil, and measures the wavefront in each subaperture with LIFT. It will allow a more efficient estimation of high orders than a regular Shack-Hartmann. The final objective is to compare existing wavefront sensors' noise propagations for the estimation on a natural star of low orders on one hand, which corresponds to LIFT's domain of application and of high orders on the other hand, for example in an extreme adaptive optics application. This study motivates the use of LIFT and the LIFTed Shack-Hartmann in adaptive optics systems.

Keywords : ADAPTIVE OPTICS ; WAVEFRONT SENSING