

# Reconstruction d'images astronomiques en interférométrie optique

Serge Meimon

▶ **To cite this version:**

Serge Meimon. Reconstruction d'images astronomiques en interférométrie optique. Astrophysique [astro-ph]. Université Paris Sud - Paris XI, 2005. Français. NNT : . tel-00011355

**HAL Id: tel-00011355**
**https://theses.hal.science/tel-00011355**

Submitted on 11 Jan 2006

**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.

UNIVERSITÉ PARIS SUD

THESE DE DOCTORAT

présentée pour obtenir le grade de

DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS SUD

par

**Serge MEIMON**

# RECONSTRUCTION D'IMAGES ASTRONOMIQUES EN INTERFÉROMETRIE OPTIQUE

Soutenue le 21 Novembre 2005 devant le jury composé de

| | |
|---|---|
| **M. Guy Demoment** | Directeur de thèse |
| **Mme. Sylvie Roques** | Rapporteur |
| **M. Peter Lawson** | Rapporteur |
| **M. Laurent Mugnier** | Co-directeur de thèse |
| **M. Guy Perrin** | Membre du jury |
| **M. Eric Thiébaut** | Membre du jury |
| **M. Guy Le Besnerais** | (Invité) Co-encadrant de thèse |

# Remerciements

*Ceux qui sont partis de rien pour arriver à pas grand chose n'ont de merci à dire à personne.*

Pierre Dac

**Les classiques**

Je tiens tout d'abord à remercier Guy Demoment, Laurent Mugnier et Guy Le Besnerais qui, malgré mon lourd passé d'ingénieur automaticien, m'ont permis d'effectuer une thèse en optique. En plus de leur confiance, ils m'ont fait cadeau, entre autres, de quelques cours d'inférence bayésienne dont la pédagogie m'a permis de saisir l'attrait de la théorie de l'estimation à laquelle j'étais jusque là hermétique ; d'une formation accélérée en Unix, Linux, IDL, XFIG, EMACS, et autres noms d'antibiotiques ; d'illustrations artistiques en marge de premiers jets laborieux, et de tirés à part de dessins de Charlie ou du Canard.

Je tiens aussi à remercier vivement Peter Lawson et Sylvie Roques qui ont eu à juger ce travail dans de très brefs délais.

J'ai trouvé en la personne d'Eric Thiébaut un amateur impénitent de grillades et de mechouïa[1], un acharné des dérivées matricielles et de Lagrangiens, et un interlocuteur disponible et méritant.

Je suis très reconnaissant envers Guy Perrin qui, non content de m'ouvrir la porte du territoire sacré du mont Hopkins, m'a montré comment on fête un anniversaire en Arizona, comment on fait la danse de la pluie, et comment on reconstruit de tête une image astronomique. Merci aussi à Sylvestre Lacour pour les joies du rallye automobile en montagne, à la recherche d'azote liquide, et pour nos discussions fructueuses et spiritueuses au lever du soleil après huit heures d'observation.

Enfin, je remercie mes collègues de l'ONERA, thésards et permanents, qui m'ont donné envie de faire de la recherche et m'ont convaincu de continuer à la faire avec eux.

**Les sucrées**

Je remercie mon père, ma mère, mon frère ma soeur et mon cousin, pour leur soutien sans failles, leur affection sans défauts, leur sens de la dérision (y compris au sujet de ma thèse), et

[1] Sur un barbecue, après les grillades, oubliez huit tomates, deux poivrons verts, un poivron rouge, un petit piment rouge et trois gousses d'ail, laissez refroidir dans un sac plastique une fois que toutes les faces sont brûlées, hachez finement avec deux couteaux après avoir retiré la peau brûlée, salez, laissez passer une nuit, disposez sur un plat peu profond, ajoutez un filet d'huile d'olive et le jus d'un citron, une boite de thon "Petit-Joli" et deux oeufs durs en quartiers. Merci à ma grand-mère pour la recette.

leur capacité à me rappeler ce qui est essentiel dans la vie : bien manger, bien boire, le tout en bonne compagnie, et surtout le dimanche midi. Merci à ma grand-mère de ne jamais m'avoir laissé mourir de faim, loin s'en faut.

Merci à mes amis et colocataires de m'avoir supporté (sans anglicisme) durant ces trois années. Je remercie en particulier Olivier Macherey et Christophe Mazeaud, dont l'amitié est allée jusqu'à m'aider à rédiger ce manuscrit, et à Ivan Favéro, qui a en même temps que moi découvert que non, décidément, la thèse n'est pas un long fleuve tranquille.

**Les salées**

Je remercie le conseiller d'orientation du Lycée Voltaire qui m'a affirmé que mon avenir était dans la chaudronnerie, le professeur de mathématiques de première S qui m'a suggéré une filière littéraire, les professeurs de mathématiques et physique de math. sup. qui m'ont enjoint de trouver une "porte de sortie" et les spécialistes en projets professionnels et tests de personnalité de l'École Centrale de Nantes qui me voyaient en publicitaire ou en manager de l'an 2000, grâce soit rendue à leur clairvoyance.

**Les pas comme partout**

Je remercie certain chauffard ivre et manchot de m'avoir remis les idées en place à coup de peugeot 206.

Ma soutenance de thèse doit beaucoup à Josquin Dehaene et Stanislas Bernard, eux qui parlent sur le bout des doigts le langage des disques durs et qui ont fait qu'une panne de portable six jours avant la soutenance n'a pas été une catastrophe.

Merci à Vincent Josso, auteur de la charte graphique des transparents de ma soutenance (modèle déposé), et contributeur du désormais fameux "théorème de Bix" (section 3.2.1.2 page 77).

# Table des matières

**Introduction** **5**

**Notations** **9**

**1 Synthèse d'ouverture optique et Haute résolution angulaire.** **11**
1.1 L'expérience des trous d'Young . . . 11
1.1.1 L'expérience de 1804 . . . 11
1.1.2 Diffraction de Fraunhofer à l'infini . . . 12
1.1.3 Processus d'imagerie . . . 13
1.1.4 Retour à l'expérience des trous d'Young . . . 15
1.1.5 Formation d'interférogrammes par deux télescopes . . . 18
1.2 La mesure interférométrique optique en astronomie . . . 23
1.2.1 Interféromètres au sol . . . 23
1.2.2 Turbulence et mesure de phase . . . 25
1.2.3 Clôtures de phase ou phases de bispectre . . . 26
1.2.4 Méthodes d'estimation des phases de visibilité à partir des clôtures . . . 27
1.2.5 Mise en équations des mesures interférométriques . . . 28
1.3 Méthodes classiques de traitement de données interférométriques . . . 31
1.3.1 Sous-détermination du problème . . . 31
1.3.2 Méthodes paramétriques . . . 31
1.3.3 Pondération et « gridding » en radio-astronomie . . . 32
1.3.4 Une méthode de référence en reconstruction interférométrique : Clean . . . 32
1.3.5 Méthode du maximum d'entropie (MEM) . . . 34
1.4 Inférence bayésienne . . . 35
1.4.1 Base théorique . . . 35
1.4.2 Lois *a priori* classiques . . . 37

**2 Reconstruction d'images à partir de visibilités complexes.** **43**
2.1 Position du problème . . . 43
2.1.1 Démarche . . . 43
2.2 Construction du critère d'attache aux données . . . 44
2.2.1 Modèle de bruit . . . 44
2.2.2 Critère d'attache aux données exact $\mathcal{J}^{data}$ . . . 44
2.2.3 $\mathcal{J}^{\mathrm{data}}$ n'est pas convexe . . . 45
2.3 Approximation convexe quadratique du critère d'attache aux données . . . 48
2.3.1 Modèle de données en radio-interférométrie . . . 48
2.3.2 Choisir la « meilleure » approximation gaussienne . . . 50
2.3.3 Le critère $\mathcal{J}_{ell}$ . . . 53
2.4 L'algorithme RIME : Reconstruction d'Images par Modélisation Elliptique . . . 57
2.4.1 Construction d'un modèle *a priori* . . . 57

2.4.2 Une méthode de descente locale utilisant le calcul du gradient . . . . . 57
2.4.3 Une reconstruction non supervisée . . . . . 58
2.5 Validation sur simulations de RIME . . . . . 62
2.5.1 Données simulées . . . . . 62
2.5.2 Modèle de DSP . . . . . 63
2.5.3 Comparaison des approximations elliptiques et circulaire . . . . . 63
2.5.4 Influence de l'hyper-paramètre . . . . . 65
2.5.5 Conclusion des simulations . . . . . 65
2.6 Traitement des données issues de l'interféromètre fibré de l'IRCOM . . . . . 67
2.6.1 Présentation du dispositif expérimental . . . . . 67
2.6.2 Justification du modèle de bruit proposé . . . . . 67
2.6.3 Analyse des reconstructions obtenues . . . . . 69
2.7 Conclusion . . . . . 70

**3 Modèle direct des données affectées par la turbulence 73**
3.1 Des visibilités carrées aux visibilités en module . . . . . 74
3.1.1 Pourquoi mesure-t-on des visibilités carrées ? . . . . . 74
3.1.2 Déduire un modèle de données sur les amplitudes de visibilité . . . . . 74
3.2 Structures de l'information de phase dans le cas fortement turbulent . . . . . 77
3.2.1 Informations de phase et noyau de $\boldsymbol{B}^{\mathrm{T}}$ . . . . . 77
3.2.2 Informations manquantes, pistons turbulents et rang de $\boldsymbol{B}$ . . . . . 79
3.3 Modèle de données et critère d'attache aux données exact . . . . . 80
3.3.1 Critère d'attache aux données sur les visibilités carrées et les clôtures . 80
3.3.2 Terme d'amplitude de visibilités . . . . . 80
3.3.3 Dégénérescences du problème . . . . . 81
3.4 Modélisation myope des données . . . . . 82
3.4.1 Informations manquantes et aberrations du système . . . . . 82
3.4.2 Ambiguïtés entières . . . . . 85
3.4.3 Construction du modèle myope . . . . . 87
3.4.4 Application à une couverture fréquentielle redondante avec un interféromètre à 3 télescopes . . . . . 88
3.4.5 Un autre traitement des données IRCOM . . . . . 90
3.5 Bilan : Un modèle myope de données sur des phases et des amplitudes . . . . . 91

**4 Reconstructions à partir de visibilités carrées et de clôtures 93**
4.1 Méthodes de reconstruction d'image en interférométrie à clôtures . . . . . 94
4.1.1 Inférence bayésienne classique . . . . . 94
4.1.2 Méthodes d'auto-calibration en radio-astronomie . . . . . 95
4.1.3 Adapter ces méthodes au domaine Optique . . . . . 98
4.2 Le critère joint approché $\mathcal{J}_{WISARD}$ . . . . . 98
4.2.1 Construction du critère joint approché $\mathcal{J}_{WISARD}$ . . . . . 98
4.2.2 Équivalence de $\mathcal{J}_{WISARD}$ et du critère exact . . . . . 99
4.2.3 Non convexité du critère $\mathcal{J}_{WISARD}$ . . . . . 101
4.2.4 Choix d'une stratégie de minimisation alternée . . . . . 102
4.3 Détermination des aberrations optimales par auto calibration . . . . . 103
4.3.1 Séparation du problème d'optimisation en aberrations . . . . . 103
4.3.2 Aberrations optimales dans le cas d'un interféromètre à 3 télescopes . . 104
4.3.3 Solution linéarisée en $\boldsymbol{\phi}_{\mathrm{res}}$ . . . . . 104
4.3.4 Solution linéarisée en $\boldsymbol{\phi}_{\mathrm{res}}$ et $\boldsymbol{a}_{\mathrm{res}}$ . . . . . 106

4.3.5 Solution itérative . . . 106
4.3.6 Une méthode robuste de détermination des aberrations . . . 107
4.4 L'algorithme WISARD : Weak-phase Interferometric Sample Alternating Reconstruction Device . . . 107
4.4.1 Objet d'initialisation . . . 108
4.4.2 Une reconstruction par fréquences croissantes . . . 108
4.4.3 Post-traitement . . . 110
4.4.4 Structure de l'algorithme WISARD . . . 110
4.5 Les versions successives de l'algorithme . . . 110
4.5.1 WISARD2 . . . 111
4.5.2 WISARD2 . . . 111
4.5.3 WISARD1 . . . 112
4.6 Conclusion . . . 113

**5 Valider WISARD par simulation puis expérimentalement 115**
5.1 Validation de WISARD sur simulations : « An Imaging Beauty Contest » . . . 115
5.1.1 Deux tests préliminaires . . . 115
5.1.2 Un concours de reconstruction myope en aveugle . . . 116
5.1.3 Amélioration des performances de l'algorithme . . . 120
5.1.4 Conclusions sur les simulations . . . 121
5.2 Préparation des missions sur l'interféromètre IOTA . . . 123
5.2.1 Déroulement de la première mission . . . 123
5.2.2 L'interféromètre IOTA . . . 123
5.2.3 Une méthode de choix de configuration pupillaire . . . 127
5.2.4 Étoiles de type Mira . . . 130
5.3 Reconstruction de $\chi$ du Cygne . . . 134
5.3.1 Réduction et première reconstruction des données . . . 134
5.3.2 Sources d'erreurs et adaptation de l'algorithme . . . 136
5.3.3 Reconstruction de $\chi$ du Cygne avec WISARD . . . 137
5.4 Conclusion . . . 139

**6 Conclusion 141**
6.1 Objectifs et résultats . . . 141
6.2 Discussion . . . 142
6.3 Perspectives . . . 143

**Bibliographie 145**

**A Approximation optimale par une loi gaussienne. 151**
A.1 Démonstration dans le cas général . . . 151
A.1.1 Définition de la distance de Kullback-Leibler . . . 151
A.1.2 Termes du premier ordre . . . 152
A.1.3 Termes du second ordre . . . 152
A.1.4 Cas particulier plan . . . 152
A.2 Application au contexte interférométrique optique . . . 153
A.2.1 Termes du premier ordre . . . 154
A.2.2 Termes du second ordre . . . 154
A.2.3 Bilan . . . 155

**B Clôtures de phase independantes.** **157**
B.1 Définitions . . . 157
B.2 Combinaisons linéaires de phases indépendantes annulant la turbulence . . . 158
B.3 Clôtures indépendantes à 3 télescopes . . . 158
B.4 Ces clôtures forment une base du noyau de $\ker(\boldsymbol{B}^{\mathrm{T}})$ . . . 159
B.5 Les opérateurs $\boldsymbol{C}$ et $\bar{\boldsymbol{B}}$ . . . 160

**C Optimisation quadratique sous contrainte de positivité** **161**
C.1 Théorème de Karush, Kuhn et Tucker . . . 161
C.1.1 Formulation générale d'un problème d'optimisation sous contraintes . . . 161
C.1.2 Caracterisation d'optimalité . . . 161
C.2 Application à l'optimisation quadratique . . . 162
C.2.1 Cas des contraintes égalités . . . 163
C.2.2 Cas des contraintes inégalités . . . 163
C.3 Application à la contrainte de positivité . . . 164
C.3.1 Pixels « gelés » . . . 164
C.3.2 Algorithme . . . 165
C.3.3 Cas non quadratique . . . 165

**D Imagerie et configuration pupillaire** **167**
D.1 L'operateur $\mathcal{H} : \Psi_{\boldsymbol{o}} \longrightarrow \Psi_{\boldsymbol{im}}$ est LIT . . . 167
D.2 L'operateur $\mathcal{G} : \boldsymbol{o} \longrightarrow \boldsymbol{im}$ est LIT . . . 167

**E Gradients de $\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha})$** **169**
E.1 Gradient de $\mathcal{J}_{WISARD}(\boldsymbol{x}|\boldsymbol{\alpha})$ par rapport à $\boldsymbol{x}$ . . . 169
E.2 Gradient de $\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})$ par rapport à $\boldsymbol{\alpha}$ . . . 169
E.2.1 Réécriture de $\mathcal{J}_{ell}$ . . . 170
E.2.2 Calcul du gradient de $\mathcal{J}^i_{ell}$ par rapport à $\phi_{\mathrm{res},i}$ . . . 171
E.2.3 Calcul du gradient de $\mathcal{J}_{ell}$ par rapport à $\boldsymbol{\phi}_{\mathrm{res}}$ . . . 171
E.2.4 Conclusion : gradient de $\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})$ par rapport à $\boldsymbol{\alpha}$ . . . 172

**F Publications.** **173**
F.1 JOSAA 2005 . . . 173
F.2 Optics Letters 2005 . . . 183

# Introduction

Les objectifs scientifiques des futures missions astronomiques, notamment la détection d'exoplanètes ou l'étude des noyaux actifs de galaxies, nécessitent des résolutions angulaires accrues - de l'ordre de la milliseconde d'arc - dans les longueurs d'onde Infra Rouge (IR) et Optique. Ces résolutions ne sont pas à la portée des télescopes actuellement en fonctionnement, dont les performances sont limitées par le diamètre de leur miroir primaire. Le pouvoir de résolution d'un télescope est proportionnel au diamètre de son miroir primaire. Or les technologies actuelles permettent de réaliser des télescopes au sol dont le diamètre n'excède pas 10 mètres environ, et les lanceurs spatiaux ne peuvent transporter que des instruments limités en poids et en encombrement.

Afin d'atteindre néanmoins les résolutions recherchées, il est possible de réaliser l'équivalent d'un télescope de grand diamètre, en recombinant interférométriquement les faisceaux de lumière issus de plusieurs télescopes (ou miroirs) de diamètres inférieurs. Il existe deux types d'instruments interférométriques. D'une part l'interféromètre de Fizeau : la lumière issue de chacun des miroirs élémentaires est renvoyée vers un miroir secondaire, puis vers le foyer de l'instrument où une image est formée. Ce miroir secondaire est situé au dessus des télescopes élémentaires, et serait donc à environ 100 mètres au dessus du sol si la distance maximale entre deux télescopes (ou ligne de base maximale) est de 100 mètres (Cf. Fig. 1 d'après Faucherre *et al.* (1989)), ce qui rend un tel instrument difficile à réaliser pour de grandes lignes de base. D'autre part l'interféromètre de Michelson : chaque instrument élémentaire est un télescope, dont la lumière est renvoyée vers un recombinateur faisant interférer les faisceaux entre eux. Ce design est plus compact, et permet la conception d'instruments de grande dimension. Toutefois, un tel instrument n'est pas un imageur. En effet, un interféromètre optique de Michelson fournit un jeu de figures d'interférences dont on extrait des données quantitatives, ou *observables*. Un traitement de données numérique est nécessaire pour estimer l'image à partir des quelques dizaines d'observables fournis par l'interféromètre. L'opération effectuée par un tel traitement de donnée, appelée *synthèse de Fourier* est particulièrement ardue, entre autres parce que le nombre de mesures - quelques dizaines à centaines d'observables - est très inférieur au nombre d'inconnues à déterminer - les pixels de l'image (typiquement plusieurs milliers).

La première application de l'interférométrie à l'observation de l'espace est due à Michelson, qui est parvenu à mesurer le diamètre des quatre principaux satellites de Jupiter en 1890. Malgré ce succès spectaculaire dans le domaine optique, l'interférométrie radio s'est développée bien plus vite, dès les années 40, bénéficiant des recherches militaires sur les radars. Les méthodes de traitement associées ont été développées à partir des années 70, permettant d'aller plus loin qu'une mesure de diamètre, et de reconstruire des images d'objets stellaires. L'algorithme CLEAN, proposé par Högbom en 1974 [Hogbom (1974)], est l'une des méthodes de reconstruction d'images les plus populaires en interférométrie radio. Parallèlement, les Méthodes à Maximum d'Entropie (MEM) [Nityananda & Narayan (1982)] tentent de reconstruire l'image la plus « lisse » compatible avec les observables.

Cette antériorité de la Radio-interférométrie peut sans doute être expliquée par le fait que les longueurs d'onde optique/IR sont environ 1000 fois plus petites que les longueurs d'onde

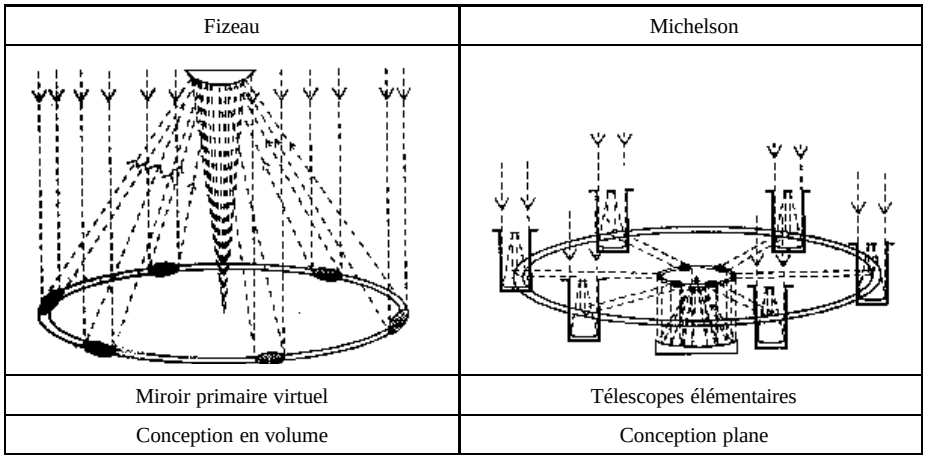



FIG. 1 – Deux types d'interféromètres optiques (d'après Faucherre *et al.* (1989)).

radio :

- La recombinaison des faisceaux, qui exige le contrôle des distances miroir-recombinateur à une fraction de longueur d'onde près, est bien plus ardue en optique –longueur d'onde de l'ordre d'un micron – qu'en radio – longueur d'onde de l'ordre du millimètre.
- De plus, les mesures optiques sont perturbées par les fluctuations très rapides (de l'ordre de 10ms) des couches turbulentes de l'atmosphère.
- Enfin, s'il est possible de mesurer la vibration des ondes radio, les capteurs optiques eux ne perçoivent qu'une quantité moyennée dans le temps, l'intensité.

Ces difficultés ont retardé la réalisation d'interféromètres optiques à longues bases, c'est-à-dire dont la distance maximale entre deux miroirs est typiquement de plusieurs dizaines de mètres au moins.

Il a fallu attendre 1987 pour que la possibilité de reconstruire des images par synthèse de Fourier à des longueurs d'onde IR/Optiques soit démontrée Haniff *et al.* (1987), et les premières reconstructions expérimentales d'images ont été effectuées en 1995 et 1996. Cependant, les traitements de données qui ont fait leurs preuves en radio-astronomie ne sont pas directement utilisables sur les données fournies par des interféromètres optiques : en raison des différences entre rayonnements radio et optiques, **les observables interférométriques optiques sont structurellement différents** de celles obtenues dans le domaine Radio. Les observables radio-interférométriques sont des *Visibilités Complexes* ; dans le domaine optique, les phases de ces visibilités complexes sont perturbées par des retards de phase turbulents aléatoires sur chaque télescope, et seules des *visibilités carrées* sont mesurables. Sans informations de phase, il devient très difficile d'estimer correctement l'objet d'intérêt. Si l'on dispose d'un modèle suffisamment simple et fiable de l'objet que l'on observe, il est toujours possible d'ajuster les paramètres de ce modèle sur les visibilités carrées uniquement. Mais la reconstruction obtenue sera extrêmement sensible aux erreurs de modèle.

Certaines informations en phase sont néanmoins mesurables lorsque l'on dispose d'un interféromètre optique d'au moins trois télescopes : les *clôtures de phase*. L'idée de former ces clôtures a été d'abord proposée par Jennison (1958) dans le contexte Radio, puis transposé au domaine Optique Jennison (1961); Rogstad (1968).

Les méthodes dites d'« auto-calibration », d'abord décrites par Cornwell et Wilkinson[Cornwell & Wilkinson (1981)], proposent des reconstructions d'images directement à partir des observables primaires de l'interférométrie optique : visibilités carrées et clôtures de phase. Ces méthodes ont été développées pour des radio-interféromètres « instables », pour

lesquels des erreurs instrumentales affectent les mesures de phase de façon analogue aux effets turbulents en optique. Cependant, dans le domaine optique, le modèle statistique du bruit sur les observables est fondamentalement différent de celui considéré par Cornwell et Wilkinson (Cf. section 4.1.2 page 95).

Il apparaît donc que les méthodes de reconstruction d'images spécifiquement dédiées à l'interférométrie optique ne sont pas encore arrivées à maturité. Cependant, le potentiel de l'imagerie par synthèse de Fourier optique a poussé la communauté astronomique à construire plusieurs interféromètres optiques (notamment COAST, CHARA, VLTI, IOTA, NPOI, KECK Interferometer), et de nombreuses données de type clôtures et visibilités carrées seront disponibles dans les années à venir. En conséquence, aux États-Unis comme en Europe, plusieurs actions visent au développement de méthodes de traitement de données adaptées. En particulier, le Jean-Marie Mariotti Center (JMMC) fédère les efforts français dans ce domaine. C'est ce qui justifie l'étude présentée dans ce manuscrit, et qui consiste à concevoir une méthode de reconstruction d'image :

- qui utilise un modèle de bruit sur les données correspondant à celui effectivement constaté sur des interféromètres optiques ;
- qui prenne en entrée des visibilités carrées ou en module, et des clôtures de phase ;
- qui soit efficace sur les données expérimentales issues d'interféromètres optiques en fonctionnement comprenant au moins trois télescopes.

**Modèle de bruit** Dans un premier temps, nous avons considéré que nous ne disposions pas de visibilités carrées et de clôtures de phase, mais directement de visibilités complexes, pour nous concentrer uniquement sur le modèle de bruit. Ce cadre correspond à des situations expérimentales actuelles (interféromètres spatiaux, interféromètres au sol à configuration redondante) et futures (systèmes à référence de phase, tel que l'instrument PRIMA du VLTI).
Nous avons donc cherché à établir une modélisation fine du bruit, puis à concevoir un algorithme de reconstruction d'images à partir de visibilités complexes exploitant cette modélisation. Nous avons ensuite entrepris de valider cet algorithme sur des données expérimentales obtenues sur l'interféromètre fibré de l'IRCOM (Chapitre 2).

**Une modélisation myope des effets turbulents** Les interféromètres en fonctionnement (IOTA, CHARA, NPOI, VLTI, COAST) recombinent au plus 6 télescopes, et les systèmes à référence de phase ne se généraliseront pas avant une dizaine d'années au moins. Aujourd'hui, il n'est donc généralement pas possible de mesurer des visibilités complexes, mais seulement des clôtures de phase et des visibilités carrées. Notre objectif à ce stade de l'étude a donc été d'introduire dans notre modélisation le manque d'information de phase dû aux effets turbulents. Nous avons opté pour une modélisation myope des informations manquantes, assimilées à des aberrations du système (Chapitre 3).
Nous avons alors entrepris de construire une méthode de reconstruction d'images astronomiques, à partir de visibilités carrées et de clôtures de phase, s'appuyant sur ce modèle myope (Chapitre 4).

**Validation expérimentale** Un tel algorithme serait de moindre valeur s'il n'était pas confronté aux données expérimentales qu'il est censé traiter. En effet, traiter des données réelles conduit toujours à découvrir des sources d'erreur imprévues dans les simulations, ou des pistes d'amélioration de la méthode. Bénéficiant des liens étroits entre l'ONERA et le Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique (LESIA), nous avons pu participer à une mission d'observation sur l'interféromètre à 3 télescopes IOTA, dans le but de mesurer

l'efficacité de notre méthode sur données réelles, et d'utiliser les reconstructions pour analyser la structure des étoiles $\chi$ du Cygne et $\alpha$ du Bouvier (Chapitre 5).

# Notations

| description | notation | exemples |
|---|---|---|
| Longueur d'onde | $\lambda$ | |
| objet (distribution d'intensité) | $o(\boldsymbol{\xi})$ | $o(\boldsymbol{\xi}) = \sum_k x_k\, f_k(\boldsymbol{\xi})$ |
| paramètres de l'objet | $\boldsymbol{x}$ | $\boldsymbol{x} = (x_1, \ldots, x_{N_{\text{pixel}}})^{\text{T}}$ |
| bases projetées | $\boldsymbol{u} = (u, v)$ | $\boldsymbol{u}_{j,k} = \boldsymbol{r}_k - \boldsymbol{r}_j$ |
| opérateur de bases | $\boldsymbol{B}$ | $\boldsymbol{u} = \boldsymbol{B} \cdot \boldsymbol{r}$ |
| fréquences spatiales | $\boldsymbol{\nu} = \boldsymbol{u}/\lambda$ | $\boldsymbol{\nu}_{j,k} = \frac{\boldsymbol{r}_k - \boldsymbol{r}_j}{\lambda}$ |
| coordonnées angulaires de la source | $\boldsymbol{\xi} = (\xi, \zeta)$ | |
| coordonnées spatiales des télescopes | $\boldsymbol{r}$ | |
| Fonction de coût, critère | $\mathcal{J}(\boldsymbol{x})$ | $\mathcal{J}(\boldsymbol{x}) = \mathcal{J}_{\text{data}}(\boldsymbol{x}) + \mathcal{J}_{\text{prior}}(\boldsymbol{x})$ |
| gradient du critère | $\frac{\partial \mathcal{J}(\boldsymbol{x})}{\partial \boldsymbol{x}}$ | |
| visibilités complexes | $\boldsymbol{y}$ | $\boldsymbol{y} = \boldsymbol{H} \cdot \boldsymbol{x} + \boldsymbol{y}_b$ |
| Opérateur de Fourier | $\boldsymbol{H}$ | $\boldsymbol{H} \cdot \boldsymbol{x}$ |
| aberrations myopes | $\boldsymbol{\alpha}$ | $\boldsymbol{H}(\boldsymbol{\alpha}) = \text{Diag}\left\{\exp \text{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}\right\} \cdot \boldsymbol{H}$ |
| amplitude de visibilité | $a$ ou $\boldsymbol{a}$ | $a_{\text{model}} = \lvert\boldsymbol{H} \cdot \boldsymbol{x}\rvert$ |
| phase de visibilité | $\phi$ ou $\boldsymbol{\phi}$ | $\boldsymbol{\phi}_{\text{model}} = \boldsymbol{\phi}(\boldsymbol{x}) = \arg(\boldsymbol{H} \cdot \boldsymbol{x})$ |
| phase de bispectre, clôtures | $\beta$ ou $\boldsymbol{\beta}$ | $\boldsymbol{\beta}(\boldsymbol{x}) = \boldsymbol{C} \cdot \boldsymbol{\phi}(\boldsymbol{x})$ |
| | | $\boldsymbol{\beta}_{\text{res}} = \text{arc}\,(\boldsymbol{\beta}_{\text{data}} - \boldsymbol{\beta}(\boldsymbol{x}))$ |
| spectre de puissance | $s$ ou $\boldsymbol{s}$ | $\boldsymbol{s}(\boldsymbol{x}) = \lvert\boldsymbol{H} \cdot \boldsymbol{x}\rvert^2$ |
| bruit | $y_b$, $a_b$, $\phi_b$ | |
| modèle | $\boldsymbol{y}(\boldsymbol{x})$, $\phi(\boldsymbol{x})$ | |
| Mesures | $y^{\text{data}}$ ou $y^{\text{d}}$ | $y^{\text{data}} = y(\boldsymbol{x}) + y_b$ |
| Résidus $= -$bruit | $a_{\text{res}}$ | $\boldsymbol{a}_{\text{res}} = \boldsymbol{a}(\boldsymbol{x}) - \boldsymbol{a}^{\text{data}} = -\boldsymbol{a}_b$ |
| Opérateur de clôture | $\boldsymbol{C}$ | $\boldsymbol{\beta} = \boldsymbol{C} \cdot \boldsymbol{\phi}$ |
| Matrice de covariance | $\boldsymbol{R}$ | $\boldsymbol{R}_a = \langle \boldsymbol{a}_{\text{res}} \cdot \boldsymbol{a}_{\text{res}}^{\text{T}} \rangle$ |
| Matrice de pondération | $\boldsymbol{W} = \boldsymbol{R}^{-1}$ | $\boldsymbol{W}_a$, $\boldsymbol{W}_\beta$ |
| $\chi^2$ | $\chi^2_a$ | $\chi^2_a = \boldsymbol{a}_{\text{res}}^{\text{T}} \cdot \boldsymbol{R}_a^{-1} \cdot \boldsymbol{a}_{\text{res}}$ |
| écart-type | $\sigma$ | |
| phase turbulente | $\varphi$ ou $\boldsymbol{\varphi}$ | |
| aberration myope | $\alpha$ ou $\boldsymbol{\alpha}$ | |
| fonction de base | $f_k(\boldsymbol{\xi})$ | $f_k(\boldsymbol{\xi}) = f(\boldsymbol{\xi} - \boldsymbol{\xi}_k)$ |
| Transformée de Fourier | $\hat{f}(\boldsymbol{\nu})$ ou $\text{TF}\,(f)$ | $\hat{f}(\boldsymbol{\nu}) = \int f(\boldsymbol{\xi})\, \text{e}^{-2\,\text{i}\,\pi\,\boldsymbol{\nu}\cdot\boldsymbol{\xi}}\, \text{d}\boldsymbol{\xi}$ |
| Transposé | $\boldsymbol{A}^{\text{T}}$ | |
| Transposé conjugué | $\boldsymbol{A}^{\text{H}}$ | |

# Chapitre 1

# Synthèse d'ouverture optique et Haute résolution angulaire.

Ce chapitre a pour but de présenter les divers aspects de la synthèse d'ouverture. Il s'inspire en partie de l'introduction du manuscrit de thèse de Frédéric Cassaing[Cassaing (1997)].

## 1.1 L'expérience des trous d'Young

### 1.1.1 L'expérience de 1804

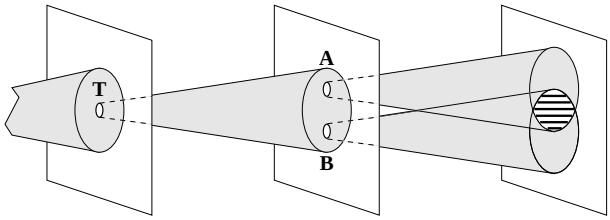



FIG. 1.1 – L'expérience des trous d'Young.

L'expérience de T. Young, publiée en 1804, a été marquante dans l'essor de la description ondulatoire de la lumière. Une source de lumière quasi ponctuelle, constituée d'un trou source $T$ percé dans un écran, éclaire deux autres trous $A$ et $B$ qui diffractent chacun un cône de lumière. À l'intersection de ces deux cônes, apparaissent des **franges d'interférence**, alternativement brillantes et sombres, qui disparaissent si l'on bouche un des trous. On arrive ainsi à la conclusion que l'« addition » de deux sources lumineuses peut créer des zones d'ombre.

Ce phénomène peut être expliqué en supposant que la lumière se propage sous la forme d'une onde sphérique sinusoïdale de période spatiale $\lambda$, la **longueur d'onde**, correspondant à la couleur du faisceau lumineux. La lumière blanche est une somme d'ondes de longueurs d'onde différentes. Les zones sombres apparaissent aux endroits où les ondes issues des deux trous sont en opposition de phase, c'est à dire qu'elles se compensent l'une l'autre. L'emplacement de ces franges sombres, et en particulier l'**interfrange** (la distance entre deux franges) dépend en particulier de la longueur d'onde de la source lumineuse. Young a ainsi pu déduire des caractéristiques de son interféromètre que la longueur d'onde de la lumière varie entre 0,4 $\mu$m (bleu) et 0,8 $\mu$m (rouge).

Supposons qu'il n'y ait pas une source en $T$, mais un ensemble de sources d'intensités différentes en entrée des trous d'Young. La figure d'interférences constatée sur l'écran est la

somme des figures que l'on obtiendrait pour chacune des sources prise isolément. La figure d'interférence formée par les trous d'Young donne donc des informations sur la distribution des sources. Pour étudier quantitativement la possibilité de retrouver la distribution de ces sources à partir de figures d'interférences, il est nécessaire d'introduire quelques notions mathématiques et optiques.

### 1.1.2 Diffraction de Fraunhofer à l'infini

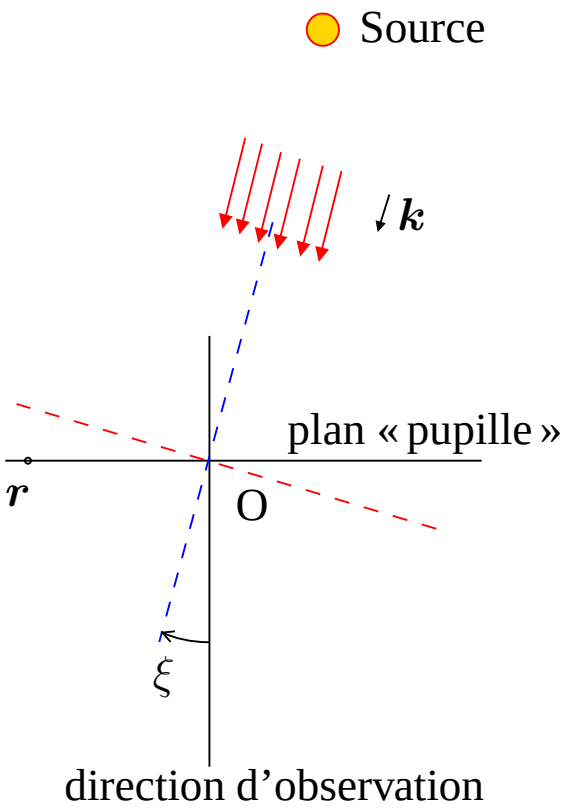



FIG. 1.2 – Diffraction de Fraunhofer à l'infini.

On considère une onde sphérique monochromatique émise par une étoile vue depuis la terre avec un angle $\boldsymbol{\xi} = (\xi, \zeta)$ par rapport à l'axe optique de l'instrument(Figure 1.1.2). Considérons un point $P$ du plan « pupille » (plan perpendiculaire à l'axe optique où se trouve la pupille d'entrée de l'instrument d'observation) repéré par $\vec{OP} = \boldsymbol{r} = (r, s)$. L'étoile étant très lointaine (séparée d'une distance grande devant $|\boldsymbol{r}|$), l'onde reçue est assimilable à une onde plane, de vecteur d'onde $\vec{k}$. Soit $\Psi(\boldsymbol{\xi})$ l'amplitude complexe du champ en $O$. Le champ électromagnétique dans le plan pupille s'écrit :

$$\Psi(\boldsymbol{r}) = \Psi(\boldsymbol{\xi}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$$

Pour simplifier, on considère une géométrie plane. L'étoile étant vue avec un angle $\xi$, alors

$$\overrightarrow{k} \cdot \boldsymbol{r} = \frac{2\pi}{\lambda} r \sin\xi$$

On se place dans l'approximation de Fraunhofer, c'est-à-dire que $\xi$ est petit. On a donc $\overrightarrow{k} \cdot \boldsymbol{r} \simeq \frac{2\pi}{\lambda} r\xi$.

Supposons maintenant qu'il n'y ait pas une seule étoile vue sous l'angle $\xi$, mais une infinité de sources. Alors le champ électromagnétique en $\boldsymbol{r}$ est la somme des contributions de chaque source :

$$\Psi(r) = \int_{\xi} \Psi(\xi) \mathrm{e}^{-2i\pi \frac{r}{\lambda}\xi} \mathrm{d}\xi$$

En deux dimensions, cette relation s'écrit :

$$\Psi(r,s) = \iint_{\xi,\zeta} \Psi(\xi,\zeta) \mathrm{e}^{-2i\pi\left(\frac{r}{\lambda}\xi + \frac{s}{\lambda}\zeta\right)} \mathrm{d}\xi \mathrm{d}\zeta$$

Si l'on considère les coordonnées réduites dans le plan pupille $\frac{\boldsymbol{r}}{\lambda}$, l'équation précédente prend alors la forme d'une transformation de Fourier bidimensionnelle :

$$\boxed{\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) = \widetilde{\Psi}(\boldsymbol{\xi})} \tag{1.1}$$

$\Psi(\xi)$ est, à un coefficient multiplicatif près dû à la propagation, l'amplitude de l'onde sphérique émise par la source située dans le ciel aux coordonnées angulaires $\boldsymbol{\xi} = (\xi, \zeta)$. Ce coefficient multiplicatif peut être considéré comme indépendant de la direction de visée $\boldsymbol{\xi}$ pour des « petits » angles.
Le résultat final est que **l'amplitude du champ propagé à une distance infinie du plan contenant les sources d'intérêt est la transformée de Fourier de la distribution d'amplitude sur ce plan.**

## 1.1.3 Processus d'imagerie

### 1.1.3.1 Processus idéal

Pour imager la distribution d'amplitude de la source $\Psi(\xi,\zeta)$, il faut donc réaliser une transformée de Fourier inverse (ou une transformée de Fourier directe plus un retournement). Ce procédé constitue le processus d'imagerie :

$$\Psi(\boldsymbol{\xi}) \xrightarrow[TF]{\text{propagation}} \Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) = \widetilde{\Psi}(\boldsymbol{\xi}) \xrightarrow[TF^{-1}]{\text{imagerie}} \Psi(\boldsymbol{\xi}) \tag{1.2}$$

Par exemple, une lentille convergente effectue cette transformation. Le champ imagé sur un écran placé au plan focal de la lentille s'écrit :

$$\Psi_{\boldsymbol{im}}(\boldsymbol{\xi}) = \mathrm{TF}^{-1}\left[\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right)\right] = \iint_{\frac{\boldsymbol{r}}{\lambda}} \Psi(\frac{\boldsymbol{r}}{\lambda}) \mathrm{e}^{2i\pi\left(\frac{\boldsymbol{r}}{\lambda}\cdot\boldsymbol{\xi}\right)} \mathrm{d}\frac{\boldsymbol{r}}{\lambda}$$

### 1.1.3.2 Fonction pupille

On définit 3 plans dans l'espace :

**Plan objet :** c'est le plan de l'objet d'intérêt, où le champ électromagnétique est $\Psi(\xi)$.

**Plan pupille ou de Fourier :** c'est le plan situé à l'infini du plan objet, dans lequel le champ électromagnétique est $\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) = \widetilde{\Psi}(\boldsymbol{\xi})$, d'où la dénomination de plan de Fourier.

**Plan focal ou image :** c'est le plan dans lequel est formée l'image, et où le champ électromagnétique est idéalement $\Psi_{\boldsymbol{im}}(\xi) = \Psi(\xi)$.

La précédente équation ne tient pas compte des imperfections de l'instrument, c'est-à-dire du procédé permettant de réaliser une Transformée de Fourier inverse. On modélise ces imperfections par une fonction pupille $p\left(\frac{\boldsymbol{r}}{\lambda}\right)$, ou transmittance. Un instrument idéal est tel que la transmittance vaut 1 partout.

L'opération d'imagerie s'écrit maintenant :

$$\Psi_{\boldsymbol{im}}(\boldsymbol{\xi}) = \mathrm{TF}^{-1}\left[p\left(\frac{\boldsymbol{r}}{\lambda}\lambda\right)\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right)\right] = \iint_{\frac{\boldsymbol{r}}{\lambda}} p\left(\frac{\boldsymbol{r}}{\lambda}\right)\Psi(\frac{\boldsymbol{r}}{\lambda})\mathrm{e}^{2i\pi(\frac{\boldsymbol{r}}{\lambda}\cdot\boldsymbol{\xi})}\mathrm{d}\frac{\boldsymbol{r}}{\lambda}$$

La pupille de l'instrument, c'est-à-dire sa surface collectrice, est d'extension finie : il n'est donc pas possible d'accéder à la totalité du spectre $\widetilde{\Psi}(\xi)$ de l'objet. En d'autres termes, la transmittance de la pupille $p(\frac{\boldsymbol{r}}{\lambda})$ vaut 0 en dehors de la surface collectrice. Si l'on appelle $D$ la plus grande dimension de la pupille, la fréquence maximale transmise est $D/\lambda$. Ainsi, toute observation optique peut être considérée comme un **filtrage spatial** sur l'objet observé. L'oeil humain a une pupille circulaire d'un diamètre $D$ de 3 à 8 mm. Un oeil parfait a une transmittance valant 1 sur un cercle de quelques millimètres, et 0 ailleurs.

#### 1.1.3.3 Détection

Une autre limitation fondamentale rend le processus d'imagerie plus complexe que le schéma idéal de l'équation 1.2 : les détecteurs optiques ne sont pas sensibles au champ électromagnétique mais à sa moyenne quadratique, l'**intensité** $\langle|\Psi|^2\rangle_t$.

$$\begin{array}{ccccc} \Psi(\boldsymbol{\xi}) & \xrightarrow[TF]{\text{propagation}} & \Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) = \widetilde{\Psi}(\boldsymbol{\xi}) & \xrightarrow[\times p]{\text{filtrage pupillaire}} & p(\frac{\boldsymbol{r}}{\lambda})\,\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) \\ p(\frac{\boldsymbol{r}}{\lambda})\,\Psi\left(\frac{\boldsymbol{r}}{\lambda}\right) & \xrightarrow[TF^{-1}]{\text{imagerie}} & \Psi_{\boldsymbol{im}}(\boldsymbol{\xi}) & \xrightarrow[\langle||^2\rangle_t]{\text{détection}} & \boldsymbol{im} = \left\langle|\Psi_{\boldsymbol{im}}(\boldsymbol{\xi})|^2\right\rangle_t \end{array} \tag{1.3}$$

Ce que l'on peut percevoir de l'objet dans un cas idéal où le processus d'imagerie est parfait est : $\langle|\Psi(\boldsymbol{\xi})|^2\rangle_t$, ce qui correspond à $p(\frac{\boldsymbol{r}}{\lambda}) = 1$, pour tout $\boldsymbol{r}$. On appellera donc cette distribution d'intensité de la source « l'objet » :

$$\boxed{\boldsymbol{o}(\boldsymbol{\xi}) = \left\langle|\Psi(\boldsymbol{\xi})|^2\right\rangle_t}$$

#### 1.1.3.4 Modèle convolutif

On définit les opérateurs $\mathcal{H}$ et $\mathcal{G}$ correspondant aux processus successifs de propagation à l'infini et d'imagerie :

$$\begin{array}{ccc} \Psi_{\boldsymbol{o}}(\boldsymbol{\xi}) & \xrightarrow[\mathcal{H}]{\text{propagation+imagerie}} & \Psi_{\boldsymbol{im}}(\boldsymbol{\xi}) \\ \text{détection}\downarrow \langle||^2\rangle_t & & \text{détection}\downarrow \langle||^2\rangle_t \\ \boldsymbol{o}(\boldsymbol{\xi}) & \xrightarrow[\mathcal{G}]{\text{propagation+imagerie}} & \boldsymbol{im}(\boldsymbol{\xi}) \end{array} \tag{1.4}$$

On peut démontrer (Cf Annexe D page 167) que non seulement $\mathcal{H}$ est Linéaire Invariant par Translation (LIT), mais aussi que, sous l'hypothèse que deux points source rayonnent de façon décorrélée, $\mathcal{G}$ l'est aussi. Cet opérateur prend donc la forme d'un produit de convolution par la Réponse Impulsionnelle (RI)(Goodman, 1968) :

$$\boxed{\boldsymbol{im} = \boldsymbol{o} * RI}$$

La RI s'exprime en fonction de la transmittance $p$ :

$$\boxed{RI = \widetilde{p \otimes p}}$$

La relation entre les TF de l'objet et de l'image est un produit simple par **la fonction de transfert optique** :

$$\boxed{\widetilde{\boldsymbol{im}} = \widetilde{\boldsymbol{o}} \times FTO}$$

La FTO s'exprime simplement comme **l'auto corrélation de la pupille**

$$\boxed{FTO = p \otimes p}$$

#### 1.1.3.5 Imagerie et configuration pupillaire

Ainsi, la **forme de la pupille** conditionne le **filtrage spatial** effectué sur l'objet. Par ailleurs, le support de l'auto-corrélation, de support double par rapport à celui de la pupille, est fini : l'instrument présente donc une **fréquence de coupure** franche au delà de laquelle la transmission fréquentielle est rigoureusement nulle. Les instruments d'optique sont des filtres passe-bas (l'auto-corrélation est maximale à l'origine), parfois aussi passe-bande.

Pour un télescope monolithique à pupille circulaire, la fonction de transfert est une classique fonction en forme de "chapeau chinois" et la réponse impulsionnelle est la tache d'Airy (fig. 1.3a). En interférométrie, compte tenu de la grande variété de formes que peut prendre la pupille, la fonction de transfert peut être toute autre. Par exemple, pour les trous d'Young (fig. 1.3b), la FTO est constituée de trois pics. Le **pic central**, auto-corrélation d'une pupille, correspond à la FTO d'un télescope seul ; il est accompagné de deux **pics "frange"** symétriques et de hauteur moitié, traduisant l'interférence entre les télescopes. La RI est donc la tache de diffraction d'un télescope, modulée par des franges sinusoïdales. Une telle configuration pupillaire ne collecte qu'une faible partie des fréquences de l'objet vues par le monolithique de la figure 1.3a. Pour mieux couvrir le plan des fréquences spatiales, il est nécessaire d'utiliser une configuration avec plus de télescopes. Avec la configuration dite « Golay-6 »[Golay (1971)], la FTO couvre le même domaine fréquentiel que le monolithique équivalent même si les hautes fréquences sont plus atténuées (fig. 1.3c).

Pour classer la grande diversité des pupilles possibles en synthèse d'ouverture, la terminologie suivante est généralement adoptée :

- toute pupille non connexe (en plusieurs parties disjointes) est dite **lacunaire** ;
- une pupille est dite **compacte** lorsque la fonction de transfert associée ne s'annule pas avant la fréquence de coupure. Sinon, elle est dite **diluée** ;
- une pupille est dite **redondante** lorsque plusieurs paires de télescopes forment la même base (en longueur et direction). Sinon, elle est **non redondante** ;

Ces notions sont illustrées en figure 1.4, sur un exemple monodimensionnel.

### 1.1.4 Retour à l'expérience des trous d'Young

#### 1.1.4.1 Réponse impulsionnelle

Dans le cas des trous d'Young, la pupille est opaque partout sauf en deux points du plan pupille, repérés par $\boldsymbol{r}_1$ et $\boldsymbol{r}2$ :

$$p\left(\frac{\boldsymbol{r}}{\lambda}\right) = p_1\delta\left(\frac{\boldsymbol{r}-\boldsymbol{r}_1}{\lambda}\right) + p_2\delta\left(\frac{\boldsymbol{r}-\boldsymbol{r}_2}{\lambda}\right)$$

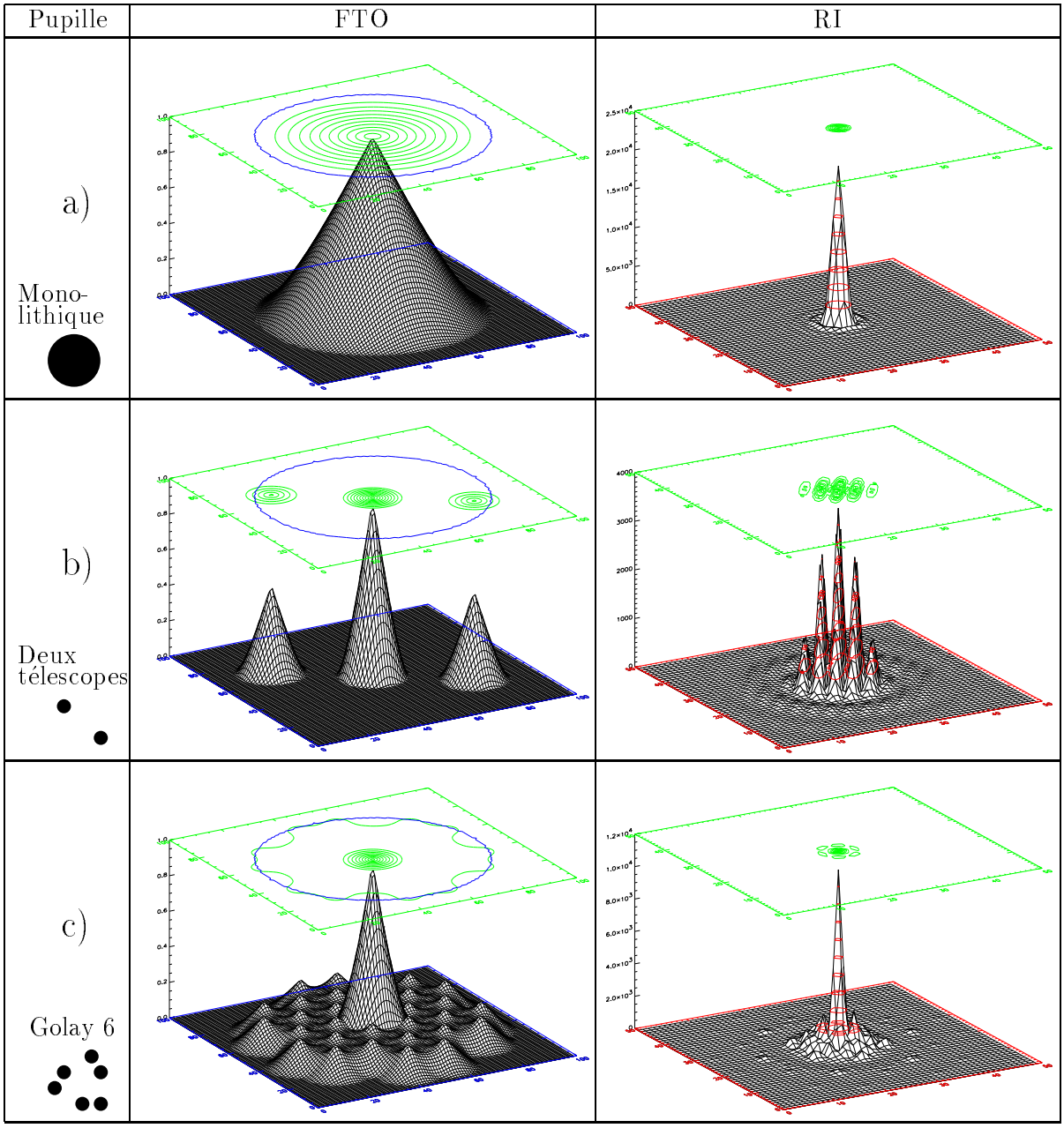



FIG. 1.3 – Influence de la configuration pupillaire sur la formation d'images[Cassaing (1997)].

Cette fonction pupille correspond à deux télescopes dont le diamètre est négligeable devant la base les séparant. Dans ce cadre, dit d'interférométrie à longues bases, chaque télescope peut être vu comme une ouverture ponctuelle (c'est le cas notamment pour le Very Large Telescope Interferometer (VLTI), et plus généralement, c'est dans ce cadre que nous nous plaçons dans ce manuscrit). On définit :

**la base** des trous d'Young par le vecteur reliant les deux trous/télescopes : $\boldsymbol{u}_{12} = \boldsymbol{r}_2 - \boldsymbol{r}_1$

**la fréquence spatiale** correspondante $\boldsymbol{\nu}_{12} = \frac{\boldsymbol{u}_{12}}{\lambda}$

On considère un objet de distribution d'intensité $\boldsymbol{o}(\boldsymbol{\xi})$. L'image formée au plan focal est :

$$\boldsymbol{im}_{12}(\boldsymbol{\xi}) = \boldsymbol{o}(\boldsymbol{\xi}) * \widetilde{(p \otimes p)}(\boldsymbol{\xi})$$

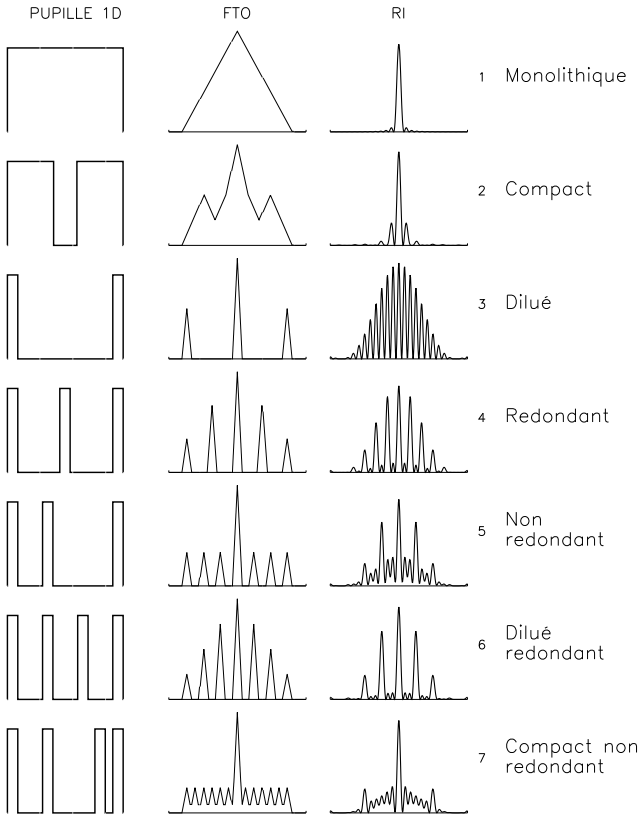



FIG. 1.4 – Pupilles monodimensionnelles, avec leurs fonction de transfert et réponse impulsionnelle[Cassaing (1997)].

Dans le cas d'un instrument dont la pupille est percée des deux trous d'Young , on a :

$$\begin{aligned} RI(\boldsymbol{\xi}) &= \widetilde{(p \otimes p)}(\xi) \\ &= |\widetilde{p}|^2(\xi) \\ &= \left| \iint p(\frac{\boldsymbol{r}}{\lambda}) \mathrm{e}^{-2i\pi(\frac{\boldsymbol{r}}{\lambda}\cdot\boldsymbol{\xi})} \mathrm{d}\frac{\boldsymbol{r}}{\lambda} \right|^2 \\ &= |p_1|^2 + |p_2|^2 + p_1 p_2^* \mathrm{e}^{2\mathrm{i}\pi(\boldsymbol{\nu}\cdot\boldsymbol{\xi})} + p_1^* p_2 \mathrm{e}^{-2\mathrm{i}\pi(\boldsymbol{\nu}\cdot\boldsymbol{\xi})} \end{aligned}$$

La RI est constituée d'un terme constant, et de deux termes sinusoïdaux. La FTO présente donc 3 pics, 1 pic central, et deux pics frange.

#### 1.1.4.2 Image formée par les trous d'Young

L'image formée sera donc :

$$\begin{aligned} \boldsymbol{im}_{12}(\boldsymbol{\xi}) &= \boldsymbol{o}(\boldsymbol{\xi}) * RI(\boldsymbol{\xi}) \\ &= \iint \boldsymbol{o}(\boldsymbol{\xi}') RI(\boldsymbol{\xi} - \boldsymbol{\xi}') \mathrm{d}\boldsymbol{\xi}' \end{aligned}$$

On pose :

$$\widetilde{\boldsymbol{o}}(\boldsymbol{\nu}_{12}) = a(\boldsymbol{\nu}_{12}) e^{i\phi(\boldsymbol{\nu}_{12})} \tag{1.5}$$

$$p_i = |p_i| e^{i\varphi_i} \tag{1.6}$$

Le calcul donne :

$$\boldsymbol{im}_{12}(\boldsymbol{\xi}) = (|p_1|^2 + |p_2|^2)a(\mathbf{0}) + 2|p_1 p_2| a(\boldsymbol{\nu}_{12}) \cos\left(2\pi\nu_{12} \cdot \boldsymbol{\xi} + \phi(\boldsymbol{\nu}_{12}) + \varphi_2 - \varphi_1\right)$$

L'interférogramme $\boldsymbol{im}_{12}(\boldsymbol{\xi})$ correspond bien à la description de l'expérience de trous d'Young évoquée en début de chapitre.

#### 1.1.4.3 Le théorème de Zernike-Van Cittert

La TF de l'objet $\widetilde{\boldsymbol{o}}(\boldsymbol{\nu}_{12}) = a(\boldsymbol{\nu}_{12})e^{i\phi(\boldsymbol{\nu}_{12})}$, c'est-à-dire la TF de la distribution angulaire d'**intensité** de la source, peut s'exprimer en fonction du *degré de cohérence complexe* $\gamma_{12}$ du **champ** dans le plan pupille $\gamma_{12} = \langle \Psi^*(\boldsymbol{r}_1/\lambda)\Psi(\boldsymbol{r}_2/\lambda)\rangle_t$. En effet, avec la relation $\Psi(\boldsymbol{r}_1/\lambda) = \widetilde{\Psi}(\boldsymbol{\xi})$, on peut écrire :

$$\begin{aligned}
\gamma_{12} &= \left\langle \int_{\boldsymbol{\xi}_1} \Psi^*(\boldsymbol{\xi}_1) \mathrm{e}^{2\mathrm{i}\pi\boldsymbol{\xi}_1 \cdot \frac{\boldsymbol{r}_1}{\lambda}} \mathrm{d}\boldsymbol{\xi}_1 \times \int_{\boldsymbol{\xi}_2} \Psi(\boldsymbol{\xi}_2) \mathrm{e}^{-2\mathrm{i}\pi\boldsymbol{\xi}_2 \cdot \frac{\boldsymbol{r}_2}{\lambda}} \mathrm{d}\boldsymbol{\xi}_2 \right\rangle_t \\
&= \left\langle \iint_{\boldsymbol{\xi}_1 \boldsymbol{\xi}_2} \Psi^*(\boldsymbol{\xi}_1)\Psi(\boldsymbol{\xi}_2) \mathrm{e}^{2\mathrm{i}\pi\left(\boldsymbol{\xi}_1 \cdot \frac{\boldsymbol{r}_1}{\lambda} - \boldsymbol{\xi}_2 \cdot \frac{\boldsymbol{r}_2}{\lambda}\right)} \mathrm{d}\boldsymbol{\xi}_1 \mathrm{d}\boldsymbol{\xi}_2 \right\rangle_t \\
&= \iint_{\boldsymbol{\xi}_1 \boldsymbol{\xi}_2} \langle \Psi^*(\boldsymbol{\xi}_1)\Psi(\boldsymbol{\xi}_2)\rangle_t \, \mathrm{e}^{2\mathrm{i}\pi\left(\boldsymbol{\xi}_1 \cdot \frac{\boldsymbol{r}_1}{\lambda} - \boldsymbol{\xi}_2 \cdot \frac{\boldsymbol{r}_2}{\lambda}\right)} \mathrm{d}\boldsymbol{\xi}_1 \mathrm{d}\boldsymbol{\xi}_2
\end{aligned}$$

L'hypothèse selon laquelle chaque point de la source observée rayonne de façon incohérente peut s'écrire :

$$\langle \Psi^*(\boldsymbol{\xi}_1)\Psi(\boldsymbol{\xi}_2)\rangle_t = \delta(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2) \left\langle |\Psi^*(\boldsymbol{\xi}_1)|^2 \right\rangle_t = \delta(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\boldsymbol{o}(\boldsymbol{\xi}_1)$$

d'où

$$\boxed{\gamma_{12} = \int_{\boldsymbol{\xi}} \boldsymbol{o}(\boldsymbol{\xi}) \mathrm{e}^{-2\mathrm{i}\pi\boldsymbol{\xi} \cdot \frac{\boldsymbol{r}_2 - \boldsymbol{r}_1}{\lambda}} \mathrm{d}\boldsymbol{\xi} = \widetilde{\boldsymbol{o}}(\boldsymbol{\nu}_{12})} \tag{1.7}$$

Ce résultat correspond au **théorème de Zernike-Van Cittert**, établi tout d'abord par Van Cittert en 1934 puis sous une autre forme par Zernike en 1938 : **le degré de cohérence complexe $\gamma$ dans un plan éloigné d'une source monochromatique et spatialement incohérente est la transformée de Fourier normalisée de la distribution d'intensité de l'objet**.

### 1.1.5 Formation d'interférogrammes par deux télescopes

On considère l'interféromètre à deux télescopes présenté figure 1.1.5. Les faisceaux issus des télescopes, situés en $\boldsymbol{r}_1$ et $\boldsymbol{r}_2$, passent respectivement par les lignes à retard LR1 et LR2, avant d'être recombinés. Le retard relatif du faisceau issu du télescope 2 par rapport à celui issu de télescope 1 est donné par :

$$d = d_2 - d_1$$

On considère une source ponctuelle monochromatique, donnant lieu à une onde plane de vecteur d'onde $\boldsymbol{k}$, avec :

$$|\boldsymbol{k}| = \frac{2\pi}{\lambda}$$

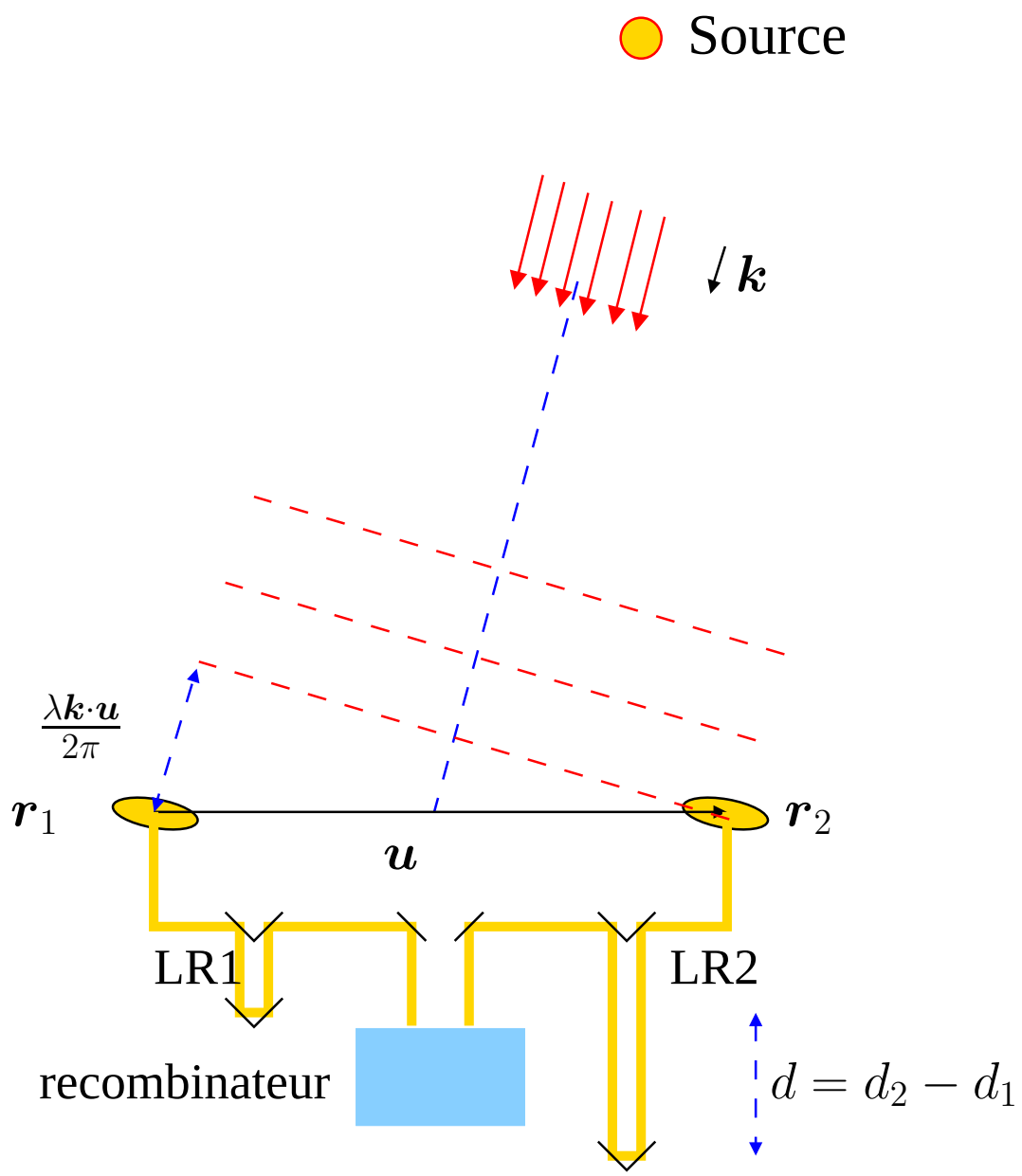



FIG. 1.5 – Un interféromètre à deux télescopes

#### 1.1.5.1 Interférogrammes en lumière monochromatique

Si le diamètre des télescopes est petit devant la base $\boldsymbol{u}$, on peut assimiler chaque télescope à une ouverture ponctuelle. On suppose que les deux télescopes transmettent autant de flux l'un que l'autre :

$$|p_1| = |p_2| = \tau$$

En l'absence de turbulence, le seul déphasage introduit par l'instrument provient des lignes à retard :

$$\varphi_2 - \varphi_1 = \frac{2\pi}{\lambda}(d_2 - d_1) = \frac{2\pi d}{\lambda}$$

L'image formée par l'interféromètre est :

$$\boldsymbol{im}(\boldsymbol{\xi}) = (|p_1|^2 + |p_2|^2)a(\mathbf{0}) + 2|p_1 p_2| a(\boldsymbol{\nu}_{12}) \cos\left(2\pi\nu_{12}\cdot\boldsymbol{\xi} + \phi(\boldsymbol{\nu}_{12}) + \frac{2\pi d}{\lambda}\right)$$

Si $p_2 = 0$ alors $\boldsymbol{im}_{12}(\boldsymbol{\xi}) = \boldsymbol{im}_1 = |p_1|^2 a(\mathbf{0}) = \tau^2 a(\mathbf{0}) = \boldsymbol{im}_{tel}$
Si $p_1 = 0$ alors $\boldsymbol{im}_{12}(\boldsymbol{\xi}) = \boldsymbol{im}_2 = |p_2|^2 a(\mathbf{0}) = \tau^2 a(\mathbf{0}) = \boldsymbol{im}_{tel}$
On peut donc réécrire l'image formée comme suit :

$$\boldsymbol{im}(\boldsymbol{\xi}) = 2\tau^2 a(\mathbf{0}) \left[1 + \frac{a(\boldsymbol{\nu}_{12})}{a(\mathbf{0})} \cos \Phi_{12}(\boldsymbol{\xi})\right]$$

avec

$$\begin{aligned}\Phi_{12}(\boldsymbol{\xi}) &= 2\pi \boldsymbol{\nu}_{12} \cdot \boldsymbol{\xi} + \phi(\boldsymbol{\nu}_{12}) + \frac{2\pi d}{\lambda} \\ &= \frac{2\pi}{\lambda} \boldsymbol{u}_{12} \cdot \boldsymbol{\xi} + \phi(\boldsymbol{\nu}_{12}) + \frac{2\pi d}{\lambda}\end{aligned}$$

$a(\mathbf{0})$ est le flux total de l'objet. Sans perte de généralité, on considère dans la suite des transformées de Fourier normalisées, ce qui revient à poser $a(\mathbf{0}) = 1$.
Le terme $\tau^2 a(\mathbf{0})$ correspond à l'image qui serait formée si l'on bouchait l'un des deux trous, ou que l'on obturait l'un des deux télescopes, que l'on notera $\boldsymbol{im}_{tel}$. Nous avons considéré des trous d'Young ponctuels. Si l'on ne néglige plus le diamètre des trous/télescopes $D_{tel}$ devant la base $|\boldsymbol{u}|$, alors $\boldsymbol{im}_{tel}$ n'est plus une constante, mais correspond à une tache d'Airy.
La longueur de la ligne à retard $d$ vient s'ajouter à la différence de marche $\boldsymbol{\nu}_{12} \cdot \boldsymbol{\xi}$. On définit donc une différence de marche totale :

$$d'_{12} \triangleq \boldsymbol{u}_{12} \cdot \boldsymbol{\xi} + d \tag{1.8}$$

Avec ceci, la phase de l'interférogramme s'écrit :

$$\Phi_{12}(d') = \frac{2\pi}{\lambda} \boldsymbol{u}_{12} \cdot \boldsymbol{\xi} + \phi(\boldsymbol{\nu}_{12}) + \frac{2\pi d}{\lambda} = k d'_{12} + \phi(\boldsymbol{\nu}_{12})$$

d'où

$$\boxed{\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d') \left[1 + a(\boldsymbol{\nu}) \cos(k d'_{12} + \phi(\boldsymbol{\nu}_{12}))\right]} \tag{1.9}$$

L'interférogramme $\boldsymbol{im}(d')$ correspond bien à l'expérience de trous d'Young évoquée en début de chapitre. Il est formé est un réseau de franges, dont :
– le contraste permet de déterminer le **module** $a(\boldsymbol{\nu}_{12})$ de la TF normalisée de l'objet à la fréquence $\boldsymbol{\nu}$, ou **amplitude de visibilité** ;
– la modulation de phase permet de déterminer la **phase** $\phi(\boldsymbol{\nu}_{12})$ de la TF de l'objet à la fréquence $\boldsymbol{\nu}_{12}$, ou **phase de visibilité**, connaissant le déphasage $k d'_{12}$.

#### 1.1.5.2 Prise en compte de la largeur spectrale du filtre

Jusqu'à présent, nous avons considéré des sources monochromatiques. Considérons à présent une source polychromatique, d'intensité spectrale $F(k - \bar{k})$, et supposons que l'interféromètre a une fonction de transfert spectrale $\eta(k - \bar{k})$, en fonction du nombre d'onde $k = 2\pi/\lambda$[1].
L'interférogramme obtenu est la somme des interférogrammes correspondant à chaque longueur d'onde, multipliés par un facteur $W(k - \bar{k}) = F(k - \bar{k})\eta(k - \bar{k})$ :

$$\boldsymbol{im}(d') = \int_k W(k - \bar{k}) \boldsymbol{im}(d', k) \mathrm{d}k$$

[1] $\bar{k}$ est un nombre d'onde arbitraire, correspondant par exemple à la transmission maximale de l'interféromètre.

avec

$$\int_k W(k-\bar{k}) = 1$$

Le terme $\boldsymbol{im}(\xi, k)$ présente une double dépendance en $k$ :

– la distribution d'intensité de la source $\boldsymbol{o}$ peut être différente selon le nombre d'onde, et la définition 1.5 page 17 devient :

$$\widetilde{\boldsymbol{o}_k}(\boldsymbol{\nu}) = a_k(\boldsymbol{\nu}) e^{i\phi_k(\boldsymbol{\nu})};$$

– la fréquence d'echantillonnage $\boldsymbol{\nu}$ s'écrit $\boldsymbol{\nu} = \frac{\boldsymbol{u}}{\lambda} = \frac{k\boldsymbol{u}}{2\pi}$

D'après (1.9), on a :

$$\boldsymbol{im}(d', k) = 2\boldsymbol{im}_{tel}(d', k)\,[1 + a_k(k\boldsymbol{u}/2\pi)\cos(kd' + \phi_k(k\boldsymbol{u}/2\pi))] \tag{1.10}$$

Le terme $\boldsymbol{im}_{tel}(d', k)$ correspond à la diffraction de l'onde plane incidente par une ouverture circulaire. Il dépend de $k$, mais avec une variation lente (c'est la somme de figures d'Airy, dont le diamètre varie selon $k$. Au niveau du maximum, la dispersion en longueurs d'onde est faible). On fait donc l'approximation $\boldsymbol{im}_{tel}(d', k) \simeq \boldsymbol{im}_{tel}(d', \bar{k})$.

Etant donnée la largeur spectrale réduite de la plupart des filtres utilisés, on peut de plus supposer que la distribution d'intensité de la source est identique à toutes les longueurs d'onde :

$$\begin{cases} a_k(k\boldsymbol{u}/2\pi) = a(k\boldsymbol{u}/2\pi) \\ \phi_k(k\boldsymbol{u}/2\pi) = \phi(k\boldsymbol{u}/2\pi) \end{cases}$$

Avec ces approximations, on a :

$$\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k}) \int_k W(k-\bar{k})\,[1 + a(k\boldsymbol{u}/2\pi)\cos(kd' + \phi(k\boldsymbol{u}/2\pi))]$$

Cette intégrale n'est pas simple à calculer dans le cas général. Il existe cependant deux cas particuliers pour lesquels l'expression précedente se simplifie.

**Cas d'une variation lente de $a(\boldsymbol{\nu})$ et $\phi(\boldsymbol{\nu})$ :** par exemple, si l'objet est un dirac au centre du champ, sa TF est uniforme et ne dépend pas de $\boldsymbol{\nu}$ :

$$a(\boldsymbol{\nu}) = a,\ \phi(\boldsymbol{\nu}) = \phi.$$

L'étude de la validité de cette hypothèse est hors du champ de ce manuscrit. Dans ce cas, on a :

$$\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k}) \int_k W(k-\bar{k})\,[1 + a\cos(kd' + \phi)]\,\mathrm{d}k$$

Posons $k' = k - \bar{k}$ :

$$\begin{aligned} \boldsymbol{im}(d') &= 2\boldsymbol{im}_{tel}(d', \bar{k}) \int_k W(k-\bar{k})\,[1 + a\cos(kd' + \phi)]\,\mathrm{d}k \\ &= 2\boldsymbol{im}_{tel}(d', \bar{k}) \left[1 + a\int_k W(k-\bar{k})\cos(kd' + \phi)\mathrm{d}k\right] \end{aligned}$$

$W$ etant réel, le calcul donne :

$$\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k}) \left[1 + a\Re\mathrm{e}\left\{\hat{W}(d')\mathrm{e}^{\mathrm{i}(d'\bar{k}+\phi)}\right\}\right]$$

Si $W$ est une fonction rectangulaire de largeur $2\Delta k$ centrée en $\bar{k}$, l'interférogramme obtenu par superposition des interférogrammes obtenus pour chaque longueur d'onde est :

$$\boxed{\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k}) \left[1 + a\mathrm{sinc}(d'\Delta k)\cos(d'\bar{k} + \phi)\right]} \tag{1.11}$$

où $\mathrm{sinc}(a) \triangleq \frac{\sin a}{a}$.
En lumière monochromatique de nombre d'onde $\bar{k}$, la figure d'interférence que l'on obtient (Eq. 1.9 page 20) s'écrit :

$$\boldsymbol{im}(d', \bar{k}) = 2\boldsymbol{im}_{tel}(d', \bar{k}) \left[1 + a\cos(\bar{k}d' + \phi)\right]$$

On constate donc un effet de modulation des franges par un sinus cardinal, dont la pseudo-période est inversement proportionnelle à la largeur du filtre (Figure 1.6).

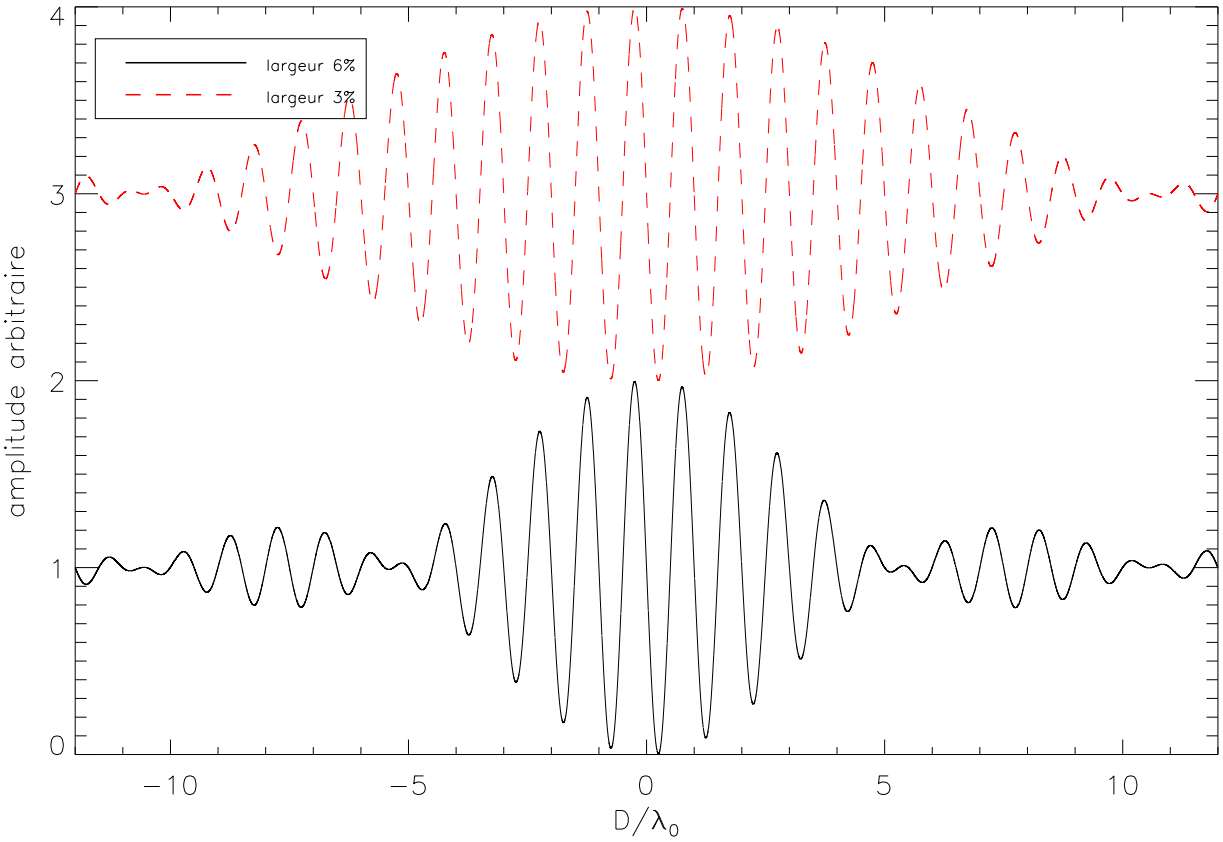



FIG. 1.6 – Modulation des franges par la dispersion spectrale, pour $2\Delta k/\bar{k} = 3\%$ et $2\Delta k/\bar{k} = 6\%$

Cette étude simplifiée permet de prévoir l'allure de l'interférogramme. Néanmoins, l'équation (1.11) peut conduire à des conclusions erronnées. En effet, le maximum du sinc correspond à $d' = 0$. En ce point, l'argument du cos serait $\bar{k}d' + \phi = \phi$, et l'on pourrait déduire la phase $\phi$ en comparant les maximums de la courbe sinusoïdale et de l'enveloppe, sans connaitre précisément la position des lignes à retard, ce qui n'est pas vrai en pratique.

**Cas d'une variation linéaire en $k$ de $a(\boldsymbol{\nu})$ et $\phi(\boldsymbol{\nu})$ :** par exemple, si l'objet est un dirac excentrée, sa TF est uniforme en module, mais sa phase évolue linéairement :

$$a(\boldsymbol{u}) = a(k\boldsymbol{u}/2\pi) = a,\ \phi(\boldsymbol{\nu}) = \phi(k\boldsymbol{u}/2\pi) = \phi_0 \frac{k\boldsymbol{u}}{2\pi} = kd_\phi.$$

Dans ce cas, on a :

$$\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k}) \int_k W(k-\bar{k})\,[1 + a\cos k(d' + d_\phi)]\,\mathrm{d}k$$

Si $W$ est une fonction rectangulaire de largeur $2\Delta k$ centrée en $\bar{k}$, l'interférogramme obtenu par superposition des interférogrammes obtenus pour chaque longueur d'onde est :

$$\boldsymbol{im}(d') = 2\boldsymbol{im}_{tel}(d', \bar{k})\,[1 + a\mathrm{sinc}\Delta k(d' + d_\phi)\cos k(d' + d_\phi)]$$

Dans ce cas, les maximums de la courbe sinusoïdale et de l'enveloppe sont confondus, et il est nécessaire de connaitre précisément la position des lignes à retard pour déterminer $\phi$.

**Cas général** Dans le cas général, il est difficile d'extraire une information de la position du maximum de l'enveloppe. De plus, pour pouvoir mieux l'exploiter, on disperse généralement l'interférogramme en plusieur sous-bandes, lesquelles sont trop étroites pour que l'enveloppe soit visible, a fortiori exploitable.
L'information de phase à extraire est la position d'*un* maximum du cosinus, tel que :

$$\bar{k}d' + \phi = 2k\pi$$

#### 1.1.5.3 Modulation de l'interférogramme avec les lignes à retard

Plutôt que d'enregistrer tout l'interférogramme avec un capteur multi-pixels, il est possible de placer un capteur monopixel en $d' = 0$, et de de faire varier $d$ en commandant les lignes à retard, de façon à faire défiler l'interférogramme sur le capteur.

## 1.2 La mesure interférométrique optique en astronomie

### 1.2.1 Interféromètres au sol

#### 1.2.1.1 Super-synthèse

Un interféromètre de Michelson est un réseau de télescopes recombinés interférométriquement deux à deux. Considérons que l'on dispose de $N_{tels}$ télescopes, dont les positions respectives projetées sur le plan d'observation à l'instant $t$ sont $\{\boldsymbol{r}_1(t), \boldsymbol{r}_2(t), \ldots, \boldsymbol{r}_{N_{tels}}(t)\}$. ces positions changent au cours de l'observation, à cause de la rotation de la terre.
Il y a autant de bases que de façons de choisir 2 points parmi $N_{tels}$, soit

$$N_B = \mathrm{card}(\boldsymbol{y}) = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Par exemple, les bases d'un interféromètre à 4 télescopes sont :

$$\begin{cases} \boldsymbol{u}_{12}(t) = \boldsymbol{r}_2(t) - \boldsymbol{r}_1(t) \\ \boldsymbol{u}_{13}(t) = \boldsymbol{r}_3(t) - \boldsymbol{r}_1(t) \\ \boldsymbol{u}_{14}(t) = \boldsymbol{r}_4(t) - \boldsymbol{r}_1(t) \\ \boldsymbol{u}_{23}(t) = \boldsymbol{r}_3(t) - \boldsymbol{r}_2(t) \\ \boldsymbol{u}_{24}(t) = \boldsymbol{r}_4(t) - \boldsymbol{r}_2(t) \\ \boldsymbol{u}_{34}(t) = \boldsymbol{r}_4(t) - \boldsymbol{r}_3(t) \end{cases}$$

Matriciellement, ces relations s'écrivent :

$$\boldsymbol{u}(t) = \boldsymbol{B} \cdot \boldsymbol{r}(t) \tag{1.12}$$

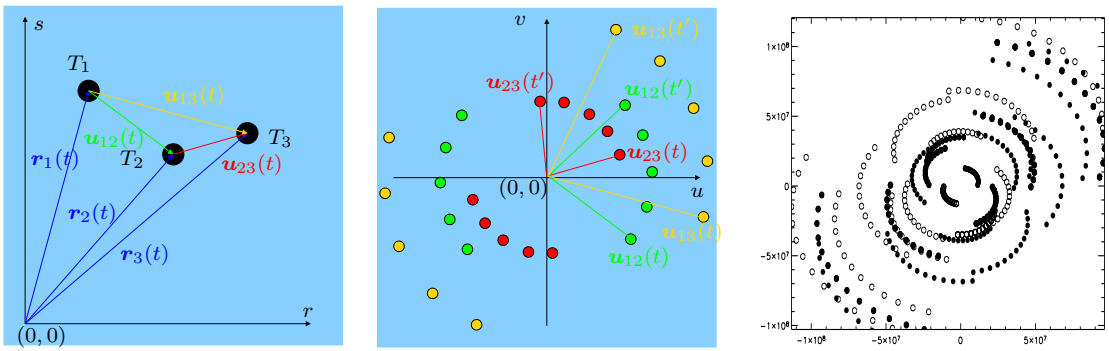
Pour $N_{tels} = 4$, l'opérateur $\boldsymbol{B}$ s'écrit :

$$\boldsymbol{B} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Les fréquences spatiales correspondantes se calculent par $\boldsymbol{\nu}(t) = \frac{\boldsymbol{u}(t)}{\lambda}$. La rotation de la terre fait que même si l'on ne change pas la configuration de l'interféromètre, les bases couvertes varient au cours du temps. Cette densification progressive de la couverture du plan $u - v$ est appelée « super-synthèse ». On obtient un ensemble de bases dans le « plan $(u, v)$ », présentant des séries de points situées sur des trajectoires elliptiques. De plus, comme l'objet $\boldsymbol{x}$ est réel positif, on a :

$$\widetilde{\boldsymbol{x}}(-\boldsymbol{\nu}) = \widetilde{\boldsymbol{x}}^*(\boldsymbol{\nu})$$

Si l'on couvre un ensemble de fréquences, on couvre donc aussi leur symétrique dans le plan fréquentiel. On indique donc dans le plan $(u, v)$ les bases symétriques de celles effectivement réalisées (1.7.b). La couverture fréquentielle obtenue a généralement l'allure de la figure (1.7.c).




(a) (b) (c)

FIG. 1.7 – (a) Positions et bases, (b) Bases et rotation de la Terre, (c) Couverture fréquentielle.

#### 1.2.1.2 Interféromètres au sol à 3 télescopes ou plus en fonctionnement

Le tableau 1.1 dresse la liste des interféromètres à 3 télescopes ou plus en fonctionnement. Ce tableau sera vite obsolète, étant donnée la rapidité avec laquelle ces projets voient le jour et évoluent. Notamment, les prochaines verront l'ajout de 4 télescopes auxiliaires sur le VLTI, et la réalisation du projet OHANA. Pour des informations actualisées et des liens vers les sites Internet de ces instruments, consulter `http://olbin.jpl.nasa.gov/`.

| Acronyme | Pays | $N_{tels}$ | Diamètre des ouvertures(cm) | Base max. (m) | $\lambda$ (microns) |
|---|---|---|---|---|---|
| ISI | US | 3 | 165 | 30 | 10 |
| COAST | UK | 5 | 40 | 22 | $0,4-0,95\,\&\,2,2$ |
| SUSI | Australie | 13 | 14 | 640 | $0,4-0,66$ |
| IOTA | US | 3 | 45 | 38 | $0,5-2,2$ |
| NPOI | US | 6 | 60 | 435 | $0,45-0,85$ |
| CHARA | US | 6 | 100 | 350 | $0,45-2,4$ |
| VLTI | Chili | 4 | 820 | 200 | $0,45-12$ |

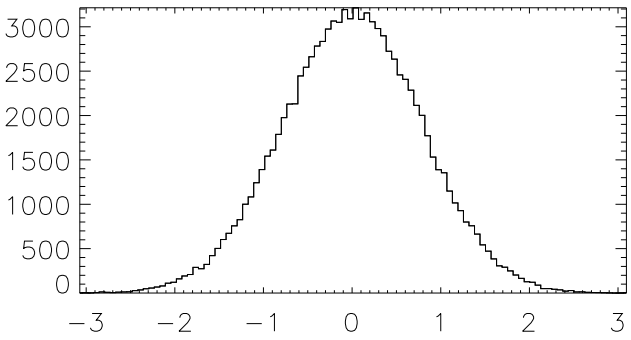
TAB. 1.1 – – Interféromètres au sol de plus de 3 télescopes

### 1.2.2 Turbulence et mesure de phase

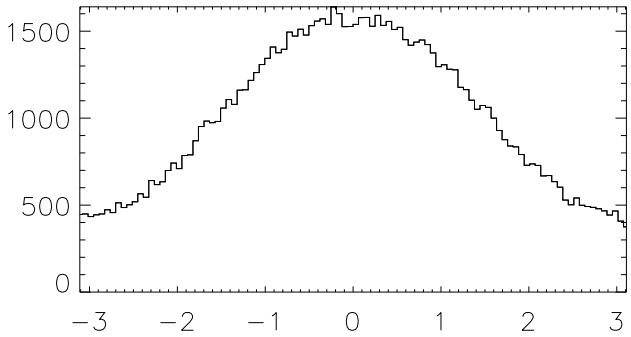
L'effet de la turbulence atmosphérique est assimilable à une différence de marche $D^{turb} = d_2^{turb} - d_1^{turb}$, fluctuant rapidement, qui vient s'ajouter à la différence de marche $D$ générée par les lignes à retard.
Ces déphasages font donc se déplacer le paquet de franges correspondant à la figure 1.6 page 22 sur l'écran, et de façon très rapide (le temps de cohérence de la turbulence est généralement de l'ordre de quelques millisecondes). L'erreur faite sur l'estimation de la phase objet est alors :

$$\varphi_2 - \varphi_1 = \frac{2\pi d_2^{turb}}{\lambda} - \frac{2\pi d_1^{turb}}{\lambda}$$

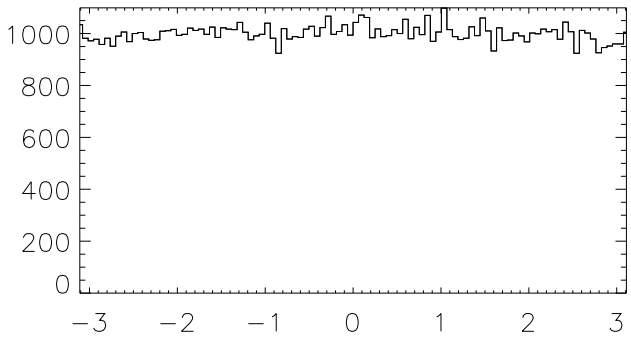
Les modèles classiques Roddier (1981) de turbulence dont on dispose permettent de modéliser la statistique des décalages de phase sur chaque télescope par des lois gaussiennes de moyennes nulles. On peut alors supposer que si l'on prend un nombre important de mesures, et que l'on en fait la moyenne dans le temps, l'influence de la turbulence sera réduite à 0.
Cependant, les décalages de phase $\varphi$ ne sont « vus » sur l'interférogramme que modulo $2\pi$. La statistique des décalages de phase effectivement constatés sur l'interférogramme est donc une loi gaussienne à moyenne nulle **repliée autour de l'intervalle** $[-\pi, \pi]$. Or une gaussienne à moyenne nulle d'écart-type supérieur à $\pi$ repliée est très proche d'une statistique uniforme, comme le montre la figure 1.8. On n'a plus aucune information statistique sur le déphasage turbulent $\varphi(\boldsymbol{r}_2) - \varphi(\boldsymbol{r}_1)$. Faire la moyenne temporelle des interférogrammes conduit donc à brouiller les franges. On ne peut donc accéder à une mesure fiable de la phase de visibilité $\phi(\boldsymbol{\nu}_{12})$ ni en faisant une moyenne temporelle, ni en exploitant des données "instantanées", ou courte pose.




Gaussienne d'écart-type $\pi/4$ modulo $2\pi$




Gaussienne d'écart-type $\pi/2$ modulo $2\pi$




Gaussienne d'écart-type $\pi$ modulo $2\pi$

FIG. 1.8 – Effet du repliement de phase.

## 1.2.3 Clôtures de phase ou phases de bispectre

### 1.2.3.1 Clôtures de phase

Il est néanmoins possible d'extraire d'interférogrammes affectés par la turbulence des informations sur la phase, si l'on dispose d'au moins 3 télescopes. Considérons un interféromètre optique à 3 télescopes (IO3T). On forme trois interférogrammes $\boldsymbol{im}_{12}$, $\boldsymbol{im}_{13}$ et $\boldsymbol{im}_{23}$. On extrait de ceux-ci les mesures suivantes :

- 3 modules de la TF de l'objet aux fréquences respectives $\boldsymbol{\nu}_{12}$, $\boldsymbol{\nu}_{13}$ et $\boldsymbol{\nu}_{23}$ dont on extrait les amplitudes de visibilité $a_{12}$, $a_{13}$ et $a_{23}$, non affectées par la turbulence
- 3 positions du maximum des franges, dont on extrait les phases de visibilités $\phi_{12}^{data}$, $\phi_{13}^{data}$ et $\phi_{23}^{data}$, reliées aux phases "vraies", ou phases objet, $\phi_{ij}^{o}$ par les relations :

$$\begin{cases} \phi_{12}^{data} = \phi_{12}^{o} + \varphi_2 - \varphi_1 \\ \phi_{13}^{data} = \phi_{13}^{o} + \varphi_3 - \varphi_1 \\ \phi_{23}^{data} = \phi_{23}^{o} + \varphi_3 - \varphi_2 \end{cases} \tag{1.13}$$

On voit facilement qu'on peut combiner ces 3 équations de façon à annuler les phases turbulentes :

$$\phi_{12}^{data} + \phi_{23}^{data} - \phi_{13}^{data} = \phi_{12}^{o} + \phi_{23}^{o} - \phi_{13}^{o}$$

Ces quantités sont appelées **clôtures de phase**, qui seront notées par la suite $\beta$ :

$$\beta_{123}^{data} = \beta_{123}^{o}$$

L'idée de former ces clôtures a été d'abord proposée par Jennison Jennison (1958) dans le contexte Radio, puis transposé au domaine Optique Jennison (1961); Rogstad (1968).

#### 1.2.3.2 Phases de bispectre

Les clôtures étant stables au cours du temps, il est avantageux de faire la moyenne de plusieurs mesures de clôtures , de façon à améliorer le rapport signal sur bruit. Cependant, faire la moyenne d'une phase modulo $2\pi$ conduit une fois encore à des erreurs d'enroulement de phase, on moyenne donc plutôt les phaseurs $\mathrm{e}^{\mathrm{i}\phi}$ que les phases $\phi$. De plus, il est avantageux de moyenner des visibilités complexes $a\mathrm{e}^{\mathrm{i}\phi}$ plutôt que les phaseurs.
Supposons que l'on prenne une première mesure parfaite $y(t_1) = 10\mathrm{e}^{\mathrm{i}\pi}$. Supposons alors qu'un problème instrumental perturbe le faisceau issu de l'un des deux télescopes, et que le contraste de l'interférogramme obtenu devienne quasi-nul. On fait deux mesures : $y(t_2) = 0,1\mathrm{e}^{2\mathrm{i}\pi}$ et $y(t_3) = 0,2\mathrm{e}^{\mathrm{i}\pi/3}$. On a :

$$\arg(y(t_1) + y(t_2) + y(t_3)) = 3,13\text{rad} \simeq \pi$$
$$\frac{1}{3}(\arg y(t_1) + \arg y(t_2) + \arg y(t_3)) = -0,7\text{rad}$$

On voit que les mesures à faible contraste ont perturbé la moyenne des phases, alors que la phase du complexe moyen reste stable. Ici, on moyenne le **bispectre**, défini par $\widetilde{o}(\boldsymbol{\nu}_{12})\widetilde{o}(\boldsymbol{\nu}_{23})\widetilde{o}^{*}(\boldsymbol{\nu}_{13})$, dont la phase est égale à la clôture de phase évoquée précédemment. Les clôtures de phase moyennes mesurées sont donc en fait les **phases de bispectres moyens**.

### 1.2.4 Méthodes d'estimation des phases de visibilité à partir des clôtures

Il existe cependant des cas particuliers pour lesquels il est possible de retrouver les phases de visibilité à partir des clôtures :

#### 1.2.4.1 Réseaux redondants, ou « Bootstrapping »

Si l'on mesure plusieurs fois la même fréquence spatiale, on peut diminuer le nombre d'inconnues en phase. Par exemple, supposons que l'on prenne une première mesure avec un IO3T, dont les télescopes ont pour positions projetées sur l'axe d'observation $\mathbf{0}$, $\boldsymbol{r}$ et $2\boldsymbol{r}$. Les fréquences spatiales mesurées sont $\frac{\vec{u}}{\lambda}$, $\frac{\vec{u}}{\lambda}$, $\frac{2\vec{u}}{\lambda}$. On peut mesurer une clôture de phase, mais seules 2 phases objet sont à déterminer. La redondance diminue le rapport nombre d'inconnues sur nombre de mesures. En l'occurrence, si l'on connaît la phase de l'objet à la fréquence $\frac{\vec{u}}{\lambda}$, la phase de l'objet à la fréquence **double** est connue elle aussi. Si l'instant d'après on déplace le télescope 3 de la position $2\boldsymbol{r}$ à $3\boldsymbol{r}$, les fréquences spatiales mesurées sont $\frac{\vec{u}}{\lambda}$, $\frac{2\vec{u}}{\lambda}$, $\frac{3\vec{u}}{\lambda}$. Les

phases correspondant à 2 de ces 3 fréquences ont déjà été mesurées. On mesure une clôture, et il n'y a qu'une inconnue à déterminer. On peut donc déterminer la phase de l'objet à la fréquence **triple** de celle fixée arbitrairement. En procédant de proche en proche, on peut mesurer les phases correspondant à une division régulière d'un axe. Si l'on procède de même dans d'autres directions, on disposera d'une couverture régulière du plan $(u, v)$, avec autant d'informations manquantes à introduire que de directions.
Cette technique, appelée aussi "bootstrapping", permet une couverture régulière du plan $(u, v)$ avec un nombre faible d'informations manquantes à introduire. On peut ainsi obtenir avec un IO3T le même type d'informations qu'avec un IONT. Cependant, il est nécessaire de mesurer plusieurs fois les mêmes fréquences. Ceci est coûteux en temps. En d'autres termes, à temps d'observation fixé, il est possible d'augmenter le degré d'information de phase, mais au prix d'une couverture fréquentielle plus pauvre.

#### 1.2.4.2 Interférométrie avec des réseaux denses de télescopes

Cette redondance de la couverture du plan fréquentiel, c'est-à-dire le fait que la même phase objet intervienne dans plusieurs clôtures, apparaît naturellement lorsque l'on dispose d'un interféromètre à plus de 3 télescopes. Lorsque l'on dispose, par exemple, de 10 télescopes reliés en réseau, à chaque instant on mesure 36 clôtures pour 45 phases objet à déterminer. Le rapport mesures/inconnues est alors de 80%, au lieu d'un tiers pour un interféromètre à 3 télescopes(cf. section 3.2.1.3 page 78).

#### 1.2.4.3 Référence de phase

Si on pointe une source connue, on peut déterminer la turbulence affectant l'instrument à l'instant du pointage et dans la direction de pointage. Si l'on pointe l'objet d'intérêt simultanément, ou après un délai inférieur au temps de cohérence de la turbulence, et si l'objet est dans le domaine isoplanétique de la source connue, alors cette dernière sert de calibrateur en phase pour l'objet : la turbulence et les différences de marche instrumentales étant connues, on peut retrouver les phases objet à partir des phases corrompues par la turbulence. La difficulté consiste à trouver de tels calibrateurs en phase proche de l'objet d'intérêt.
On utilise en général des sources non résolues. Pour plus de détails, l'on pourra consulter l'introduction à la clôture de phase de Monnier (1999).

### 1.2.5 Mise en équations des mesures interférométriques

#### 1.2.5.1 De l'objet aux visibilités

La distribution angulaire d'intensité est notée $\boldsymbol{o}(\boldsymbol{\xi})$. Il n'est possible de reconstruire, au mieux, qu'une version discrétisée de l'objet, c'est-à-dire décomposée sur une base de fonctions $\{f_k\}$ : $\boldsymbol{o}(\boldsymbol{\xi}) \simeq \sum_k x_k\, f(\boldsymbol{\xi} - \boldsymbol{\xi}_k)$. La TF de l'objet s'écrit alors :

$$\widetilde{\boldsymbol{o}}(\boldsymbol{\nu}) \simeq \mathrm{TF}\left(\sum_k x_k\, f(\boldsymbol{\xi} - \boldsymbol{\xi}_k)\right) = \sum_k x_k\, \widetilde{f}(\boldsymbol{\nu}) \mathrm{e}^{-2\mathrm{i}\pi \boldsymbol{\nu}\cdot\boldsymbol{\xi}_k} \tag{1.14}$$

Nous avons choisi une fonction $f$ telle que $\widetilde{f}(\boldsymbol{\nu})$ soit égal à 1 pour toutes les fréquences spatiales couvertes par l'interféromètre :

$$\widetilde{f}(\boldsymbol{\nu}) = \begin{cases} 1 \text{ si} \quad |\nu_x| \leq |\boldsymbol{\nu}|_{max} \quad \text{et} \quad |\nu_y| \leq |\boldsymbol{\nu}|_{max} \\ 0 \text{ sinon,} \end{cases} \tag{1.15}$$

où $|\boldsymbol{\nu}|_{max}$ est le plus grand module des fréquence spatiales couvertes par l'instrument. On en déduit :

$$f(\boldsymbol{\xi}) = f(\xi, \zeta) = \text{sinc}\left(2\pi\frac{\xi}{\Delta}\right)\text{sinc}\left(2\pi\frac{\zeta}{\Delta}\right),$$

avec $\Delta = \frac{1}{2|\boldsymbol{\nu}|_{max}}$. Une étude plus approfondie de l'influence du choix d'une base adaptée reste à faire (voir Lannes *et al.* (1994) pour une proposition de base particulière).

On note $y(\boldsymbol{\nu})$ la TF de l'objet. On a, avec les équations 1.14 et 1.15 :

$$y(\boldsymbol{\nu}_i) = \sum_k x_k \mathrm{e}^{-2\mathrm{i}\pi\boldsymbol{\nu}_i\cdot\boldsymbol{\xi}_k}.$$

Si on note $\boldsymbol{H}$ la matrice des coefficients $h_{ik} = \mathrm{e}^{-2\mathrm{i}\pi\boldsymbol{\nu}_i\cdot\boldsymbol{\xi}_k}$, alors la relation entre $\boldsymbol{x}$ et $\boldsymbol{y}$ s'écrit :

$$\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{H}\,\boldsymbol{x}$$

$\boldsymbol{y}$ est un vecteur de **visibilités complexes**.

– le vecteur des **modules de visibilité** ou **amplitudes** de la TF de l'objet s'écrit[2]

$$\boldsymbol{a}(\boldsymbol{x}) = |\boldsymbol{y}(\boldsymbol{x})| = |\boldsymbol{H}\,\boldsymbol{x}|$$

– le vecteur des **phases de visibilité** ou phases de la TF de l'objet s'écrit

$$\boldsymbol{\phi}(\boldsymbol{x}) = \arg\boldsymbol{y}(\boldsymbol{x}) = \arg(\boldsymbol{H}\,\boldsymbol{x})$$

#### 1.2.5.2 Des bases aux clôtures

Considérons un interféromètre à 4 télescopes. Si les phases turbulentes $\boldsymbol{\varphi}$ affectent les phases mesurées de la façon suivante :

$$\begin{cases} \phi_{12}^{data} = \phi_{12}(\boldsymbol{x}) + \varphi_2 - \varphi_1 \\ \phi_{13}^{data} = \phi_{13}(\boldsymbol{x}) + \varphi_3 - \varphi_1 \\ \phi_{14}^{data} = \phi_{14}(\boldsymbol{x}) + \varphi_4 - \varphi_1 \\ \phi_{23}^{data} = \phi_{23}(\boldsymbol{x}) + \varphi_3 - \varphi_2 \\ \phi_{24}^{data} = \phi_{24}(\boldsymbol{x}) + \varphi_4 - \varphi_2 \\ \phi_{34}^{data} = \phi_{34}(\boldsymbol{x}) + \varphi_4 - \varphi_3 \end{cases} \tag{1.16}$$

Sous forme vectorielle, ces équations s'écrivent :

$$\boldsymbol{\phi}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi} \tag{1.17}$$

Le but de former des clôtures de phase est d'éliminer l'influence des phases turbulentes. La clôture $\beta_{123}$ s'écrit :

$$\phi_{12}^{data} + \phi_{23}^{data} - \phi_{13}^{data} = [1, -1, 0, 1, 0, 0]\boldsymbol{\phi}^{data}$$

Si on nomme $\boldsymbol{C}$ l'opérateur $[1, -1, 0, 1, 0, 0]$, on retrouve bien :

$$\boldsymbol{C}\boldsymbol{\phi}^{data} = \boldsymbol{C}\boldsymbol{\phi}^{o} + \boldsymbol{C}\boldsymbol{B}\boldsymbol{\varphi} = \boldsymbol{C}\boldsymbol{\phi}^{o}$$

[2]Les opérateurs $|.|$ et arg sur des vecteurs sont appliqu'és point à point.

En effet, l'opérateur de clôture[3] $\boldsymbol{C}$ est construit précisément de façon à avoir :

$$\boldsymbol{C}\boldsymbol{B} = \boldsymbol{0} \tag{1.18}$$

On peut former autant de clôtures de phase que de triplets de télescopes. Cependant, la clôture $\beta_{234}$ s'écrit comme la somme $\beta_{123} - \beta_{124} + \beta_{134}$. On ne voudrait considérer qu'un ensemble de clôtures **indépendantes**. On démontre (Cf Annexe B page 157) que l'on a un ensemble maximal de clôtures indépendantes en considérant toutes les clôtures incluant le télescope situé en $\boldsymbol{r}_1$. L'opérateur $\boldsymbol{C}$ correspondant est de la forme suivante pour 4 télescopes :

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

L'ensemble des définitions et notions à propose des phases et clôtures utilisées dans ce manuscrit sont reprises chapitre 3. Le vecteur des clôtures mesurées s'écrit :

$$\boldsymbol{\beta}(\boldsymbol{x}) = \boldsymbol{C}\,\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{C}\,\arg(\boldsymbol{H}\,\boldsymbol{x})$$

#### 1.2.5.3 Mesures interférométriques et turbulence

Dans le cas d'interféromètres spatiaux ou d'interféromètres à référence de phase (Cf section1.2.4.3), l'effet de la turbulence est soit suffisamment réduit pour être assimilé à un bruit de mesure, soit mesurable et corrigeable. Dans ce cas, les observables sont des phases et modules de visibilité. L'instant de mesure $t$ peut dans ce cas ne pas être précisé, car toutes les mesures jouent le même rôle.
Par contre, dans le cas de la formation de clôtures de phase, l'opérateur de clôture $\boldsymbol{C}$ est défini pour un ensemble de bases *simultanées*. Il est donc important de mentionner l'instant de mesure correspondant à chaque mesure.

On distinguera donc deux cas :

**Turbulence faible ou corrigée :** on rassemble toutes les fréquences spatiales couvertes à chaque instant en un vecteur $\boldsymbol{\nu} = \{\boldsymbol{\nu}(t_1), \ldots, \boldsymbol{\nu}(t_n)\}$, et on calcule $\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{\nu})$. Les observables sont alors

- Les amplitudes de visibilités, affectées d'un bruit $\boldsymbol{a}_b$ :

$$\begin{aligned} \boldsymbol{a}^{data} &= \boldsymbol{a}(\boldsymbol{x}) + \boldsymbol{a}_b \\ &= |\boldsymbol{H}\,\boldsymbol{x}| + \boldsymbol{a}_b \end{aligned}$$

- Les **phases** de visibilités, affectées d'un bruit $\boldsymbol{\phi}_b$ :

$$\begin{aligned} \boldsymbol{\phi}^{data} &= \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\phi}_b \\ &= \arg(\boldsymbol{H}\,\boldsymbol{x}) + \boldsymbol{\phi}_b \end{aligned}$$

C'est le cas traité au chapitre 2.

**Turbulence forte :** à chaque instant, les fréquences spatiales couvertes sont les $\boldsymbol{\nu}(t)$, et on calcule $\boldsymbol{H}(t) = \boldsymbol{H}(\boldsymbol{\nu}(t))$. On forme des clôtures de phase, et les observables à chaque instant $t$ sont

[3] Une définition rigoureuse de cette opérateur est présentée en annexe B

– Les amplitudes de visibilités, affectées d'un bruit $\boldsymbol{a}_b$ :

$$\begin{aligned}\boldsymbol{a}^{data}(t) &= \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}_b(t)\\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}| + \boldsymbol{a}_b(t)\end{aligned}$$

– Les **clôtures** de phase, affectées d'un bruit $\boldsymbol{\beta}_b$ :

$$\begin{aligned}\boldsymbol{\beta}^{data}(t) &= \boldsymbol{\beta}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \boldsymbol{\beta}_b(t)\end{aligned}$$

C'est le cas traité aux chapitres 3, 4 et 5.

On indicera, pour une variable $v$, $v_{data}$ ou $v_d$ la quantité mesurée, et $v_b$ le bruit sur cette mesure.

## 1.3 Méthodes classiques de traitement de données interférométriques

### 1.3.1 Sous-détermination du problème

Étant donnés :
– le nombre restreint d'interféromètres optiques au sol à plus de 3 télescopes,
– le nombre d'astronomes demandant l'accès à ces interféromètres,
– les aléas météorologiques limitant le nombre de nuits d'observations,
il est illusoire d'espérer avoir plus de quelques centaines de points de Fourier mesurés sur une cible avec un interféromètre optique au sol. Or on cherche à reconstruire une image contenant typiquement plusieurs milliers de pixels. Même dans le cas simple où l'on s'est affranchi de l'effet de la turbulence, le problème de synthèse de Fourier, c'est-à-dire de reconstruction d'image à partir de quelques échantillons de sa TF, est donc extrêmement **sous-déterminé**.
La seule manière de traiter la sous-détermination consiste à introduire des modèles *a priori* de l'objet, de sorte à réduire le nombre de paramètres indépendants à estimer, et à obtenir de nouveau un problème correctement déterminé, selon un procédé que l'on appelle la régularisation du problème. Bien entendu, l'ajout de ces informations *a priori* est délicate, puisqu'il faut éviter de biaiser les données vers des solutions fausses. Pour autant, il est rare qu'un objet astronomique sur lequel on pointe soit totalement inconnu. Nous présentons dans la suite diverses approches proposées pour introduire ces modèles *a priori* dans le procédé de reconstruction de l'objet, en s'inspirant en particulier de la radio-astronomie qui a été pionnière dans la reconstruction d'images à partir de données interférométriques.
Ensuite, nous présentons l'approche bayésienne, qui permet de reformuler dans un cadre d'inférence commun la plupart des méthodes de reconstruction régularisées.

### 1.3.2 Méthodes paramétriques

L'interférométrie moderne a vu le jour avec les travaux de Michelson en 1890, qui a mesuré le diamètre de Jupiter en combinant interférométriquement les faisceaux optiques issus de deux télescopes. Sur un tel interféromètre, la formation de clôtures de phase proposée par Jennison (1958) est impossible, et l'on ne dispose que d'informations sur les **amplitudes** de visibilités. Sans informations de phase, il reste possible d'estimer des grandeurs comme des diamètres stellaires, en utilisant un modèle contraint de l'objet.

Ces méthodes paramétriques, dont la mesure du diamètre de Jupiter par Michelson en 1890 constitue un exemple historique, sont encore largement utilisées en astronomie[Dyck (1999)] pour estimer notamment :

- les masses stellaires d'étoiles binaires ;
- les effets d'assombrissement centre-bord, dus à l'éventuelle atmosphère enveloppant une source ;
- la structure d'une enveloppe circumstellaire.

Les travaux de Perrin *et al.* (2004) sont un exemple récent d'estimation des paramètres d'un modèle laminaire d'étoiles de type Mira. Cependant, la reconstruction obtenue sera évidemment extrêmement sensible aux erreurs de modèle.
La construction d'interféromètres optiques mesurant des clôtures de phase permet aujourd'hui de faire de l'imagerie non-paramétrique en optique, même si l'utilisation de modèles *a priori*, bien que moins contraints, reste nécessaire. La construction d'une telle méthode d'imagerie constitue le sujet de ce manuscrit.

### 1.3.3 Pondération et « gridding » en radio-astronomie

Une méthode d'imagerie intuitive consiste à construire une carte de la TF de l'objet d'intérêt à partir des mesures de visibilités, de valeur nulle pour les fréquences non mesurées, et à faire une TF inverse de cette carte. Le problème étant sous-déterminé, cette TF inverse est en réalité une inverse généralisée de la TF discrète. Cette technique n'est applicable que lorsque l'on dispose à la fois des phases et modules pour chaque visibilité mesurée, et que les données sont suffisamment nombreuses. Elle est par conséquent généralement utilisée en interférométrie radio.
En radio-astronomie, on procède toujours à un ré-échantillonnage des données sur une grille [Schwab (1983)], ou « *gridding* », et simultanément à une pondération des données[Mugnier & Le Besnerais (2001)]. Le ré-échantillonnage permet l'utilisation de la transformation de Fourier rapide, indispensable en radio-astronomie où le nombre de données est très important.
La pondération des données se fait en général en deux étapes :

- Une première pondération permet de tenir compte du rapport signal sur bruit (RSB) sur chaque visibilité – en diminuant l'influence de données à faible RSB – et de compenser l'inhomogénéité de la couverture fréquentielle – en diminuant l'influence relative des zones où la densité de données est élevée ;
- Un fenêtrage, c'est-à-dire une sous pondération des données haute fréquence, a pour objectif de lisser la reconstruction.

L'image obtenue par TF inverse des données ré-échantillonnées et pondérées est appelée « **Dirty Map** ».

### 1.3.4 Une méthode de référence en reconstruction interférométrique : Clean

La méthode CLEAN, proposée pour la première fois par Hogbom (1974), est probablement la plus populaire en radio-astronomie. Elle s'appuie sur l'*a priori* que la scène astronomique à reconstruire est essentiellement un ensemble de points sources. La Dirty Map est alors interprétée comme l'image résultant d'une distributions de points sources observée par l'interféromètre, à laquelle s'additionne un carte dite « résiduelle », comprenant bruit de mesure et erreur de modèle.

Högbom définit deux notions utiles pour expliquer le fonctionnement de CLEAN :

**Le « Dirty Beam » (DB)** Il est défini comme la transformée de Fourier inverse d'une fonction $g(\nu)$ valant 1 pour les fréquences mesurées et 0 ailleurs. C'est donc l'image qui serait formée au foyer de l'interféromètre s'il était dirigé vers un objet ponctuel, centré. En d'autres termes, c'est la Fonction d'Etalement de Point (FEP) ou la réponse impulsionnelle de l'instrument : $DB = TF^{-1}(g)$. Le DB est généralement constitué d'un lobe principal, de forme proche du Clean Beam décrit dans le paragraphe suivant, et de plusieurs lobes secondaires d'intensité inférieure.

**Le « Clean Beam » (CB)** La définition fournie par Högbom n'en est pas très précise. On peut le voir[Mugnier & Le Besnerais (2001)] comme la FEP d'un instrument à pupille monolithique, dont le diamètre $D$ est donné par la fréquence maximale mesurée $f_{max}$ par l'interféromètre : $D = 2\lambda f_{max}$. C'est une réponse impulsionnelle idéale.

L'algorithme CLEAN a la structure suivante :

**Initialiser** la reconstruction $\hat{\boldsymbol{x}}_{\mathrm{CLEAN}}$ à 0 et la carte résiduelle (CR) à la Dirty Map.

**Calculer** la position du maximum de la carte résiduelle $I_{\max} = \max_{l,m} CR[l,m]$ et sa position $[l_{\max}, m_{\max}] = \arg\max CR[l,m]$ .

**Tant que** $I_{\max}$ est supérieur au niveau de bruit $\sigma$, **faire**

. Supposer que le maximum de la carte résiduelle correspond à un point source.

. Soustraire à la carte résiduelle une fraction $\gamma$ de l'image de ce point source, c'est-à-dire d'un Dirac convolué par le Dirty Beam :

$$CR \leftarrow DM - \frac{\gamma I_{\max}}{\max_{l,m} DB[l,m]} DB * \boldsymbol{\delta}(l_{\max}, m_{\max}).$$

. Ajouter à la reconstruction la même fraction $\gamma$ de l'image « idéale » de ce point source, c'est-à-dire d'un Dirac convolué par le Clean Beam :

$$\hat{\boldsymbol{x}}_{\mathrm{CLEAN}} \Leftarrow \hat{\boldsymbol{x}}_{\mathrm{CLEAN}} + \frac{\gamma I_{\max}}{\max_{l,m} DB[l,m]} CB * \boldsymbol{\delta}(l_{\max}, m_{\max}).$$

. Calculer $I_{\max} = \max_{l,m} CR[l,m]$ et $[l_0, m_0] = \arg\max CR[l,m]$ .

**Fin Tant que**

**Ajouter** la carte résiduelle à la reconstruction.

Cet algorithme de référence en radio-interférométrie a donné lieu à de nombreuses implantations[Guilloteau (1998a)], dont le Multi-Résolution CLEAN(MRC) de Wakker & Schwarz (1988), qui reconstruit séparément une carte de sources étendues et une carte de points sources[4]. On peut cependant noter des limitations importantes de cette méthode :

- La solution obtenue **n'est pas compatible avec les données**Guilloteau (1998a). En effet, c'est une carte de points source qui est compatible avec les données à la carte résiduelle près, et non la carte de Diracs convolués avec le Clean Beam ;
- La convolution par le Clean Beam revient à une multiplication dans l'espace de Fourier par un profil décroissant. Cette apodisation des données vient se superposer aux opérations préalables de pondération et de fenêtrage, sans une interprétation globale du pré-traitement opéré sur les données.

[4]Des développements récents en reconstruction multi-résolution de données interférométriques ont été proposés par Giovannelli & Coulais (2005), incluant une revue bibliographique des différentes hypothèses faites en interférométrie sur la structure de l'objet à reconstruire.

– Dans la méthode d'origine proposée par Högbom, il n'est pas fait mention d'un modèle explicite du bruit, par l'intermédiaire duquel nous pourrions prendre en compte les spécificités importantes du cadre interférométrique optique dans lequel nous nous plaçons.

Ce dernier point est abordé dans les travaux de Anterrieu (1998) et Lannes *et al.* (1997), qui proposent WIPE, une implantation moderne de l'algorithme CLEAN, incluant une évaluation *a posteriori* du résultat et un choix de pondération des données basé sur le conditionnement de la matrice du problème. Comme nous le verrons au chapitre 2, cependant, WIPE n'est pas directement applicable au contexte interférométrique optique.

### 1.3.5 Méthode du maximum d'entropie (MEM)

Parallèlement à CLEAN, qui suppose un objet principalement composé de points sources, la méthode du maximum d'entropie (MEM) suppose un modèle de sources lisses et positives. Cette méthode, très populaire en radio-astronomie, consiste à reconstruire l'image la plus douce possible, c'est-à-dire d'entropie maximale, qui soit cohérente avec les données (étant donné le nombre important de travaux sur ces méthodes , il existe une grande variétés d'interprétations de l'entropie. Les auteurs Nityananda & Narayan (1982) proposent une analyse de ces différentes approches, en fonction de l'effet produit par le terme entropique sur la carte reconstruite). Généralement, pour déterminer cette image optimale, on minimise une **fonction de coût** composite, dont une partie $\mathcal{J}^{data}$ mesure l'écart de la reconstruction aux données mesurées, et dont l'autre terme pénalise les solutions de faible entropie :

$$\mathcal{J}(\boldsymbol{x}) = \mathcal{J}^{data}(\boldsymbol{x}) - \text{Entropie}(\boldsymbol{x}) = \mathcal{J}^{data}(\boldsymbol{x}) - \sum_k f(x_k)$$

Le terme d'entropie contient les informations *a priori* que l'on a sur l'objet à reconstruire, et permet de « régulariser » le problème inverse sous-déterminé à résoudre (Voir Titterington (1985), Idier (2001) et Demoment (1989) pour plus d'informations sur la régularisation). Il en existe plusieurs définitions mathématiques de l'entropie, les plus populaires étant :

$$f_1(x_k) = -x_k \ln x_k$$
$$f_2(x_k) = \ln x_k$$

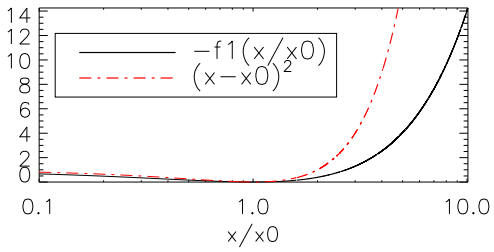
Ces régularisations entropiques attribuent en particulier un coût tendant vers l'infini aux pixels de l'objet se rapprochant de 0. Elles imposent donc implicitement la **positivité** de l'objet, qui est un *a priori* indéniable et fortement informatif sur l'objet.
Pour des *a priori* à pixels indépendants, tels que l'entropie, on peut exprimer un terme de régularisation comme un « ressort de rappel » vers un objet *a priori* $\boldsymbol{x}^0$. Par exemple, la fonction $f_1(x_k)$ est modifiée comme suit :

$$f_1(x_k) = -x_k \ln\left(\frac{x_k}{x^0}\right) + x_k - x_k^0$$

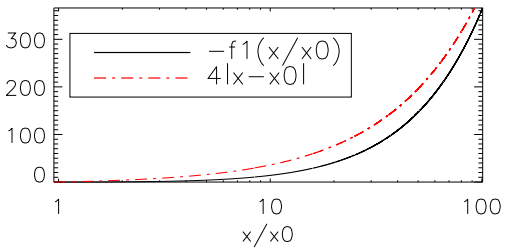
On pose $x_k = x_k^0 + \delta x$.

– Pour $\delta x$ petit, un équivalent de $f_1(x_k)$ est $-(\delta x)^2/x_k^0$. Le rappel vers la valeur *a priori* du pixel est **quadratique**. En d'autres termes, dans un voisinage de l'objet *a priori*, la régularisation entropique se comporte comme un terme $\|\boldsymbol{x} - \boldsymbol{x}^0\|^2$ (Cf. figure1.3.5.a).
– Pour $\delta x$ grand, un équivalent de $f_1(x_k)$ est $-\delta x \ln\left(\frac{\delta x}{x^0}\right)$. Cet équivalent est négligeable devant $-(\delta x)^p$, $\forall p > 1$. Le comportement de la fonction de régularisation devient donc quasiment linéaire, c'est-à-dire que la régularisation entropique se comporte comme un terme $\|\boldsymbol{x} - \boldsymbol{x}^0\|$ (Cf. figure1.3.5.b).

Les auteurs Nityananda et Narayan analysent les reconstructions MEM par rapport à leur capacité à adoucir le fond de l'image, et à faire ressortir les pics. Cette capacité est attribué à un variation de pente rapide proche de la valeur de rappel, puis plus lente loin de cette valeur. Avec cette approche, toute fonction $f$ ayant ces caractéristiques correspond à une régularisation entropique, et les fonctions classiques $f_1$ et $f_2$ n'en sont qu'un exemple parmi d'autres.




(a) comportement quadratique proche de $x_0$




(b) comportement linéaire loin de $x_0$

FIG. 1.9 – Régularisation entropique.

Pour d'autres informations sur les méthodes MEM, voir notamment Skilling & Bryan (1984), en pour les applications à l'astronomie, Narayan & Nityananda (1986), Le Besnerais *et al.* (1991) et Le Besnerais (1993a).

L'utilisation d'un critère composite, constitué d'un terme d'attache aux données et d'un terme de régularisation *a priori* s'interprète en terme probabilistes dans le cadre de l'inférence bayésienne (cf. section suivante), qui généralise la plupart des méthodes évoquées précédemment et offre les réponses les plus complètes à des problèmes laissés ouverts dans les autres approches, comme le choix d'hyperparamètres (qui règlent en particulier les poids relatifs des termes d'attache aux données et de régularisation) ou l'optimisation de critère multimodaux [Demoment & Goussard (2001)], qui concerne la majeure partie de ce manuscrit.

## 1.4 Inférence bayésienne

Cette section rappelle quelques rudiments d'inférence bayésienne utilisés dans ce manuscrit, et s'inspire largement de [Demoment & Goussard (2001)].

### 1.4.1 Base théorique

On cherche à relier la probabilité du produit logique $AB$, c'est-à-dire de l'évènement « $A$ et $B$ », aux probabilités séparées de $A$ et de $B$. On raisonne sur les probabilités conditionnelles $p(A|C)$, $p(B|C)$ et $p(AB|C)$, $C$ décrivant l'ensemble des informations que l'on a sur le problème à résoudre. Les deux principales lois utilisées sont :

– la règle du produit :

$$p(AB|C) = p(A|BC)p(B|C) = p(B|AC)p(A|C); \tag{1.19}$$

– la règle de Bayes qui s'en déduit :

$$p(B|AC) = \frac{p(A|BC)p(B|C)}{p(A|C)}. \tag{1.20}$$

Lorsque l'on doit résoudre un problème d'inférence, on dispose :

– de données $D$
– d'informations générales $I$ sur le problème (par exemple d'un modèle de bruit)

et on cherche à déterminer un ensemble de paramètres inconnus $x$.
La démarche consiste à construire une loi de probabilité $p(x|DI)$ sur $x$ tenant compte des données $D$ et des informations générales $I$, appelée loi *a posteriori*.

En appliquant la règle de Bayes, on a :

$$p(x|DI) = \frac{p(D|xI)p(x|I)}{p(D|I)}$$

$p(D|I)$ est une constante vis-à-vis de $x$. L'équation précédente s'écrit donc :

$$p(x|DI) \propto p(D|xI)p(x|I)$$

**Vraisemblance** $p(D|xI)$ Le plus souvent, on dispose d'un modèle de formation des données, c'est-à-dire que l'on sait évaluer la probabilité d'obtenir les données $D$ connaissant $x$. Ce modèle fait partie des informations générales $I$. En d'autres termes, on connaît la probabilité $p(D|xI)$, ou probabilité *directe*. Lorsqu'on étudie la dépendance de cette loi vis-à-vis de $x$, cette loi est appelée *vraisemblance*. En effet, les inconnues $x_{MV}$ les plus vraisemblables connaissant les données $D$ sont celles pour lesquelles la probabilité directe $p(D|x_{MV}I)$ sera maximale. $x_{MV}$ est l'estimation de $x$ au sens du maximum de vraisemblance :

$$\hat{x}_{MV} = \arg\max p(D|xI)$$

**Loi *a priori*** $p(x|I)$ D'autre part, on peut disposer d'une information *a priori* sur $x$, c'est-à-dire d'une loi de probabilité $p(x|I)$. Par exemple, le fait que les pixels de l'image sont d'intensité positive correspond à :

$$\begin{cases} p(x|I) = 0 & \text{pour } x \in \mathbb{R}^{-*}; \\ p(x|I) = \text{cste} & \text{pour } x \in \mathbb{R}^{+}. \end{cases} \tag{1.21}$$

Si l'on n'a aucune information *a priori*, cela signifie que toutes les valeurs de $x$ sont équiprobables, et donc que $p(x|I) = \text{cst}$.
La loi *a priori* n'est pas une distribution de probabilité au sens strict du terme, basée sur un grand nombre de réalisations de $x$. Elle est une traduction mathématique des informations *a priori* dont on dispose, c'est-à-dire qu'elle prend des valeurs d'autant plus faibles que $x$ est incompatible avec les *a priori*.

#### 1.4.1.1 L'estimateur du maximum *a posteriori* (MAP)

La loi *a posteriori* résume l'ensemble des connaissances sur $x$ à l'issue de la prise de mesures. On l'utilise pour définir un estimateur de $x$, le plus classique étant l'estimateur du maximum *a posteriori* (MAP), qui consiste à choisir $x$ maximisant la loi *a posteriori* :

$$\hat{x}_{MAP} = \arg\max p(x|DI) = \arg\max p(D|xI)p(x|I)$$

Lorsqu'on ne dispose pas d'*a priori*, c'est-à-dire que $p(x|I) = \text{cst}$, alors les estimations du maximum *a posteriori* et du maximum de vraisemblance sont identiques.

Les lois statistiques utilisées dans la pratique sont souvent exponentielles. Il est donc pratique de poser $\mathcal{J}(x) = -\log p(x|DI)$. On a alors :

$$\begin{aligned}
\hat{x}_{MAP} &= \arg\max p(x|DI) \\
&= \arg\min(-\log p(x|DI)) \\
&= \arg\min \mathcal{J}(x) \\
&= \arg\min \left[-\log(p(D|xI)p(x|I)) + cst\right] \\
&= \arg\min \left[\underbrace{-\log(D|xI)}_{\mathcal{J}_{data}} \underbrace{-\log p(x|I)}_{\mathcal{J}_{prior}}\right] \\
\hat{x}_{MAP} &= \arg\min \mathcal{J}(x) = \arg\min \left[\mathcal{J}_{data} + \mathcal{J}_{prior}\right]
\end{aligned}$$

Pour calculer l'estimation de $x$ au sens du maximum *a posteriori*, on minimise un critère composite contenant :

**un terme d'attache aux données** $\mathcal{J}_{data}$ , appelé aussi anti-log-vraisemblance, et qui est l'opposé du logarithme de la loi de vraisemblance des inconnues sachant les données, et qui traduit donc bien l'adéquation de l'estimée aux données. Dans le cas d'un modèle additif $D = Hx + B$ avec $B$ un bruit gaussien, de moyenne $\bar{B}$ et d'écart-type $\sigma_B$, $\mathcal{J}_{data}$ s'écrit :

$$2\frac{(D - Hx - \bar{B})^2}{\sigma_B^2} = 2\chi_B^2(x)$$

et l'estimateur du maximum de vraisemblance est dans ce cas identique à l'estimateur aux moindres carrés ;

**un terme d'*a priori*** $\mathcal{J}_{prior}$ , qui est l'opposé du logarithme de la loi statistique *a priori* suivie par les inconnues. On introduit généralement un paramètre $\mu$ permettant de régler le poids relatif de l'*a priori* par rapport aux données :

$$\mathcal{J}(\boldsymbol{x}) = \mathcal{J}_{\text{data}}(\boldsymbol{x}) + \mu\mathcal{J}_{prior}(x) \tag{1.22}$$

Le MAP est souvent interprété comme un maximum de vraisemblance pénalisé, le terme d'*a priori* $\mathcal{J}_{prior}$ pénalisant les solutions de forte vraisemblance mais incompatibles avec les informations *a priori* sur $x$.
L'utilisation de critères pénalisés est une méthode classique de régularisation, à laquelle on peut arriver sans faire intervenir le cadre de l'inférence bayésienne : nous en avons vu un exemple avec les méthodes du Maximum d'Entropie (MEM). A ce stade, on peut donc réinterpréter le MEM comme une méthode bayésienne correspondant à un choix particulier de loi *a priori*, et qui tend à pénaliser les solutions de faible entropie :

$$p(\boldsymbol{x}) \propto \mathrm{e}^{\text{Entropie}(\boldsymbol{x})}. \tag{1.23}$$

Une fois le modèle direct de formation des données, et donc la vraisemblance, fixés, les différentes inférences sur $x$ se distinguent par le choix du terme de régularisation, c'est-à-dire par le type d'informations *a priori* injectées dans la reconstruction. La section suivante reprend les principales lois *a priori* utilisées en interférométrie.

### 1.4.2 Lois *a priori* classiques

Le choix de la loi *a priori* utilisée ne dépend pas seulement de sa capacité à améliorer la qualité de la reconstruction. Elle doit en outre conduire à un critère composite simple à

minimiser. Ainsi, on préférera les termes de régularisation convexes – donc à minimum unique –, et même quadratiques.
Certains *a priori* ne sont pas introduits explicitement dans l'inversion sous forme d'un terme de régularisation. C'est notamment le cas pour CLEAN (cf. section 1.3.4), où l'hypothèse d'un objet constitué de points sources est pris en compte par l'intermédiaire de l'algorithmie même, ou pour l'algorithme de Richardson-Lucy [Lucy (1974)], où la solution est régularisée par arrêt avant convergence. De la même façon, le choix d'une base de discrétisation de l'objet contient implicitement des informations *a priori*, telles que la fréquence maximale de l'objet. Cette section décrit les types principaux d'*a priori explicites* utilisés en reconstruction d'image.

#### 1.4.2.1 Contrainte de positivité

Cet *a priori* étant toujours valide, il est utilisé de façon quasi-systématique, et souvent en conjonction avec une autre régularisation. Elle est particulièrement utile sur des objets de type astronomique, c'est-à-dire à support réduit. En effet, la positivité empêche l'apparition de variations haute fréquence hors du support de l'objet (des modes haute fréquence entraîneraient l'apparition de pixels négatifs). L'utilisation de cette contrainte nécessite une bonne connaissance du flux de l'objet.

La positivité peut s'interpréter en termes statistiques comme une probabilité nulle pour les valeurs négatives des pixels, et constante pour les valeurs positives. Le terme $\mathcal{J}_{prior}$ correspondant est infini sur $\mathbb{R}^{-*}$ et constant sur $\mathbb{R}^{+}$.
Néanmoins, les méthodes de minimisation de critères étant souvent à base de descente locale, le gradient infini au passage de 0 qu'un tel terme de pénalisation entraînerait serait dommageable à la convergence de l'algorithme.
De fait, la contrainte de positivité est la plupart du temps prise en compte au niveau de l'optimisation par une modification de l'algorithme.
Il existe plusieurs techniques d'implantation de la contrainte de positivité :

**Entropie** Comme nous l'avons vu section 1.3.5, la contrainte de positivité est implicitement imposée par un critère de régularisation entropique.

**Reparamétrisation** L'objet $\boldsymbol{X}$ étant positif, il peut s'écrire par exemple $\boldsymbol{X} = \boldsymbol{a}^2$. On a alors $\frac{\partial J(\boldsymbol{X})}{\partial \boldsymbol{a}} = \frac{\partial J(\boldsymbol{X})}{\partial \boldsymbol{X}} \cdot \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{a}} = 2\boldsymbol{a}\frac{\partial J(\boldsymbol{X})}{\partial \boldsymbol{X}}$. La dérivée du critère s'annule pour tout pixel nul. Tout pixel atteignant 0 lors de la minimisation sera donc gelé. Ce défaut peut être contourné en utilisant d'autres paramétrisations, par exemple $\boldsymbol{X} = \exp \boldsymbol{a}$. Cette reparamétrisation a cependant tendance à ralentir la minimisation.

**Activation des contraintes** Cette méthode repose sur l'étude du support de l'objet positif optimal, c'est-à-dire sur les pixels non-nuls de l'objet positif optimal :
- Si le pixel $i$ de l'objet positif optimal $o^{+*}$ est non nul, alors le gradient du critère par rapport à la valeur de ce pixel est nul. Sinon, en faisant évoluer la valeur de ce pixel, tout en restant positive, on pourrait faire diminuer la valeur du critère et $o^{+*}$ ne serait pas optimal.
- Si le pixel $i$ de l'objet positif optimal $o^{+*}$ est nul, alors le gradient du critère par rapport à la valeur de ce pixel est positif ou nul. Pour faire diminuer la valeur du critère, il faudrait diminuer la valeur du pixel $i$, c'est-à-dire lui donner une valeur strictement négative.

Ces conditions, dites de Karush, Kuhn et Tucker, sont largement utilisées en optimisation mathématique (pour plus d'informations, voir annexe C page 161) .

**« Slack variables »** Une autre façon de traiter le problème est d'imposer les contraintes $\boldsymbol{X} = \boldsymbol{s}$ et $\boldsymbol{s} > 0$. Les $\boldsymbol{s}$ sont appelées « slack variables ». Leur utilisation permet d'accélérer

la convergence d'une minimisation sous contrainte de positivité (voir Giovannelli & Coulais (2005)).

Nous avons opté pour la méthode d'activation des contraintes. L'utilisation de « Slack variables » constituerait une amélioration du programme, et est une des suites envisagées à ces travaux.

#### 1.4.2.2 Régularisation quadratique ou L2

Il a déjà été mentionné qu'en inférence bayésienne, un terme de régularisation s'interprète comme une loi *a priori* sur l'objet selon l'analogie formelle :

$$\mathcal{J}_{prior}(\boldsymbol{x}) = -\log p(\boldsymbol{x})$$

Une régularisation quadratique consiste à considérer un terme de régularisation (quadratique) de la forme :

$$\mathcal{J}_{prior}(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}^0)^{\mathrm{T}} \boldsymbol{W}_{\boldsymbol{x}} (\boldsymbol{x} - \boldsymbol{x}^0)$$

Elle correspond donc à une statistique *a priori* de l'objet gaussienne :

$$p(\boldsymbol{x}) \propto \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^0)^{\mathrm{T}} \boldsymbol{W}_{\boldsymbol{x}} (\boldsymbol{x}-\boldsymbol{x}^0)}$$

où $x_k^0$ est le $k^{\mathrm{e}}$ pixel d'un objet *a priori* $\boldsymbol{x}^0$. Le caractère convexe d'un tel terme de régularisation permet de restreindre le nombre de nombre de solutions équivalentes (en terme de valeur de critère) du problème inverse considéré. La régularisation quadratique est entièrement spécifiée par le choix de l'objet *a priori*, et surtout de la matrice $\boldsymbol{W}_{\boldsymbol{x}}$, qui s'interprète comme l'inverse de la matrice de covariance sur les pixels de l'objet. Selon la forme de cette matrice, on distingue deux types de régularisation quadratique.

**Régularisation à pixels indépendants** Cette régularisation classique consiste à choisir $\boldsymbol{W}_x = \mathrm{Diag}\{w_k\}$ :

$$\mathcal{J}_{prior}(\boldsymbol{x}) = \sum_k w_k (x_k - x_k^0)^2$$

**Régularisation paramétrée par la Densité Spectrale de Puissance (DSP)** Le terme d'*a priori* précédent est sensible, dans le cas général, à une translation de l'objet. En effet, pour que cet *a priori* soit invariant par translation, il faut que $\boldsymbol{W}_x$ soit d'une structure particulière, dite « Toeplitz à blocs Toeplitz ». Pour faciliter l'implantation algorithmique, on fait l'approximation que cette matrice est circulante à blocs circulants. Avec cette approximation, $\boldsymbol{W}_{\boldsymbol{x}}$ diagonalise dans la base de Fourier :

$$(\boldsymbol{x} - \boldsymbol{x}^0)^{\mathrm{T}} \boldsymbol{W}_{\boldsymbol{x}} (\boldsymbol{x} - \boldsymbol{x}^0) = (\widetilde{\boldsymbol{x}} - \widetilde{\boldsymbol{x}}^0)^{\mathrm{T}} \widetilde{\boldsymbol{W}}_{\boldsymbol{x}} (\widetilde{\boldsymbol{x}} - \widetilde{\boldsymbol{x}}^0) = (\widetilde{\boldsymbol{x}} - \widetilde{\boldsymbol{x}}^0)^{\mathrm{T}} \mathrm{Diag}\left\{DSP^{-1}\right\} (\widetilde{\boldsymbol{x}} - \widetilde{\boldsymbol{x}}^0)$$

où DSP est la **densité spectrale de puissance de l'objet**, c'est-à-dire la répartition en fréquence de l'énergie de la TF de l'objet. Le critère *a priori* sur l'objet s'écrit donc :

$$\boxed{\mathcal{J}_{prior}(\boldsymbol{x}) = \sum_{\boldsymbol{\nu}} \frac{|\widetilde{x}(\boldsymbol{\nu}) - \widetilde{x}^0(\boldsymbol{\nu})|^2}{DSP(\boldsymbol{\nu})}}$$

Cette distribution *a priori* correspond à une distribution gaussienne du coefficient de la TF de l'objet pour **toute fréquence $\boldsymbol{\nu}$ de la grille d'échantillonnage de la TF discrète** :

$$p(\widetilde{x}(\boldsymbol{\nu})) \propto \mathrm{e}^{-\frac{1}{2}\frac{|\widetilde{x}(\boldsymbol{\nu}) - \widetilde{x}^0(\boldsymbol{\nu})|^2}{DSP(\boldsymbol{\nu})}} \tag{1.24}$$

C'est ce modèle de DSP qui joue le rôle de pondération de l'écart à l'objet moyen. La pondération des données utilisée dans l'algorithme WIPE [Lannes *et al.* (1997)] peut donc s'interpréter comme une régularisation quadratique. La plupart du temps, on cherche à préserver les basses fréquences de l'objet reconstruit, et à régulariser les hautes fréquences – qui concentrent la majorité du bruit – en les ramenant vers celles de l'objet *a priori*, c'est-à-dire généralement vers 0 (l'objet moyen est en première approche fixé à un objet constant, de flux égal au flux mesuré). Le profil de DSP doit donc être passe-bas, afin d'opérer un lissage des hautes fréquences tout en préservant les basses fréquences.
En particulier, en l'absence d'une contrainte de positivité, la pondération et le fenêtrage évoqués section 1.3.3) sont équivalents à une pénalisation quadratique du problème dans le domaine fréquentiel– résultat connu par ailleurs (mais sous forme duale) en analyse spectrale[Giovannelli & Idier (2001)]. En effet, ces opérations, comme la régularisation par DSP, modifient les poids relatifs des données en fonction de leur fréquence radiale. L'utilisation de la pondération des données pour régulariser le problème a été étudiée en particulier par Lannes *et al.* (1987).

Comme nous l'avons déjà mentionné, l'enjeu de la régularisation quadratique est moins le choix de l'objet par défaut $\boldsymbol{x}^0$ (que l'on peut généralement fixer à un objet constant) que la longueur de corrélation de l'objet, c'est-à-dire sa matrice de covariance ou son profil de DSP. Une méthode de détermination de profil de DSP sur les données est décrit section 2.4.3. L'utilisation d'un critère de régularisation quadratique conduit, dans le cadre d'un bruit additif gaussien, à un critère composite quadratique lui aussi, dont l'annulation du gradient se résout par une simple inversion linéaire (plus ou moins « rentable » selon la taille de la matrice à inverser), lorsqu'il n'y a pas de contrainte de positivité. Dans le cas contraire, la minimisation sous contrainte de positivité d'un critère quadratique reste un problème classique que l'on sait traiter efficacement.

#### 1.4.2.3 Préserver les bords francs : L1L2

Le terme de régularisation L2 décrit précédemment pénalise de façon identique un bruit haute fréquence et les éventuels bords francs de l'objet, qu'il tend donc à « adoucir ». Pour y remédier, il est possible d'utiliser une régularisation qui pénalise quadratiquement les faibles gradients, et linéairement les forts gradients [Green (1990)]. La partie quadratique assure l'atténuation du bruit tandis que le comportement linéaire pour des forts gradients permet de préserver les bords francs de l'image [Bouman & Sauer (1993)]. De tels *a priori* sont appelés « linéaires-quadratiques », ou L1L2

Laurent Mugnier *et al.*proposent un terme de régularisation linéaire quadratique isotrope [Mugnier *et al.* (2004b)] sur le gradient de l'objet :

$$\mathcal{J}_{L1L2}(\boldsymbol{x}) = \mu\delta^2 \sum_{l,m} f_{L1L2}\left(\frac{\|\nabla \boldsymbol{x}\|\,(l,m)}{\delta}\right) \tag{1.25}$$

avec

$$f_{L1L2}(a) = a - \log(1 + \|\nabla \boldsymbol{x}\|) \tag{1.26}$$

La fonction $f_{L1L2}$ est quadratique proche de 0 et linéaire loin de 0, le critère $\mathcal{J}_{L1L2}$ a donc bien le comportement recherché.

#### 1.4.2.4 *A priori* « blancs »

Cependant, une telle régularisation, en pénalisant, de façon certes linéaire, les forts gradients, lie les pixels voisins de l'objet entre eux. Si l'on souhaite ne pas empêcher l'apparition de pics isolés à la limite de résolution, *i. e.* un pixel très brillant entouré de pixels sombres, il est possible d'utiliser ce terme de régularisation L1L2 blanc, noté L1L2$_w$, sur la valeur des pixels et non sur la norme du gradient :

$$\mathcal{J}_{L1L2_w}(\boldsymbol{x}) = \sum_k f_{L1L2}(x_k) = \sum_k x_k - \log(1 + x_k)$$

Si l'on souhaite que les pixels ne soient pas ramenés vers 0 mais vers une valeur *a priori* $x_k^0$, il suffit de modifier le critère en

$$\mathcal{J}_{L1L2_w}(\boldsymbol{x}, \boldsymbol{x}^0) = \sum_k f_{L1L2}(x_k - x_k^0) = \sum_k x_k - x_k^0 - \log(1 + x_k - x_k^0)$$

Ce terme de régularisation s'interprète en inférence bayésienne comme une loi de probabilité de $\boldsymbol{x}$ de la forme :

$$p(\boldsymbol{x}) = \prod_k \mathrm{e}^{-\frac{f}{2}_{L1L2}(x_k - x_k^0)},$$

c'est-à-dire que les valeurs de deux pixels, même voisins, sont deux variables considérées indépendantes.

Des *a priori* L2 du type

$$\mathcal{J}_{prior}(\boldsymbol{x}) = \sum_k w_k (x_k - x_k^0)^2,$$

ou de positivité, sont eux aussi des *a priori* blancs, c'est-à-dire ne liant pas les pixels entre eux. Ils s'interprètent comme des « ressorts de rappel » ramenant l'objet $\boldsymbol{x}$ vers l'objet *a priori* $\boldsymbol{x}^0$, avec une raideur variable (quadratique par exemple en régularisation L2).

La régularisation entropique est un autre exemple d'*a priori* blanc, dont le comportement est proche de celui d'un L1L2$_w$. En effet, pour Nityananda & Narayan (1982), toute pénalisation de la forme

$$\mathcal{J}_{prior}(\boldsymbol{x}) = \sum_k f(x_k - x_k^0),$$

où la fonction $f$ présente une variation de pente rapide proche de la valeur de rappel, puis plus lente loin de cette valeur, peut être considérée comme entropique.

La régularisation L1L2$_w$ peut donc s'interpréter, selon la définition de Nityananda & Narayan (1982), comme une régularisation entropique. Cependant, la fonction $f_{L1L2}$ est strictement croissante, alors que la fonction $-f_1(x) = x * \log(x) - x$ utilisée généralement dans les méthodes entropiques présente un minimum (Figure 1.10). Cette dernière tend donc à ramener le fond vers une valeur non nulle, ce qui n'est pas le cas de la fonction $f_{L1L2}$.

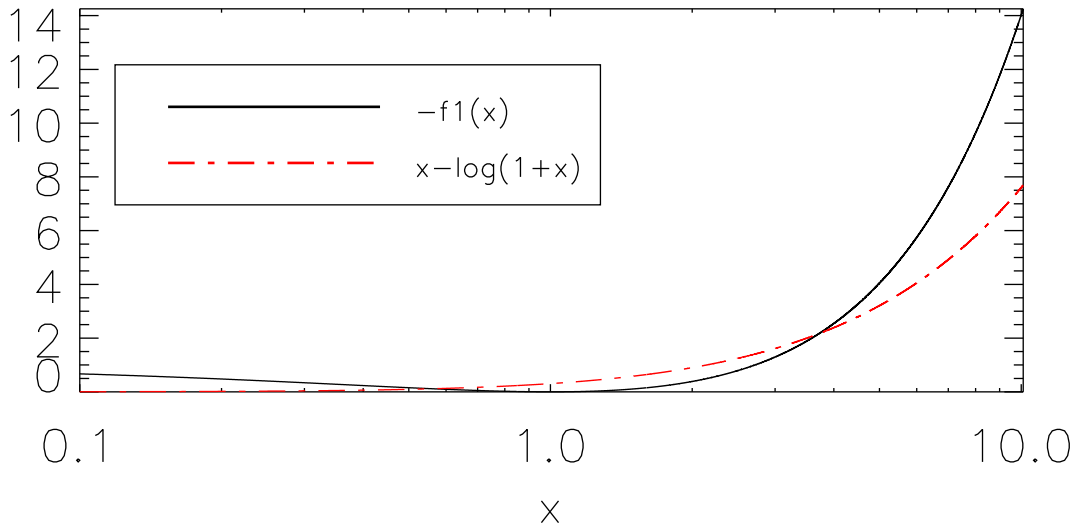



FIG. 1.10 – Comparaison des fonctions de coût $-f1$ et $f_{L1L2}$.

# Chapitre 2

# Reconstruction d'images à partir de visibilités complexes.

## 2.1 Position du problème

Dans ce chapitre, nous nous plaçons dans le cas d'une turbulence faible ou corrigée. Les observables sont

– Les amplitudes de visibilités, affectées d'un bruit $\boldsymbol{a}_\mathrm{b}$ :

$$\begin{aligned} \boldsymbol{a}^d &= \boldsymbol{a}(\boldsymbol{x}) + \boldsymbol{a}_\mathrm{b} \\ &= |\boldsymbol{H}\,\boldsymbol{x}| + \boldsymbol{a}_\mathrm{b} \end{aligned} \tag{2.1}$$

– Les **phases** de visibilités, affectées d'un bruit $\boldsymbol{\phi}_\mathrm{b}$ :

$$\begin{aligned} \boldsymbol{\phi}^d &= \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\phi}_\mathrm{b} \\ &= \arg(\boldsymbol{H}\,\boldsymbol{x}) + \boldsymbol{\phi}_\mathrm{b} \end{aligned} \tag{2.2}$$

Le problème à résoudre est un problème de synthèse de Fourier, c'est-à-dire de reconstruction de l'objet à partir d'échantillons, bruités ou pas, de sa TF. Selon le niveau de bruit sur ces données, leur nombre et la couverture fréquentielle correspondante, le problème est plus ou moins ardu. Dans le cadre de la radio-astronomie, le problème de synthèse de Fourier a été très largement traité (voir notamment Thompson *et al.* (1986)et Lannes *et al.* (1996)). En optique, la difficulté vient de l'extrême sous-détermination du problème (cf. section 1.3.1) du fait des difficultés d'acquisition évoquées au chapitre 1. En particulier, une opération de gridding (section 1.3.3) s'avère inutile (en terme de coût de calcul) et même pénalisante, puisqu'elle correspond à une approximation de la couverture fréquentielle.
Une autre spécificité de l'optique tient à un modèle de bruit en phase et module. A notre connaissance, à l'exception de l'algorithme MIRA [Thiébaut *et al.* (2003)] (décrit section 4.1), aucune méthode de synthèse de Fourier tenant compte à la fois du faible nombre de données et du bruit en phase et modules, caractéristiques de l'interférométrie optique, n'a été publiée. En particulier, les travaux de Lannes (1998a) supposent un bruit sur des visibilités complexes, et non deux bruits séparés sur phase et module.

### 2.1.1 Démarche

L'inversion bayésienne du problème nécessite d'établir un modèle de formation des données, et de choisir un critère de régularisation basé sur les informations *a priori* disponibles. De plus, lorsque l'on impose la positivité ou un *a priori* non quadratique, l'inversion ne peut pas

se faire analytiquement. Une méthode d'inversion itérative est donc nécessaire.
Notre démarche a consisté à :

- établir un modèle de formation des données (Section 2.2) ;
- déduire de ce modèle un critère à minimiser (Section 2.2.2) ;
- établir un modèle *a priori* de l'objet(Section 1.4.2) ;
- concevoir une méthode de descente du critère ainsi formé (Section 2.4.2) ;
- valider la méthode sur simulations (Section 2.5).

## 2.2 Construction du critère d'attache aux données

### 2.2.1 Modèle de bruit

Estimer la statistique du bruit sur des données interférométriques est un problème difficile, et hors du cadre de ce manuscrit. Nous avons cherché à obtenir des méthodes utilisables sur différents instruments, le modèle que nous utilisons devait donc être suffisamment général pour ne pas être restrictif. Nous avons donc considéré que seuls les écarts-type sur chaque mesure étaient estimés, et que les moyennes des bruits de mesure étaient nuls (ou que les données étaient débiaisées).
En l'absence d'autres informations, les bruits sur chaque mesure sont supposés gaussiens à moyenne nulle, et décorrélés.

### 2.2.2 Critère d'attache aux données exact $\mathcal{J}^{data}$

Les bruits $\boldsymbol{a}^b$ et $\boldsymbol{\phi}_\mathrm{b}$ sont considérés gaussiens blancs à moyenne nulle, d'écarts-type respectifs $\boldsymbol{\sigma_a}$ et $\boldsymbol{\sigma_\phi}$. Dans la suite de ce manuscrit, on note $\boldsymbol{R}_a$ (resp. $\boldsymbol{R}_\phi$) la matrice de covariance diagonale de coefficients $\sigma_a^2$ (resp. $\sigma_\phi^2$), et $\boldsymbol{W}_a$ (resp. $\boldsymbol{W}_\phi$) son inverse. Puisque $\boldsymbol{a}$ et $\boldsymbol{\phi}$ sont des vecteurs gaussiens, à moyenne nulle, et décorrélés, leur distribution de probabilité est de la forme :

$$\begin{aligned} p(\boldsymbol{a},\boldsymbol{\phi}) &\propto \mathrm{e}^{-\frac{1}{2}\left[\boldsymbol{a}^\mathrm{T}\boldsymbol{W}_a\boldsymbol{a}+\boldsymbol{\phi}_\mathrm{b}^\mathrm{T}\boldsymbol{W}_\phi\boldsymbol{\phi}_\mathrm{b}\right]} \\ &\propto \exp -\frac{1}{2}\begin{bmatrix}\boldsymbol{a}_\mathrm{b}\\ \hline \boldsymbol{\phi}_\mathrm{b}\end{bmatrix}^\mathrm{T}\left[\begin{array}{c|c} \boldsymbol{W}_a & 0 \\ \hline 0 & \boldsymbol{W}_\phi \end{array}\right]\begin{bmatrix}a_\mathrm{b}\\ \hline \phi_\mathrm{b}\end{bmatrix} \end{aligned} \tag{2.3}$$

Le critère d'attache aux données s'écrit comme un terme de $\chi^2$ :

$$\begin{aligned} \mathcal{J}^{\mathrm{data}}(\boldsymbol{x}) &= \boldsymbol{a}_\mathrm{b}^\mathrm{T}\boldsymbol{W}_a\boldsymbol{a}_\mathrm{b}+\boldsymbol{\phi}_\mathrm{b}^\mathrm{T}\boldsymbol{W}_\phi\boldsymbol{\phi}_\mathrm{b} \\ &= \left[\boldsymbol{a}^d-|H\boldsymbol{x}|\right]^\mathrm{T}\boldsymbol{W}_a\left[\boldsymbol{a}^d-|H\boldsymbol{x}|\right]+\left[\boldsymbol{\phi}^d-\arg(H\boldsymbol{x})\right]^\mathrm{T}\boldsymbol{W}_\phi\left[\boldsymbol{\phi}^d-\arg(H\boldsymbol{x})\right] \end{aligned}$$

Dans la suite, on notera $\chi_a^2(\boldsymbol{x})$ le terme $\left[\boldsymbol{a}^d-|H\boldsymbol{x}|\right]^\mathrm{T}\boldsymbol{W}_a\left[\boldsymbol{a}^d-|H\boldsymbol{x}|\right]$. Avec cette notation, on a :

$$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x}) = \chi_a^2(\boldsymbol{x})+\chi_\phi^2(\boldsymbol{x})$$

$\mathcal{J}^{\mathrm{data}}(\boldsymbol{x})$ est une somme de carrés pour chaque fréquence spatiale $\boldsymbol{\nu}_i$ :

$$j_i^{\mathrm{data}}(\boldsymbol{x}) = \left(\frac{|y_i^d|-|y_i(x)|}{\sigma_{a,i}}\right)^2+\left(\frac{\arg y_i^d-\arg y_i(x)}{\sigma_{\phi,i}}\right)^2$$

avec $\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{H}\,\boldsymbol{x},\ \boldsymbol{y}^d = a^d\mathrm{e}^{\mathrm{i}\phi^d}$.

### 2.2.3 $\mathcal{J}^{\text{data}}$ n'est pas convexe

Dans l'objectif de la minimisation du critère *a posteriori* $\mathcal{J} = \mathcal{J}^{\text{data}} + \mathcal{J}^{prior}$, le terme d'*a priori* restant à définir, il est important de s'intéresser aux caractéristiques mathématiques du terme d'attache aux données. En particulier, l'existence ou non de minima locaux de $\mathcal{J}^{\text{data}}$ influencera le choix d'une régularisation.
La convexité d'une fonction de plusieurs variable est une condition suffisante de l'absence de minima locaux, nous avons donc étudié la convexité de $\mathcal{J}^{\text{data}}$.

#### 2.2.3.1 Étude de $j_i^{\text{data}}$

La fonction $j_i^{\text{data}}$ est une mesure de l'écart entre la visibilité mesurée $y^d$ et le modèle $y(\boldsymbol{x})$, mesure dictée par la statistique du bruit. On peut l'interpréter comme le fait que l'espace des visibilités complexes est muni d'une « norme » (ce n'est pas mathématiquement une norme), qui n'est pas convexe.

Nous avons tracé les courbes iso-valeur de $j_i^{\text{data}}(y)$ dans le plan complexe. Ces courbes relient les valeurs de $y$ de même probabilité :

$$\text{Iso}_k(j_i^{\text{data}}) \triangleq \left\{y \in \mathbb{C}, j_i^{\text{data}}(y) = k\right\} \tag{2.4}$$

Ces courbes définissent des *ensembles de niveau* définis par

$$\bar{\text{Iso}}_k(j_i^{\text{data}}) \triangleq \left\{y \in \mathbb{C}, j_i^{\text{data}}(y) \leq k\right\} \tag{2.5}$$

Or, si une fonction est convexe, ses ensembles de niveaux sont des ensembles convexes. Le réseau de courbes obtenu est présenté dans la figure 2.1. Ce tracé démontre la non-convexité de cette fonction. En effet, si l'on trace un segment (représenté en pointillés sur la figure 2.1) d'un bout à l'autre d'un des ensembles de niveau, on sort de cet ensemble.

#### 2.2.3.2 Démonstration analytique

Cette non-convexité se démontre aussi analytiquement. Nous montrons ici que chacun des deux termes de $j_i^{\text{data}}(\boldsymbol{x})$ pris séparément – terme en phase et terme en module – est non convexe.

Sans perte de généralité, on étudie la convexité de $j_i^{\text{data}}(\boldsymbol{x})$. De plus, puisque $y_i$ se déduit de $\boldsymbol{x}$ par une opération linéaire, on étudie la convexité de $j_i(y) = j_i^{\text{data}}(\boldsymbol{x}) = \left(\frac{|y_i^d|-|y|}{\sigma_{a,i}}\right)^2 + \left(\frac{\arg y_i^d - \arg y}{\sigma_{\phi,i}}\right)^2$.

Si $j_i^{\text{data}}$ est convexe, elle vérifie : $j_i^{\text{data}}\left(\frac{y_1+y_2}{2}\right) < \frac{1}{2}\left(j_i^{\text{data}}(y_1) + j_i^{\text{data}}(y_2)\right)$
On choisit : $y_1 = y_i^d \, \mathrm{e}^{\frac{2i\pi}{3}}$ et $y_2 = y_i^d \mathrm{e}^{\frac{-2i\pi}{3}}$. On a alors :

$$\begin{aligned} j_i^{\text{data}}\left(\frac{y_1+y_2}{2}\right) &= j_i^{\text{data}}(-y_i^d/2) \\ &= \frac{(|y_i^d| - |-y_i^d/2|)^2}{2\sigma_{a,i}^2} + \frac{(\arg y_i^d - \arg(-y_i^d/2))^2}{2\sigma_{\phi,i}^2} \\ &= \frac{(|y_i^d|/2)^2}{2\sigma_{a,i}^2} + \frac{(\arg y_i^d - (\arg y_i^d + \pi)^2}{2\sigma_{\phi,i}^2} \\ &= \frac{|y_i^d|^2}{8\sigma_{a,i}^2} + \frac{\pi^2}{2\sigma_{\phi,i}^2} \end{aligned}$$

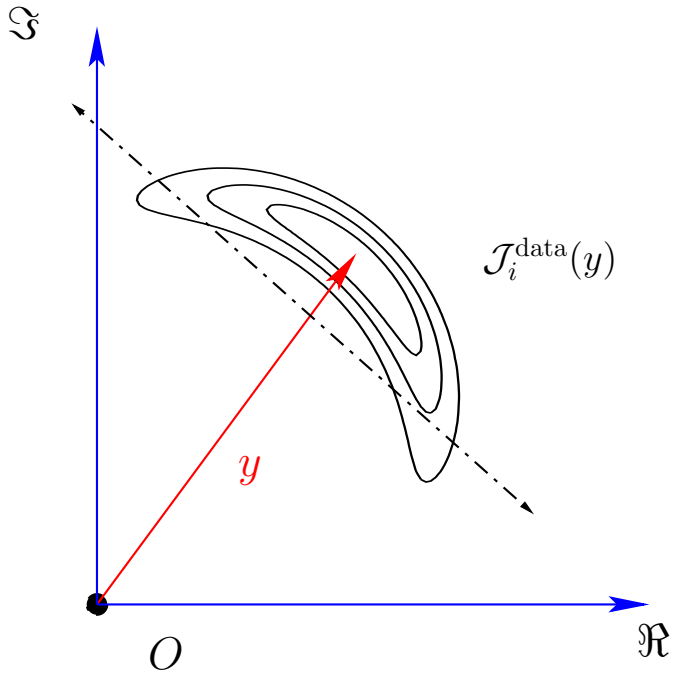



FIG. 2.1 – Courbes iso-valeur de $j_i^{\text{data}}(y)$, pour $\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$.

D'autre part, on calcule :

$$\begin{aligned} j_i^{\text{data}}(y_1) &= \frac{(|y_i^d| - |y_i^d\,\mathrm{e}^{\frac{2i\pi}{3}}|)^2}{2\sigma_{a,i}^2} + \frac{(\arg y_i^d - \arg y_i^d\,\mathrm{e}^{\frac{2i\pi}{3}})^2}{2\sigma_{\phi,i}^2} \\ &= \frac{0^2}{2\sigma_{a,i}^2} + \frac{(\frac{2\pi}{3})^2}{2\sigma_{\phi,i}^2} \\ &= \frac{2\pi^2}{9\sigma_{\phi,i}^2} \\ j_i^{\text{data}}(y_2) &= j_i^{\text{data}}(y_1) \end{aligned}$$

On a donc :

$$\begin{aligned} \frac{1}{2}\left(j_i^{\text{data}}(y_1) + j_i^{\text{data}}(y_2)\right) &= \frac{4\pi^2}{9\sigma_{\phi,i}^2} \\ &\leq \frac{\pi^2}{2\sigma_{\phi,i}^2} \\ &\leq \frac{\pi^2}{2\sigma_{\phi,i}^2} + \frac{|y_i^d|^2}{8\sigma_{a,i}^2} \\ \frac{1}{2}\left(j_i^{\text{data}}(y_1) + j_i^{\text{data}}(y_2)\right) &\leq j_i^{\text{data}}\left(\frac{y_1+y_2}{2}\right) \end{aligned}$$

Ceci prouve que chacun des deux morceaux de $j_i^{\text{data}}$ n'est pas convexe, et donc que $j_i^{\text{data}}$ ne l'est pas.

#### 2.2.3.3 Définition d'un complexe gaussien

Nous avons montré que $\mathcal{J}^{\text{data}}$ n'est pas convexe. Pourtant, la distribution de bruit associée est additive gaussienne :

$$\begin{cases} \boldsymbol{a}^d = \boldsymbol{a}(\boldsymbol{x}) + \text{ bruit gaussien} \\ \boldsymbol{\phi}^d = \boldsymbol{\phi}(\boldsymbol{x}) + \text{ bruit gaussien,} \end{cases}$$

et les lois gaussiennes sont des fonctions convexes. La contradiction vient de l'espace vectoriel dans lequel on se place.
Une variable $x$ est gaussienne si et seulement si sa distribution s'écrit :

$$p(x) \propto \exp -\frac{x-\bar{x}}{2\sigma^2},\ \sigma \neq 0$$

Un vecteur est gaussien si et seulement si toute combinaison linéaire de ses composantes est une variable gaussienne. Pour un complexe, on utilise le fait que $\mathbb{C}$ est isomorphe au plan $\mathbb{R}^2$. On en déduit qu'un complexe $b$ est gaussien si et seulement si le vecteur constitué de ses parties réelle et imaginaire est gaussien :

$$p(b) \propto \exp -\frac{1}{2} \begin{bmatrix} \Re(b-\bar{b}) \\ \Im(b-\bar{b}) \end{bmatrix}^{\mathrm{T}} \boldsymbol{W} \begin{bmatrix} \Re(b-\bar{b}) \\ \Im(b-\bar{b}) \end{bmatrix}$$

avec $\boldsymbol{W}$ une matrice symétrique réelle définie positive.

Cette définition est évidemment valable pour n'importe quelle base du plan complexe $\mathbb{C}$, qui se déduit de la base canonique par combinaison linéaire des vecteurs $(1,0)$ et $(0,1)$ : un complexe est gaussien si et seulement si le vecteur constitué de ses coordonnées dans une base quelconque de $\mathbb{C}$ suit une statistique gaussienne.

Il sera utile par la suite d'expliciter la distribution de probabilité d'un complexe gaussien dans une base quelconque du plan complexe. Toute base orthonormée $(\boldsymbol{e}_{\text{rad}}, \boldsymbol{e}_{\tan})$ du plan complexe correspond à la rotation d'un angle $\psi$ de la base canonique $\{\Re; \Im\}$ (cf. Fig. 2.2). La

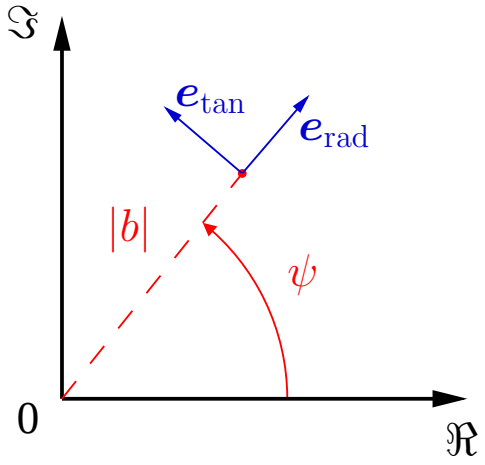



FIG. 2.2 – Bases orthonormées du plan complexe $\mathbb{C}$.

matrice de passage est une matrice de rotation :

$$\mathcal{R}(\psi) = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix}. \tag{2.6}$$

Si $b$ a pour coordonnées $b_{\rm rad}, b_{\rm tan}$ dans une base orthonormée $\{\boldsymbol{e}_{\rm rad}; \boldsymbol{e}_{\rm tan}\}$ de $\mathbb{C}$ de matrice de passage $\mathcal{R}(\psi)$, on a :

$$\begin{bmatrix} \Re(b_i) \\ \Im(b_i) \end{bmatrix} = \mathcal{R}(\psi) \begin{bmatrix} b_{\rm rad} \\ b_{\rm tan} \end{bmatrix} \quad \text{où } b_{\rm rad} = \Re(b\mathrm{e}^{-\mathrm{i}\psi}) \text{ et } b_{\rm tan} = \Im(b\mathrm{e}^{-\mathrm{i}\psi}) \tag{2.7}$$

Il est alors facile de voir que si $\begin{bmatrix} \Re(b_i) \\ \Im(b_i) \end{bmatrix}$ est gaussien, de covariance inverse $\boldsymbol{W}$, alors $\begin{bmatrix} b_{\rm rad} \\ b_{\rm tan} \end{bmatrix}$ est gaussien, de covariance inverse

$$\tilde{\boldsymbol{W}} = \mathcal{R}(\psi)^{\mathrm{T}} \boldsymbol{W} \mathcal{R}(\psi) \tag{2.8}$$

## 2.3 Approximation convexe quadratique du critère d'attache aux données

Une solution pour s'affranchir des problèmes liés à la non-convexité du critère d'attache aux données est de remplacer le bruit « réel » par un bruit approché donnant lieu à un critère d'attache aux données convexe.

### 2.3.1 Modèle de données en radio-interférométrie

Ce problème de non-convexité du critère d'attache aux données est une spécificité optique, due au fait que l'on mesure séparément phase et module des visibilités complexes. Dans le domaine des ondes radio, on mesure directement un complexe. De ce fait, le modèle de bruit est différent et conduit à un **critère d'attache aux données convexe**. En effet, on considère en radio-astronomie un modèle de bruit additif sur la mesure de visibilité complexe :

$$\boldsymbol{y}^d = \boldsymbol{y}(\boldsymbol{x}) + \boldsymbol{y}_{\rm b}$$

La distribution de $\boldsymbol{y}_{\rm b}$ est assimilée à une distribution gaussienne

$$p(\boldsymbol{y}_{\rm b}) \propto \mathrm{e}^{-\frac{1}{2}{\boldsymbol{y}_{\rm b}}^{\rm H} \boldsymbol{W}_y \boldsymbol{y}_{\rm b}}.$$

où ${\boldsymbol{y}_{\rm b}}^{\rm H}$ est le transposé conjugué du vecteur complexe $\boldsymbol{y}_{\rm b}$, et où $\boldsymbol{W}_y$ est une matrice diagonale de coefficients $\boldsymbol{\sigma}_y^{-2}$.

Le critère d'attache aux données s'écrit comme un terme de $\chi^2$ sur un complexe :

$$\begin{aligned} \mathcal{J}^{\rm data}(\boldsymbol{x}) &= {\boldsymbol{y}_{\rm b}}^{\rm H} \boldsymbol{W}_y \boldsymbol{y}_{\rm b} \\ &= \left[\boldsymbol{y}^d - H\boldsymbol{x}\right]^{\rm H} \boldsymbol{W}_y \left[\boldsymbol{y}^d - H\boldsymbol{x}\right] \end{aligned}$$

$\mathcal{J}^{\rm data}(\boldsymbol{x})$ est une somme de carrés pour chaque fréquence spatiale $\boldsymbol{\nu}_i$ :

$$j_i^{\rm data}(\boldsymbol{x}) = \left( \frac{|y_i^d - y_i(x)|}{\sigma_{y,i}} \right)^2$$

avec $\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{H}\,\boldsymbol{x}$, $\boldsymbol{y}^d = a^d \mathrm{e}^{\mathrm{i}\phi^d}$. On a associé au plan complexe des visibilités complexes $y$ une distance qui est, au facteur $\boldsymbol{\sigma}_{y,i}^{-2}$ près, la norme euclidienne canonique. Les lignes isocritère (cf. eq. 2.4) sont :

$$\mathrm{Iso}_k(j_i^{\rm data}) = \left\{ y,\ |y_i^d - y|^2 = k\sigma_{y,i}^2 \right\}$$

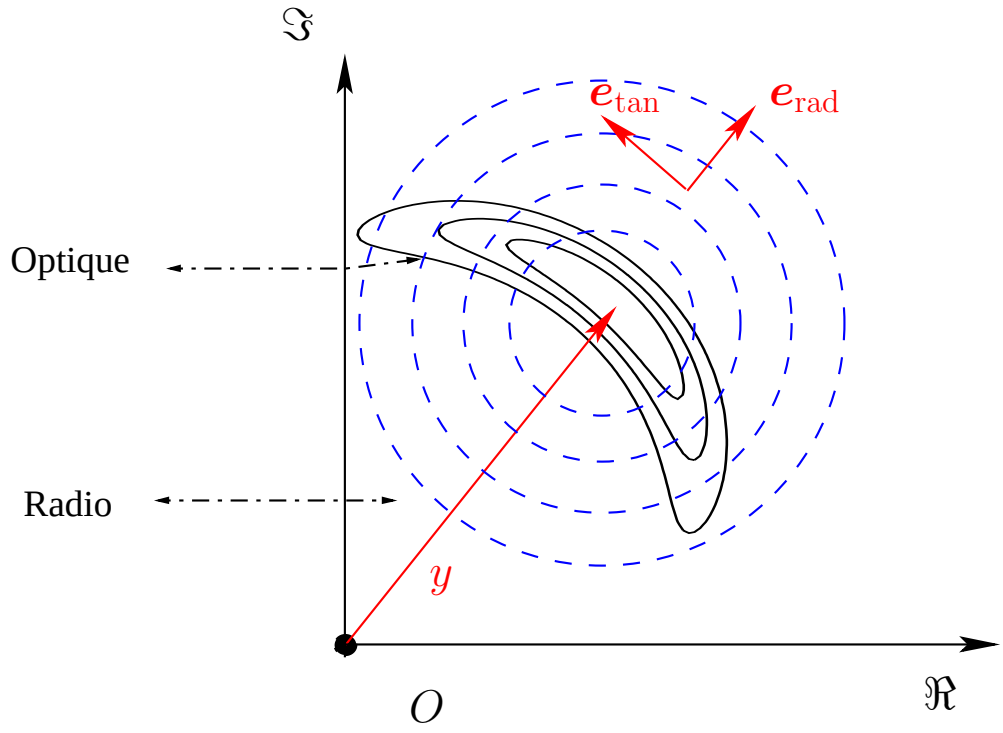



FIG. 2.3 – Courbes isocritère pour les modèles , pour de bruit optique ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$) et radio ($\sigma_{y,i} = \sqrt{a^d \sigma_{a,i}\sigma_{\phi,i}}$).

Ce sont donc des cercles concentriques, d'où l'appellation de bruit « circulaire ». La figure 2.3 présente les courbes isocritère obtenues pour les modèles de bruit radio et optique.

La plupart des méthodes de traitement de données interférométriques ont été développées dans le domaine des ondes radio, et postulent un modèle de bruit circulaire. C'est ce modèle qui est utilisé notamment dans les travaux de André Lannes en interférométrie optique Lannes (1998b). Il est donc naturel de chercher dans un premier temps à approcher le modèle de bruit optique par un modèle de bruit circulaire. Le problème est qu'il faut déterminer un écart-type $\sigma_{y,i}$ en fonction de $\sigma_{a,i}$ et de $\sigma_{\phi,i}$. On peut choisir l'écart-type $\sigma_{y,i}$ de telle sorte que les ensembles de niveau $\bar{\mathrm{Iso}}_k$ (cf. eq. 2.5) pour les modèles circulaire et réel soient de même aire :

$$k\pi\sigma_{y,i}^2 = \iint_{(a,\phi),ae^{\mathrm{i}\phi}\in\bar{\mathrm{Iso}}_k j_i^{\mathrm{data}}} a\mathrm{d}a\mathrm{d}\phi$$

Le calcul donne :

$$\sigma_{y,i} = \sqrt{a^d \sigma_{a,i}\sigma_{\phi,i}} \tag{2.9}$$

Néanmoins, on voit qu'avec l'approximation circulaire, on perd le caractère *directionnel* du modèle de bruit réel. Nous nous sommes donc posé le problème suivant :

- Existe-t-il d'autres approximations gaussiennes qui seraient plus proches de la distribution réelle ?
- Plus généralement, comment mesurer la qualité d'une approximation convexe ?

### 2.3.2 Choisir la « meilleure » approximation gaussienne

Notre objectif étant de comparer le modèle réel du bruit (gaussien en phase et module) à un modèle de bruit complexe additif $\boldsymbol{y}^d = \boldsymbol{y}(\boldsymbol{x}) + \boldsymbol{y}_\mathrm{b}$, nous avons entrepris de mettre le modèle de bruit optique sous la forme d'un bruit complexe additif (section 2.3.2.1), et non sous la forme d'un bruit en phase et d'un bruit en module. Nous avons alors tenté de définir un critère de qualité pour une approximation gaussienne, et de déterminer l'approximation gaussienne optimale selon ce critère.
Dans cette section, on ne considère qu'une seule fréquence spatiale $\boldsymbol{\nu}_0$. Les visibilités complexe $y_0$, en amplitude $a_0$ ou en phase $\phi_0$ seront notées $y$, $a$ et $\phi$, pour ne pas surcharger les calculs.

#### 2.3.2.1 Bruit complexe additif équivalent

On a (cf. Eqs.2.1 et 2.2)

$$\begin{cases} a^d = a(\boldsymbol{x}) + a_b \\ \phi^d = \phi(\boldsymbol{x}) + \phi_b \end{cases} \tag{2.10}$$

et on veut un modèle de la forme

$$y^d = y(\boldsymbol{x}) + y_b. \tag{2.11}$$

On forme donc des pseudo-mesures de visibilité complexes $y^d = a^d \mathrm{e}^{\mathrm{i}\phi^d}$ :

$$\begin{aligned} y^d &= (a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}(\phi(\boldsymbol{x})+\phi_b)} \\ &= a(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} + \left[(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b}\right]\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} - a(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} \\ &= a(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} + \left[(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b} - a(\boldsymbol{x})\right]\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} \\ &= y(\boldsymbol{x}) + \left[(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b} - a(\boldsymbol{x})\right]\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} \end{aligned}$$

En identifiant cette expression avec l'équation 2.11, on obtient :

$$y_b = \left[(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b} - a(\boldsymbol{x})\right]\mathrm{e}^{\mathrm{i}\phi(\boldsymbol{x})} \tag{2.12}$$

Il est évident, avec ces notations, que le bruit complexe $y_b$ n'est pas gaussien, et que le modèle de bruit n'est pas additif.

Pour étudier ce bruit $y_b$ , on choisit la base orthonormée correspondant à une rotation d'un angle $\psi = \phi(\boldsymbol{x})$ où $\phi(\boldsymbol{x})$ est la phase « vraie », c'est-à-dire la phase de la TF de l'objet à la fréquence considérée (Eq.2.2). Dans cette base, les coordonnées de $y_b$ sont (Eqs. 2.12 et 2.7) :

$$\begin{cases} y_b.\boldsymbol{e}_\mathrm{rad} = \Re(y_b \mathrm{e}^{-i\phi(\boldsymbol{x})}) \\ \qquad\qquad = \Re\left\{(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b} - a(\boldsymbol{x})\right\} \\ \qquad\qquad = (a(\boldsymbol{x}) + a_b)\cos\phi_b - a(\boldsymbol{x}) \\ y_b.\boldsymbol{e}_\mathrm{tan} = \Im(y_b \mathrm{e}^{-i\phi(\boldsymbol{x})}) \\ \qquad\qquad = \Im\left\{(a(\boldsymbol{x}) + a_b)\,\mathrm{e}^{\mathrm{i}\phi_b} - a(\boldsymbol{x})\right\} \\ \qquad\qquad = (a(\boldsymbol{x}) + a_b)\sin\phi_b \end{cases} \tag{2.13}$$

– La distribution exacte est donnée par (Eq. 2.3 page 44, pour une seule fréquence) :

$$p(a_\mathrm{b}, \phi_\mathrm{b}) \propto \exp -\frac{1}{2} \begin{bmatrix} a_\mathrm{b} \\ \phi_\mathrm{b} \end{bmatrix}^\mathrm{T} \begin{bmatrix} \sigma_a^{-2} & 0 \\ 0 & \sigma_\phi^{-2} \end{bmatrix} \begin{bmatrix} a_\mathrm{b} \\ \phi_\mathrm{b} \end{bmatrix}$$

– Une distribution gaussienne approchée s'écrit :

$$p(y_\mathrm{b}) \propto \exp -\frac{1}{2} \begin{bmatrix} (y_b - \bar{y_\mathrm{b}}).\boldsymbol{e}_\mathrm{rad} \\ (y_b - \bar{y_\mathrm{b}}).\boldsymbol{e}_\mathrm{tan} \end{bmatrix}^\mathrm{T} \boldsymbol{W}_y \begin{bmatrix} (y_b - \bar{y_\mathrm{b}}.\boldsymbol{e}_\mathrm{rad} \\ (y_b - \bar{y_\mathrm{b}}).\boldsymbol{e}_\mathrm{tan} \end{bmatrix} \tag{2.14}$$

– Le cas particulier d'une distribution circulaire correspond à $\boldsymbol{W}_y = \sigma_y^{-2}\mathbf{Id}$, avec $\sigma_y$ pouvant être choisi selon l'équation 2.9 .

#### 2.3.2.2 Distance de Kullback

Il est nécessaire de définir une distance entre deux distributions, de façon à déterminer la distribution gaussienne la plus « proche » de la distribution réelle. Il existe plusieurs distances entre distributions, par exemple les distances de Minkowski, de Bhattacharyya, du $\chi^2$. La distance de Kullback-Leibler est définie par :

$$\Delta(f,g) = -\int f(u) \log \frac{g(u)}{f(u)} \mathrm{d}u = -\mathrm{E}_f\{\log g\} + cst,$$

avec $\mathrm{E}_f\{\}$ l'espérance selon la loi $f$. Cette distance nous a semblé adaptée pour deux raisons :

– C'est une « log-distance », facile à calculer pour des lois gaussiennes ;

– Cette distance n'est pas symétrique : $\Delta(f,g) \neq \Delta(g,f)$. On la notera donc plutôt

$$\Delta_f(g) = -\mathrm{E}_f\{\log g\} + cst$$

Dans le cas qui nous occupe, les deux lois à comparer n'ont pas le même statut. Une des deux lois, $f$, est la loi décrivant le modèle **réel**, et l'autre, $g$, en est une **approximation**. Cette dissymétrie de la distance de Kullback Leibler est cohérente avec notre problème. La distance d'une loi gaussienne $g$ à la vraie loi $f$ sera égale à l'espérance selon $f$ de $\log g(u)$.

On définit donc un critère de qualité d'une approximation gaussienne $g$ de la loi vraie $f$ par :

$$\Delta_f(g) = -\mathrm{E}_f\{\log g\}$$

#### 2.3.2.3 Approximation quadratique optimale : l'approximation elliptique

Un résultat vraisemblablement classique, mais dont nous proposons une démonstration succincte en annexe A, permet d'affirmer que pour toute distribution $f(\boldsymbol{x})$, la distribution Gaussienne définie par

$$g(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi \det \boldsymbol{M}}} \exp -\frac{1}{2} (\boldsymbol{x} - \bar{\boldsymbol{x}})^t \boldsymbol{M}^{-1} (\boldsymbol{x} - \bar{\boldsymbol{x}})$$

où $\boldsymbol{M}$ est une matrice symétrique réelle définie positive, et qui atteint le minimum de la Distance de Kullback-Leibler $\delta(f,g)$ est telle que :

$$\begin{aligned} \bar{\boldsymbol{x}} &= \mathrm{E}_f\{\boldsymbol{x}\} \\ \boldsymbol{M} &= \mathrm{E}_f\left\{(\boldsymbol{x} - \mathrm{E}_f\{\boldsymbol{x}\})(\boldsymbol{x} - \mathrm{E}_f\{\boldsymbol{x}\})^\mathrm{T}\right\} \\ &= \boldsymbol{R}_{\boldsymbol{x}} \end{aligned} \tag{2.15}$$

Nous avons donc appliqué ces résultats au bruit complexe $y_b$ dont les coordonnées dans une base orthonormée particulière (Section 2.2.3.3) s'écrivent (Eq. 2.13) :

$$\begin{cases} y_b.\boldsymbol{e}_{\text{rad}} = (a(\boldsymbol{x}) + a_b)\cos\phi_b - a(\boldsymbol{x}) \\ y_b.\boldsymbol{e}_{\text{tan}} = (a(\boldsymbol{x}) + a_b)\sin\phi_b \end{cases}$$

L'équation 2.15, appliquée au vecteur $\begin{matrix} y_b.\boldsymbol{e}_{\text{rad}} \\ y_b.\boldsymbol{e}_{\text{tan}} \end{matrix}$, donne (voir annexe A) :

$$\begin{cases} \begin{bmatrix} \bar{y_{\text{b}}}.\boldsymbol{e}_{\text{rad}} \\ \bar{y_{\text{b}}}.\boldsymbol{e}_{\text{tan}} \end{bmatrix} = \begin{bmatrix} a(\boldsymbol{x})\left(\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} - 1\right) \\ 0 \end{bmatrix} \\ \boldsymbol{R}_y = \left[\begin{array}{c|c} \sigma_{\text{rad}}^2 & 0 \\ \hline 0 & \sigma_{\text{tan}}^2 \end{array}\right] = \left[\begin{array}{c|c} \frac{1+\mathrm{e}^{-\sigma_\phi^2}}{2}\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}a^2(\boldsymbol{x}) & 0 \\ \hline 0 & \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}a^2(\boldsymbol{x}) \end{array}\right] \end{cases} \tag{2.16}$$

On constate que :

- $\boldsymbol{R}_y$ est diagonale dans la base particulière $(\boldsymbol{e}_{\text{rad}}, \boldsymbol{e}_{\text{tan}})$ décrite section 2.2.3.3, ce qui en justifie le choix ;
- dans le cas général, l'approximation optimale n'est pas une loi circulaire, car $\boldsymbol{R}_y \neq \sigma^2\mathbf{Id}$,c'est-à-dire que $\sigma_{\text{rad}}^2 \neq \sigma_{\text{tan}}^2$.

#### 2.3.2.4 Approximation quadratique au premier ordre en $\sigma_a^2$, $\sigma_\phi^2$ et $a$

Pour pouvoir analyser qualitativement cette approximation optimale, nous avons fait un développement au premier ordre en $\sigma_a^2$ et $\sigma_\phi^2$ des paramètres $\bar{y_{\text{b}}}.\boldsymbol{e}_{\text{rad}}$, $\sigma_{\text{rad}}^2$ et $\sigma_{\text{tan}}^2$ :

$$\begin{cases} \bar{y_{\text{b}}}.\boldsymbol{e}_{\text{rad}} = a(\boldsymbol{x})\left(\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} - 1\right) \\ \qquad\qquad \simeq -\dfrac{a(\boldsymbol{x})\sigma_\phi^2}{2} \\ \qquad\qquad \simeq 0 \text{ si } a(\boldsymbol{x}) \ll 1 \\ \sigma_{\text{rad}}^2 = \dfrac{1+\mathrm{e}^{-\sigma_\phi^2}}{2}\sigma_a^2 + \dfrac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}a^2(\boldsymbol{x}) \\ \qquad\qquad \simeq \sigma_a^2 + \dfrac{a^2(\boldsymbol{x})\sigma_\phi^2}{2} \\ \qquad\qquad \simeq \sigma_a^2 \text{ si } a^2(\boldsymbol{x}) \ll 1 \\ \sigma_{\text{tan}}^2 = \dfrac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\sigma_a^2 + \dfrac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}a^2(\boldsymbol{x}) \quad \simeq \dfrac{a(\boldsymbol{x})^2\sigma_\phi^2}{2} \end{cases} \tag{2.17}$$

On voit qu'au premier ordre en $\sigma_a^2$ et $\sigma_\phi^2$, et pour des visibilités en amplitude $a(\boldsymbol{x})$ faibles devant 1, l'approximation elliptique n'est pas biaisée. Cette hypothèse sur les amplitudes est l'hypothèse la plus forte, et est donc susceptible de générer des erreurs de reconstruction. La principale différence entre l'approximation optimale et son développement au premier ordre est le biais dans la direction radiale. La différence de performance entre les deux approximations elliptiques, si différence il y a, devra être mise en évidence par simulations.
L'écart-type dans la direction radiale est, de façon prévisible, très proche de l'écart-type en module de la distribution originale. L'hypothèse faite sur les visibilités est $a^2(\boldsymbol{x}) \ll 1$, qui est

une hypothèse moins forte que $a(\boldsymbol{x}) \ll 1$.
Enfin, l'écart type dans la direction tangentielle est, au premier ordre, la longueur de l'arc de rayon $a(\boldsymbol{x})$ et d'angle $\sigma_\phi$ (Figure 2.4).

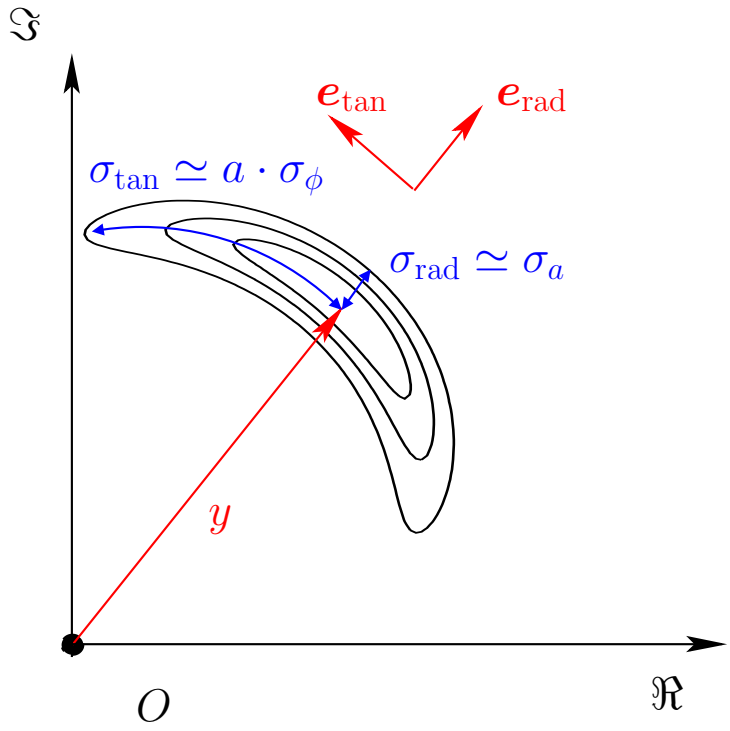



FIG. 2.4 – Interprétation graphique des écarts-type de l'approximation optimale développée au premier ordre ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$).

De plus, on peut faire les remarques suivantes :
– Le développement limité au premier ordre des paramètres de l'approximation optimale conduit à la même loi qu'un développement limité de la loi « vraie » autour du maximum ;
– pour la loi « optimale » développée au premier ordre, les ensembles de niveau ont la même aire que ceux de la loi vraie, et que la loi circulaire (cf. eq 2.9)

La figure 2.5 montre la forme des lignes isocritère pour les distributions « vraie », « circulaire » et « optimale » développée au premier ordre[1]. On retrouve pour la distribution optimale des lignes isocritère elliptiques, dont les axes sont orientés selon les vecteurs $(\boldsymbol{e}_{\mathrm{rad}}, \boldsymbol{e}_{\tan})$.

Par simplicité des calculs, on considérera dans la suite le développement limité au premier ordre de l'approximation optimale (avec l'hypothèse de faibles visibilités).

### 2.3.3 Le critère $\mathcal{J}_{ell}$

Dans cette section, on étudie le critère d'attache aux données $\mathcal{J}_{ell}$ obtenu avec l'approximation *elliptique* optimale du bruit. Les indices $i$ que nous avions omis dans la section précédente

[1]L'approximation elliptique est tout aussi valable dans le cas où $\sigma_{a,i} > a_i^d \sigma_{\phi,i}$ [Meimon *et al.* (name)].

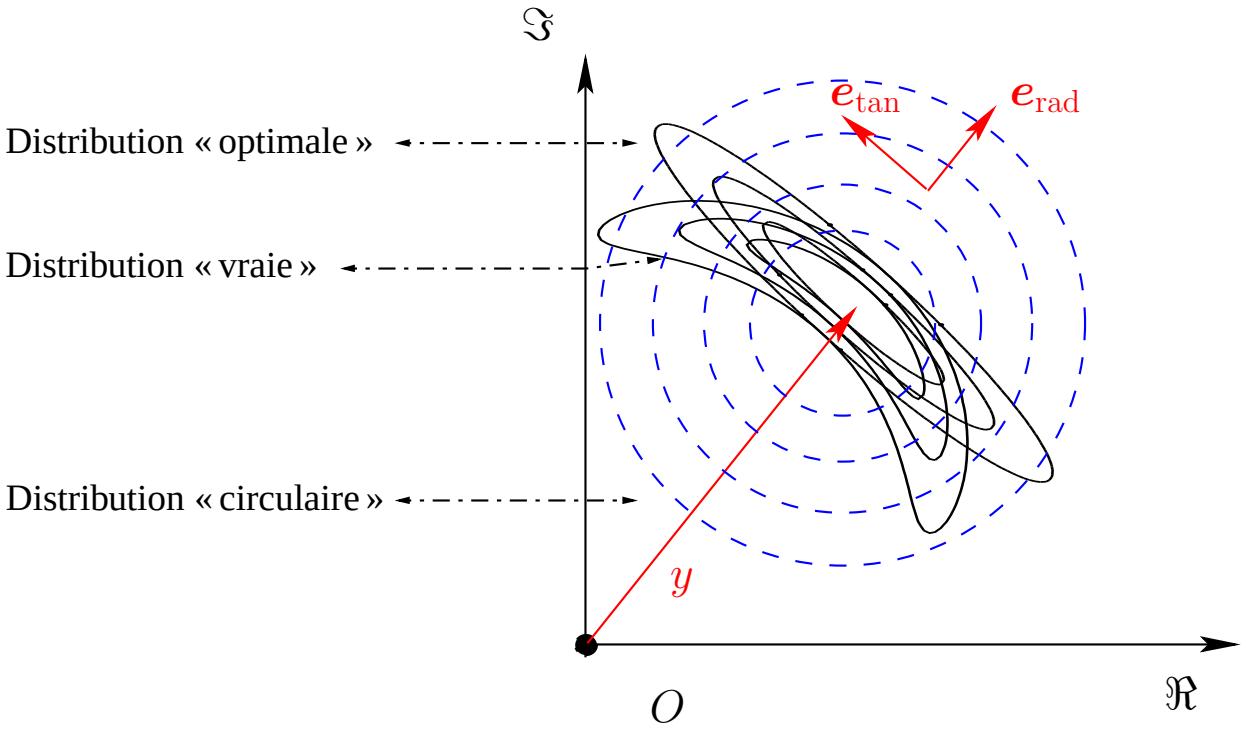



FIG. 2.5 – Courbes isocritère pour les distributions « vraie », « circulaire » et « optimale » ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$).

doivent ici être mentionnés.

#### 2.3.3.1 Forme du critère $\mathcal{J}^i_{ell}$

Avec notre approximation, le bruit $y_{\mathrm{b},i}$ suit une statistique gaussienne

$$p(y_{\mathrm{b},i}) \propto \exp -\frac{1}{2} \begin{bmatrix} y_{\mathrm{b},i}.\boldsymbol{e}_{\rm rad} \\ y_{\mathrm{b},i}.\boldsymbol{e}_{\tan} \end{bmatrix}^{\mathrm{T}} \tilde{\boldsymbol{W}}_{y,i} \begin{bmatrix} y_{\mathrm{b},i}.\boldsymbol{e}_{\rm rad} \\ y_{\mathrm{b},i}.\boldsymbol{e}_{\tan} \end{bmatrix}$$

avec

$$\tilde{\boldsymbol{W}}_{y,i} = \boldsymbol{R}^{-1}_{y,i} \simeq \mathrm{Diag}\left\{\sigma^{-2}_{\mathrm{rad},i}, \sigma^{-2}_{\tan,i}\right\}$$

et

$$\sigma_{\mathrm{rad},i} \simeq \sigma_{a,i}, \qquad \sigma_{\tan,i} \simeq a_i(\boldsymbol{x})\sigma_{\phi,i}$$

On peut donc réécrire la loi de probabilité :

$$p(y_{\mathrm{b},i}) \propto \exp -\frac{1}{2} \left[ \frac{(y_{\mathrm{b},i}.\boldsymbol{e}_{\rm rad})^2}{\sigma^2_{\mathrm{rad},i}} + \frac{(y_{\mathrm{b},i}.\boldsymbol{e}_{\tan})^2}{\sigma^2_{\tan,i}} \right]$$

On en déduit le critère d'attache aux données approché :

$$\mathcal{J}^i_{ell} = \frac{((y^{\rm data}_i - y_i(\boldsymbol{x})).\boldsymbol{e}_{\mathrm{rad},i})^2}{\sigma^2_{\mathrm{rad},i}} + \frac{((y^{\rm data}_i - y_i(\boldsymbol{x})).\boldsymbol{e}_{\tan,i})^2}{\sigma^2_{\tan,i}} \tag{2.18}$$

Cette expression est une norme carrée sur $y_{\mathrm{res},i} = y_i(\boldsymbol{x}) - y_i^{\mathrm{data}}$ :

$$\begin{aligned} \mathcal{J}_{ell}^i &= \|y_{\mathrm{res},i}\|_{ell,i}^2 \\ \|z\|_{ell,i} &= \frac{(z.\boldsymbol{e}_{\mathrm{rad},i})^2}{\sigma_{\mathrm{rad},i}^2} + \frac{(z.\boldsymbol{e}_{\mathrm{tan},i})^2}{\sigma_{\mathrm{tan},i}^2} \\ &= \frac{\Re^2(z\mathrm{e}^{-\mathrm{i}\,\phi_i^{\mathrm{data}}})}{\sigma_{\mathrm{rad},i}^2} + \frac{\Im^2(z\mathrm{e}^{-\mathrm{i}\,\phi_i^{\mathrm{data}}})}{\sigma_{\mathrm{tan},i}^2} \end{aligned} \tag{2.19}$$

#### 2.3.3.2 Forme matricielle du critère $\mathcal{J}_{ell}^i$

Le critère $\mathcal{J}_{ell}^i$ a été construit de façon à obtenir un critère convexe quadratique sur $\boldsymbol{x}$. Nous avons cherché à cette étape à exprimer $\mathcal{J}_{ell}^i$ sous la forme

$$\mathcal{J}_{ell}(\boldsymbol{x}) = (\boldsymbol{D} - \boldsymbol{M}\boldsymbol{x})^{\mathrm{T}} \cdot \boldsymbol{W} \cdot (\boldsymbol{D} - \boldsymbol{M}\boldsymbol{x}) \,,$$

où $\boldsymbol{D}$ contient les parties imaginaire et réelle des données, $\boldsymbol{M}$ dépendant des fréquences mesurées, et $\boldsymbol{W}$ dépendant des $\sigma_{\mathrm{rad},i}$ et$\sigma_{\mathrm{tan},i}$.

On définit pour un nombre $z$, un vecteur $\boldsymbol{v}$ ou une matrice $\boldsymbol{M}$ complexes quelconques :

$$\bar{\bar{z}} = \begin{bmatrix} \Re z \\ \Im z \end{bmatrix}, \; \bar{\bar{\boldsymbol{v}}} = \begin{bmatrix} \Re \boldsymbol{v} \\ \hline \Im \boldsymbol{v} \end{bmatrix}, \; \bar{\bar{\boldsymbol{M}}} = \begin{bmatrix} \Re \boldsymbol{M} \\ \hline \Im \boldsymbol{M} \end{bmatrix}. \tag{2.20}$$

La loi de probabilité gaussienne approchée du bruit complexe $y_{b,i}$, développée au premier ordre, est :

$$p(y_{b,i}) \propto \exp -\frac{1}{2} \bar{\bar{y}}_{b,i}^{\mathrm{T}} \, \boldsymbol{W}_{y,i} \, \bar{\bar{y}}_{b,i}$$

avec

$$\begin{cases} \boldsymbol{W}_{y,i} = \mathcal{R}(\phi_i(\boldsymbol{x}))^{\mathrm{T}} \tilde{\boldsymbol{W}}_{y,i} \mathcal{R}(\phi_i(\boldsymbol{x})) \\ \tilde{\boldsymbol{W}}_{y,i} = \mathrm{Diag}\left\{\sigma_{\mathrm{rad},i}^{-2}, \sigma_{\mathrm{tan},i}^{-2}\right\} \\ \qquad\quad \simeq \mathrm{Diag}\left\{\sigma_{a,i}^{-2}, (a_i(\boldsymbol{x})\sigma_{\phi,i})^{-2}\right\} \end{cases}$$

Ne connaissant pas $\phi_i(\boldsymbol{x})$ et $a_i(\boldsymbol{x})$, on utilise pour construire le modèle de bruit les mesures correspondantes $\phi_i^d$ et $a_i^d$ :

$$\begin{cases} \boldsymbol{W}_{y,i} = \mathcal{R}(\phi_i^d)^{\mathrm{T}} \tilde{\boldsymbol{W}}_{y,i} \mathcal{R}(\phi_i^d) \\ \tilde{\boldsymbol{W}}_{y,i} = \mathrm{Diag}\left\{\sigma_{\mathrm{rad},i}^{-2}, \sigma_{\mathrm{tan},i}^{-2}\right\} \\ \qquad\quad \simeq \mathrm{Diag}\left\{(\sigma_a^d)^{-2}, (a_i^d\sigma_{\phi,i})^{-2}\right\} \end{cases}$$

$$\mathcal{J}_{ell}^i(\boldsymbol{x}) = \left(\bar{\bar{y}}_i^d - \bar{\bar{y}}_i(\boldsymbol{x})\right)^{\mathrm{T}} \, \boldsymbol{W}_{y,i} \, \left(\bar{\bar{y}}_i^d - \bar{\bar{y}}_i(\boldsymbol{x})\right)$$

#### 2.3.3.3 De $\mathcal{J}_{ell}^{i}$ à $\mathcal{J}_{ell}$

Le critère $\mathcal{J}_{ell}$ est simplement la somme des termes $\mathcal{J}_{ell}^{i}$ pour toutes les mesures :

$$\begin{aligned}
\mathcal{J}_{ell}(\boldsymbol{x}) &= \sum_i \mathcal{J}_{ell}^{i}(\boldsymbol{x}) \\
&= \sum_i \left(\bar{\bar{y}}_i^d - \bar{\bar{y}}_i(\boldsymbol{x})\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{y,i} \cdot \left(\bar{\bar{y}}_i^d - \bar{\bar{y}}_i(\boldsymbol{x})\right) \\
&= \begin{bmatrix} \Re(y_1^d - y_1(\boldsymbol{x})) \\ \Im(y_1^d - y_1(\boldsymbol{x})) \\ \hline \Re(y_2^d - y_2(\boldsymbol{x})) \\ \Im(y_2^d - y_2(\boldsymbol{x})) \\ \hline \vdots \\ \hline \Re(y_N^d - y_N(\boldsymbol{x})) \\ \Im(y_N^d - y_N(\boldsymbol{x})) \end{bmatrix}^t \underbrace{\left[\begin{array}{c|c|c|c} \boldsymbol{W}_{y,1} & 0 & \dots & 0 \\ \hline 0 & \boldsymbol{W}_{y,2} & \ddots & \dots \\ \hline \vdots & \ddots & \ddots & 0 \\ \hline 0 & \dots & 0 & \boldsymbol{W}_{y,N} \end{array}\right]}_{\boldsymbol{W}_y} \begin{bmatrix} \Re(y_1^d - y_1(\boldsymbol{x})) \\ \Im(y_1^d - y_1(\boldsymbol{x})) \\ \hline \Re(y_2^d - y_2(\boldsymbol{x})) \\ \Im(y_2^d - y_2(\boldsymbol{x})) \\ \hline \vdots \\ \hline \Re(y_N^d - y_N(\boldsymbol{x})) \\ \Im(y_N^d - y_N(\boldsymbol{x})) \end{bmatrix}
\end{aligned}$$

On souhaite exprimer cette relation en fonction de $\bar{\bar{y}}^d$, $\bar{\bar{\boldsymbol{H}}}$ et $\boldsymbol{x}$ (Eq.2.20). On réarrange donc l'ordre des coefficients :

$$\begin{bmatrix} \Re(y_1^d - y_1(\boldsymbol{x})) \\ \Im(y_1^d - y_1(\boldsymbol{x})) \\ \hline \Re(y_2^d - y_2(\boldsymbol{x})) \\ \Im(y_2^d - y_2(\boldsymbol{x})) \\ \hline \Re(y_3^d - y_3(\boldsymbol{x})) \\ \Im(y_3^d - y_3(\boldsymbol{x})) \\ \hline \vdots \\ \hline \Re(y_N^d - y_1(\boldsymbol{x})) \\ \Im(y_N^d - y_N(\boldsymbol{x})) \end{bmatrix} = \underbrace{\left[\begin{array}{cccccc|ccccc} 1 & 0 & \dots & \dots & \dots & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 & 0 & \dots & \dots & 0 \\ \hline 0 & 1 & 0 & \dots & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ \hline 0 & 0 & 1 & 0 & \dots & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline 0 & \dots & \dots & \dots & \dots & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 & \dots & \dots & \dots & 1 \end{array}\right]}_{\boldsymbol{P}} \underbrace{\begin{bmatrix} \Re(y_1^d - y_1(\boldsymbol{x})) \\ \Re(y_2^d - y_2(\boldsymbol{x})) \\ \Re(y_3^d - y_3(\boldsymbol{x})) \\ \vdots \\ \Re(y_N^d - y_N(\boldsymbol{x})) \\ \hline \Im(y_1^d - y_1(\boldsymbol{x})) \\ \Im(y_2^d - y_2(\boldsymbol{x})) \\ \Im(y_3^d - y_3(\boldsymbol{x})) \\ \vdots \\ \Im(y_N^d - y_N(\boldsymbol{x})) \end{bmatrix}}_{\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{y}}}(\boldsymbol{x})}$$

$$\begin{aligned}
&= \boldsymbol{P}\left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{y}}}(\boldsymbol{x})\right) \\
&= \boldsymbol{P}\left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)
\end{aligned}$$

Le critère $\mathcal{J}_{ell}(\boldsymbol{x})$ s'écrit alors :

$$\begin{aligned}
\mathcal{J}_{ell}(\boldsymbol{x}) &= \left[\boldsymbol{P}\left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)\right]^{\mathrm{T}} \cdot \boldsymbol{W}_y \cdot \left[\boldsymbol{P}\left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)\right] \\
&= \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)^{\mathrm{T}} \cdot \underbrace{\boldsymbol{P}^{\mathrm{T}}\boldsymbol{W}_y\boldsymbol{P}}_{\boldsymbol{W}_{\bar{\bar{y}}}} \cdot \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)
\end{aligned}$$

D'où

$$\boxed{\mathcal{J}_{ell}(\boldsymbol{x}) = \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}} \cdot \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)} \tag{2.21}$$

Sous cette forme, il est évident que le critère d'attache aux données $\mathcal{J}_{ell}(\boldsymbol{x})$ est quadratique et donc convexe par rapport à $\boldsymbol{x}$.

## 2.4 L'algorithme RIME : Reconstruction d'Images par Modélisation Elliptique

A ce stade de l'étude plusieurs questions se sont posées, concernant :
- l'efficacité de notre « convexification » du critère, vis-à-vis d'une approximation circulaire plus classique.
- le critère proposé en fin de section précédente, qui est une version développée au premier ordre du critère obtenu par approximation optimale. La différence entre ces deux versions du critère elliptique n'était pas à ce stade quantifiée.
- la capacité à reconstruire une image fidèle à l'objet d'intérêt, sur données simulées ou expérimentales, avec un critère d'attache aux données approché.

Pour pouvoir répondre à ces questions, la première étape a donc été d'élaborer une méthode de reconstruction d'images à partir de visibilités complexes, incluant, entre autres, l'adjonction d'un critère *a priori* (section 2.4.1) et l'élaboration d'une méthode de minimisation de critère (section 2.4.2).

### 2.4.1 Construction d'un modèle *a priori*

Parmi la variété d'*a priori* présentés section 1.4.2, nous avons opté pour une régularisation quadratique sous contrainte de positivité. Ce choix à été motivé par les arguments suivants :
- Étant donnée la sous-détermination du problème, il est nécessaire d'utiliser une contrainte de positivité ;
- le principal apport de notre méthode concerne une prise en compte fine du modèle de bruit. Pour étudier le critère d'attache aux données qui s'en déduit, il est logique de se placer dans le cas le plus simple, en termes d'*a priori* ;
- les objets sur lesquels ont été faites les simulations (section 2.5) ou le traitement de données expérimentales (section 2.6) sont des objets plutôt lisses.

Le choix d'un modèle de DSP et de l'hyperparamètre réglant les poids relatifs du terme d'attache aux données et du terme d'*a priori* est traité section 2.4.3.

L'utilisation de la contrainte de positivité rend l'utilisation d'une méthode d'inversion linéaire impossible. Cependant, le fait que sont quadratiques à la fois le critère d'attache aux données –grâce à l'utilisation de l'approximation convexe quadratique décrite section 2.3– et le terme de pénalisation nous à conduits à considérer une méthode de descente de type quasi-Newton.

### 2.4.2 Une méthode de descente locale utilisant le calcul du gradient

#### 2.4.2.1 Une méthode de type quasi-Newton

L'algorithme OP-VMLMB de minimisation sous contrainte de positivité que nous utilisons est un algorithme BFGS (Broyden-Fletcher-Goldfarb-Shanno) de type quasi-Newton, élaboré par E. Thiébaut Thiébaut (2002). La positivité y est imposée par une méthode d'activation des contraintes décrit plus en détail dans l'annexe C. A chaque étape de cette méthode, il faut minimiser (sous contrainte que quelques pixels sont « gelés ») un critère, qui dans le cas que nous traitons est quadratique. Cette minimisation peut donc se faire en un pas si l'on dispose du gradient et du Hessien en chaque point. Le calcul analytique du Hessien étant trop coûteux en place mémoire, cet algorithme en calcule une approximation à partir du gradient. Dans le cas non quadratique, même si la minimisation est plus difficile, il est néanmoins avantageux

d'utiliser un tel algorithme, qui interpole jusqu'au second ordre la fonction à minimiser.
Le gradient étant fournie sous forme analytique à l'algorithme, la section suivante présente le calcul du gradient du critère d'attache aux données.

#### 2.4.2.2 Calcul du gradient

Le terme $\mathcal{J}_{prior}$ dépend de l'*a priori* utilisé, et est généralement simple à calculer. Le gradient du terme d'attache aux données décrit section 2.3.3 (équation 2.21 page 56) est donné par :

$$\mathcal{J}_{ell}(\boldsymbol{x}) = \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}} \cdot \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right)$$

Sous cette forme, le critère d'attache aux données se dérive très simplement :

$$\begin{aligned}\frac{\partial \mathcal{J}_{ell}(\boldsymbol{x})}{\partial \boldsymbol{x}} &= \frac{\partial \bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}}{\partial \boldsymbol{x}} \cdot \frac{\partial \mathcal{J}_{ell}(\boldsymbol{x})}{\partial \bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}} \\ &= -(\bar{\bar{\boldsymbol{H}}})^{\mathrm{T}} \cdot 2\boldsymbol{W}_{\bar{\bar{y}}} \left(\bar{\bar{\boldsymbol{y}}}^d - \bar{\bar{\boldsymbol{H}}}\boldsymbol{x}\right) \\ &= 2(\bar{\bar{\boldsymbol{H}}})^{\mathrm{T}} \boldsymbol{W}_{\bar{\bar{y}}} \left(\bar{\bar{\boldsymbol{H}}}\boldsymbol{x} - \bar{\bar{\boldsymbol{y}}}^d\right)\end{aligned}$$

#### 2.4.2.3 Critère d'arrêt

Plusieurs possibilités existent. On peut par exemple arrêter les itérations quant la diminution du critère d'une étape à l'autre n'est plus quantifiable, c'est-à-dire lorsqu'elle est inférieure à la précision machine.
Une autre option consiste à tester non pas la diminution du critère, mais la valeur atteinte par la norme du gradient, puisqu'on calcule ce gradient à chaque étape. Néanmoins, nous n'avons pas constaté de différence entre les résultats obtenus pour ces deux critères d'arrêt.

### 2.4.3 Une reconstruction non supervisée

Il est illusoire d'espérer établir une stratégie de minimisation qui fonctionnera quel que soit le type d'objet à reconstruire, sans intervention de l'utilisateur. Par contre, la possibilité de calculer automatiquement un modèle *a priori* et une valeur d'hyperparamètre **proche** de l'optimum est un apport notable. L'utilisateur de la méthode pourra alors commencer par une inversion « automatique », ou **non supervisée**, pour ensuite affiner les réglages autour des paramètres conseillés.
Il existe des méthodes de choix automatique d'hyperparamètres dans un cas quadratique, notamment les approches dites « L-curve », et qui visent à équilibrer la valeur des termes d'attache aux données et de régularisation [Idier (2001)]. Une telle approche est néanmoins coûteuse en temps de calcul, car elle nécessite plusieurs tentatives avant de pouvoir estimer l'hyperparamètre.
Parallèlement, les travaux de Lannes *et al.* (1994) proposent une règle de choix d'hyperparamètre en fonction du conditionnement de la matrice du problème (dans notre cas, la matrice $\boldsymbol{H}$). Cependant, cette méthode nécessite une évaluation itérative du conditionnement de la matrice régularisée, et ce pour chaque valeur test des paramètres de régularisation. De plus, elle ne permet pas d'assurer l'optimalité de l'hyperparamètre obtenu lorsqu'une contrainte de positivité est imposée.
C'est dans ce but que nous avons proposé une méthode simple de choix de DSP pour une régularisation quadratique, dont la structure correspond aux travaux de Conan *et al.* (1998). Une

méthode d'estimation des paramètres de ce modèle de DSP, inspirée de Blanc *et al.* (2003), est décrite section 2.4.3.3. Nous ne prétendons pas atteindre un choix optimal d'hyperparamètre, mais simplement un réglage correct à partir duquel l'opérateur peut affiner la reconstruction.

#### 2.4.3.1 Estimation de DSP par Maximum de Vraisemblance

Le modèle de formation de données que nous avons élaboré (voir section 2.3) s'écrit :

$$\boldsymbol{y}^{\text{data}} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{y}_{\text{b}}$$

avec une statistique gaussienne du bruit. L'opérateur $\boldsymbol{H}$ étant l'opérateur de Fourier associé aux **fréquences mesurées** $\boldsymbol{\nu}_i$, cette relation s'écrit :

$$y_i^{\text{data}} = \widetilde{\boldsymbol{x}}(\boldsymbol{\nu}_i) + y_{\text{b},i}$$

où la TF considérée est une TF discrète.
On a vu qu'une régularisation quadratique paramétrée par la DSP de l'objet revient à choisir une loi gaussienne pour le coefficient de la TF de l'objet à une fréquence quelconque $\boldsymbol{\nu}$ de la grille d'échantillonnage de la TF discrète(Eq.1.24) :

$$p(\widetilde{x}(\boldsymbol{\nu})) \propto \mathrm{e}^{-\frac{1}{2}\frac{|\widetilde{x}(\boldsymbol{\nu})-\widetilde{x}^0(\boldsymbol{\nu})|^2}{DSP(\boldsymbol{\nu})}}$$

En particulier, aux fréquences couvertes par l'interféromètre, on a :

$$p(\widetilde{x}(\boldsymbol{\nu}_i)) \propto \mathrm{e}^{-\frac{1}{2}\frac{|\widetilde{x}(\boldsymbol{\nu})-\widetilde{x}^0(\boldsymbol{\nu})|^2}{DSP(\boldsymbol{\nu}_i)}}$$

Avec cet *a priori*, le vecteur $\boldsymbol{H}\boldsymbol{x}$ est un vecteur blanc gaussien, de matrice de covariance $\boldsymbol{S}_x = \text{Diag}\,\{DSP(\boldsymbol{\nu}_i)\}$.
Supposons que le bruit $\boldsymbol{y}_{\text{b}}$ est blanc et gaussien, de matrice de covariance $\boldsymbol{R}_b = \text{Diag}\,\{\sigma_i^2\}$.
Si $\boldsymbol{H}\boldsymbol{x}$ et $\boldsymbol{y}_{\text{b}}$ sont blancs et gaussiens, de matrices de covariance respectives $\boldsymbol{S}_x$ et $\boldsymbol{R}_b$, alors le vecteur des mesures $\boldsymbol{y}^{\text{data}}$ est lui aussi blanc gaussien, de matrice de covariance $\boldsymbol{R}_y = \boldsymbol{S}_x + \boldsymbol{R}_b$ :

$$\begin{aligned} p(\boldsymbol{y}^{\text{data}}) &= \frac{1}{\sqrt{2\pi|\boldsymbol{R}_y|}} \exp -\frac{1}{2}(\boldsymbol{y}^{\text{data}} - \boldsymbol{y}^0)^{\text{H}}(\boldsymbol{R}_y)^{-1}(\boldsymbol{y}^{\text{data}} - \boldsymbol{y}^0) \\ &= \frac{1}{\sqrt{2\pi|\boldsymbol{S}_x + \boldsymbol{R}_b|}} \exp -\frac{1}{2}(\boldsymbol{y}^{\text{data}} - \boldsymbol{y}^0)^{\text{H}}(\boldsymbol{S}_x + \boldsymbol{R}_b)^{-1}(\boldsymbol{y}^{\text{data}} - \boldsymbol{y}^0) \\ &= \frac{1}{\sqrt{2\pi\prod_i (DSP(\boldsymbol{\nu}_i) + \sigma_i^2)}} \exp -\frac{1}{2}\sum_i \frac{|\boldsymbol{y}_i^{\text{data}} - \boldsymbol{y}_i^0|^2}{DSP(\boldsymbol{\nu}_i) + \sigma_i^2} \end{aligned}$$

De façon inverse, cette fonction est aussi la vraisemblance de la DSP de l'objet et de la matrice de covariance du bruit $\boldsymbol{R}_b = \text{Diag}\,\{\sigma_i^2\}$, connaissant les mesures $\boldsymbol{y}^{\text{data}}$ :

$$\mathcal{L}(\boldsymbol{S}_x, \boldsymbol{R}_b|\boldsymbol{y}^{\text{data}}) = \frac{1}{\sqrt{2\pi\prod_i (DSP(\boldsymbol{\nu}_i) + \sigma_i^2)}} \exp -\frac{1}{2}\sum_i \frac{|\boldsymbol{y}_i^{\text{data}} - \boldsymbol{y}_i^0|^2}{DSP(\boldsymbol{\nu}_i) + \sigma_i^2}$$

#### 2.4.3.2 Modèle paramétrique de $\boldsymbol{R}_b$ et $DSP(\boldsymbol{\nu})$

On dispose déjà d'un modèle de bruit gaussien (section 2.3.3 page 54). Cependant, pour déterminer la DSP de l'objet, les calculs sont plus simples si le bruit est supposé uniforme : $\boldsymbol{R}_b = r_b\,\mathbf{Id}$ :

$$\mathcal{L}(\boldsymbol{S}_x, \sigma_b|\boldsymbol{y}^{\text{data}}) = \frac{1}{\sqrt{2\pi \prod_i (DSP(\boldsymbol{\nu}_i) + \sigma_b^2)}} \exp -\frac{1}{2} \sum_i \frac{|\boldsymbol{y}_i^{\text{data}} - \boldsymbol{y}_i^0|^2}{DSP(\boldsymbol{\nu}_i) + \sigma_b^2}$$

De plus, on considère le modèle paramétrique de DSP proposé Conan *et al.* (1998) :

$$DSP(\boldsymbol{\nu}) = \frac{K}{\left(\frac{|\boldsymbol{\nu}|}{\rho_0}\right)^p + 1}.$$

Les paramètres $\sigma_b$ ,$K$, $\rho_0$ et $p$ peuvent donc être déduits des données par Maximum de Vraisemblance, c'est-à-dire en maximisant :

$$\mathcal{L}(r_b, K, \rho_0, p) = \frac{1}{\sqrt{2\pi \prod_i (\frac{K}{\left(\frac{|\boldsymbol{\nu}_i|}{\rho_0}\right)^p + 1} + \sigma_b^2)}} \exp -\frac{1}{2} \sum_i \frac{|\boldsymbol{y}_i^{\text{data}} - \boldsymbol{y}_i^0|^2}{\frac{K}{\left(\frac{|\boldsymbol{\nu}_i|}{\rho_0}\right)^p + 1} + \sigma_b^2}$$

La figure 2.4.3.2 montre la modélisation obtenue pour les données simulées présentées section 2.5 page 62.

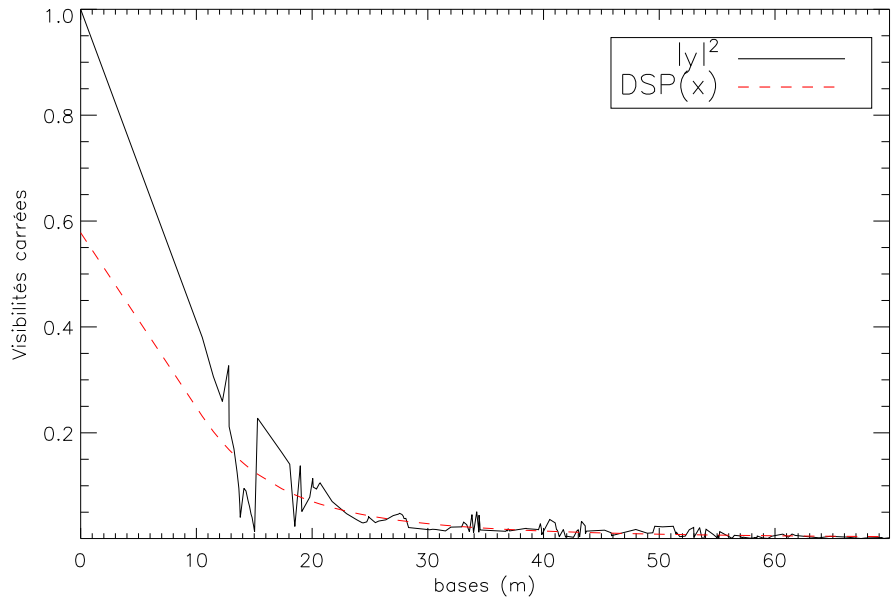



FIG. 2.6 – Modèle de DSP.

#### 2.4.3.3 Calcul automatique d'un terme de régularisation L2

Comme l'ont montré Blanc *et al.* (2003) dans le cadre de l'estimation d'aberrations par diversité de phase, les mesures interférométriques à « quelques » fréquences spatiales permettent d'ajuster les paramètres de la régularisation par DSP, qui, elle, porte sur toutes les fréquences

de la grille d'échantillonnage de la TF discrète :

$$\mathcal{J}_{\text{prior}}(\boldsymbol{x}) \propto \sum_{\boldsymbol{\nu}} \frac{|\widetilde{\boldsymbol{x}}(\nu) - \widetilde{\boldsymbol{x}}^0(\nu)|^2}{\text{DSP}(\nu)}$$

Les auteurs Conan *et al.* (1998) notent que $K$ règle le **poids** $\mu$ de la régularisation (Équation 1.22 page 37), qui peut donc être estimé directement sur les données.
Le paramètre ainsi déterminé serait asymptotiquement optimal (pour un grand nombre de mesures) si le modèle de bruit était effectivement blanc gaussien et uniforme, et si l'on n'imposait pas la positivité. Cette méthode permet néanmoins d'obtenir une valeur proche de l'optimum.

## 2.5 Validation sur simulations de RIME

Nous avons alors procédé à des simulations, dans le but de
- valider la capacité de RIME à reconstruire une image fidèle à l'objet d'intérêt, sur données simulées ou expérimentales ;
- quantifier l'apport de notre « convexification » du critère, vis-à-vis d'une approximation circulaire plus classique ;
- mesurer l'impact du développement au premier ordre effectué sur les paramètres du critère obtenu par approximation optimale (Section 2.3.2.4 page 52) ;
- vérifier que la méthode de réglage automatique des paramètres de régularisation (Section 1.4.2 page 37) conduit bien à des estimées de l'objet proche du réglage optimal de ces paramètres.

### 2.5.1 Données simulées

#### 2.5.1.1 L'objet

L'objectif de ces simulations était de valider la méthode de reconstruction dans une situation aussi réaliste que possible. Nous avons donc choisi comme objet de test un modèle de source astrophysique, en l'occurrence le modèle de micro jet de l'étoile Ru Lupus, proposé par Paulo Garcia et al. Garcia *et al.* (2001) (Figure 2.7).

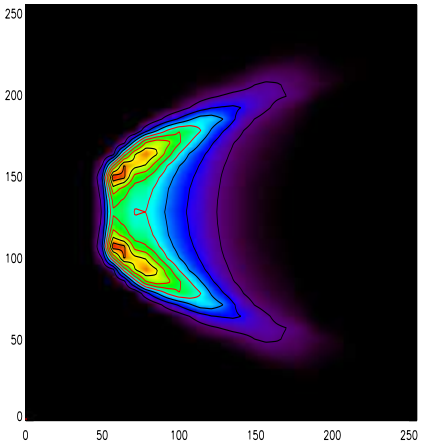



FIG. 2.7 – Modèle de Micro-jet de l'étoile Ru Lupus (Garcia *et al.*, 2001)

#### 2.5.1.2 La couverture fréquentielle

Nous avons simulé la couverture fréquentielle obtenue avec 3 des télescopes du VLTI, pour 6 nuits d'observation de la même source, avec 20 points de mesure par nuit. La couverture fréquentielle est optimiste, en ce sens qu'elle est plus dense que celle que l'on obtiendra réellement lorsque le VLTI sera opérationnel (limitée par le nombre de nuits d'observation disponibles, la position de la source dans le ciel, etc...). En effet, il est nécessaire d'avoir « beaucoup » de données pour pouvoir juger de la qualité du critère d'attache aux données choisi. Si l'on dispose de très peu de données, c'est le choix de la régularisation qui sera prépondérant dans la reconstruction obtenue. A l'inverse, si l'on dispose de très nombreuses données à bon rapport signal à bruit, toutes les reconstructions seront quasi-parfaites, et difficile à comparer. Nous avons

tenté de trouver une configuration suffisamment réaliste pour avoir valeur d'exemple, tout en permettant de comparer les critères d'attache aux données considérés (Figure 2.8).

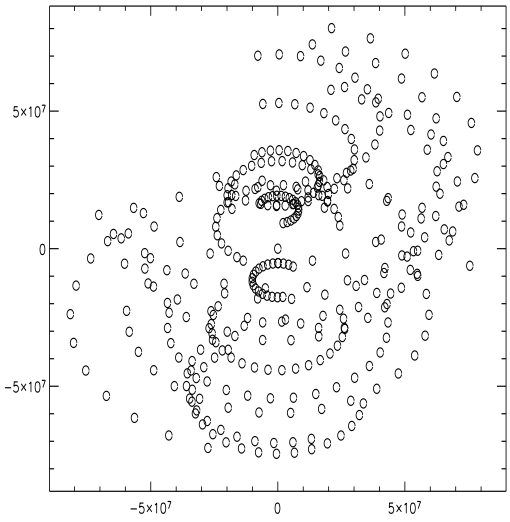



FIG. 2.8 – Couverture fréquentielle obtenue avec 3 des télescopes du VLTI, pour 6 nuits d'observation de la même source, avec 20 points de mesure par nuit.

#### 2.5.1.3 Le modèle de bruit

Les visibilités complexes sont bruitées selon le modèle

$$\boldsymbol{a}^d = \boldsymbol{a}(\boldsymbol{x}) + \boldsymbol{a}_b$$
$$\boldsymbol{\phi}^d = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\phi}_b$$

$\boldsymbol{a}_b$ suit une statistique gaussienne decorrélée d'écart-type $\boldsymbol{\sigma_a} = 4{,}65.10^{-2}\boldsymbol{a}^d$.
$\boldsymbol{a}_\phi$ suit une statistique gaussienne decorrélée d'écart-type $\boldsymbol{\sigma_\phi} = 0{,}27$ radians.

### 2.5.2 Modèle de DSP

Le modèle de DSP de l'objet est estimée directement sur les données (cf. Section 1.4.2 page 37). Sur la figure 2.4.3.2 page 60, on peut constater que le modèle paramétrique de la DSP de l'objet permet de décrire correctement les données.

### 2.5.3 Comparaison des approximations elliptiques et circulaire

#### 2.5.3.1 Comparaison de l'approximation optimale et de son développement limité au premier ordre

Nous avions proposé (Section 2.3.2.4 page 52) un développement limité de l'approximation quadratique optimale. Sur l'ensemble des simulations réalisées, aucune différence de reconstruction entre ces deux méthodes n'a été constatée.
Il est possible qu'avec un rapport signal sur bruit (RSB) bien plus faible, le développement limité que nous avons utilisé ne soit plus valide.Cependant, avec un faible RSB, le terme d'*a*

*priori* devient prépondérant sur le terme d'attache aux données, et la différence de performances entre modèles de bruit s'atténue.
Dans la suite de ce manuscrit, on ne considérera que l'approximation elliptique développée au premier ordre.

#### 2.5.3.2 Comparaison des approximations elliptique et circulaire

L'objectif suivant a été de quantifier l'apport de notre approximation convexe du critère, vis-à-vis d'une approximation circulaire. Pour cela, nous avons généré 100 séries de visibilités complexes, chacune affectée d'un échantillon de bruit différent.
Pour chaque jeu de données, nous avons effectué deux reconstructions, l'une avec l'approximation circulaire, l'autre avec l'approximation elliptique. Pour quantifier la différence de qualité entre une reconstruction elliptique $\hat{\boldsymbol{x}}_{RIME}$ et une reconstruction circulaire $\hat{\boldsymbol{x}}_{circ}$ de l'objet $\boldsymbol{x}^0$, nous avons considéré l'**Amélioration de l'Erreur Quadratique Moyenne** (AEQM) en décibels (dB) définie par :

$$AEQM = 20 \log_{10} \left( \frac{\sum (\hat{\boldsymbol{x}}_{circ} - \boldsymbol{x}^0)^2}{\sum (\hat{\boldsymbol{x}}_{RIME} - \boldsymbol{x}^0)^2} \right)$$

Une AEQM positive correspond à une reconstruction meilleure avec notre approximation elliptique.

La figure (2.9) présente l'histogramme de répartition de l'AEQM pour les 100 tirages de bruit.

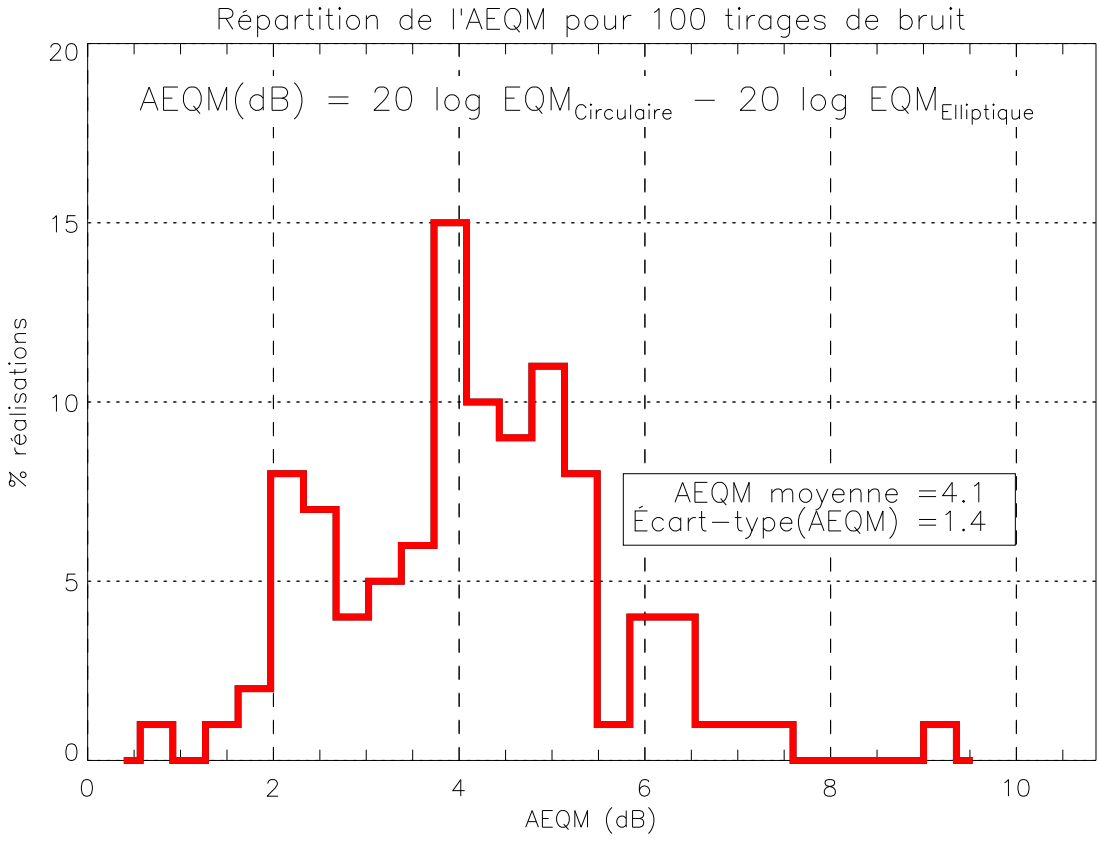



FIG. 2.9 – Histogramme de répartition de l'AEQM

L'amélioration est de 4 dB en moyenne, et 95% des reconstructions présentent une AEQM de plus de 2 dB. On considère qu'une différence de qualité entre deux images est visible pour une AEQM supérieure à 0,5dB. Nous pouvons donc conclure que notre approximation elliptique est bien plus performante que l'approximation circulaire, en termes de qualité d'image reconstruite.

#### 2.5.3.3 Illustration

Pour mieux illustrer l'intérêt de l'utilisation de l'approximation elliptique, nous avons sélectionné parmi les 100 tirages de bruit 3 tirages particuliers :

- Un cas **moyen**, pour lequel l'AEQM est proche de la moyenne des AEQM (4,16dB) ;
- Un **pire** cas, pour lequel l'AEQM est minimum (0,56dB) ;
- Un **meilleur** cas, pour lequel l'AEQM est maximum (9,35dB).

Comme le montre la figure Fig 2.10, notre méthode apporte clairement une amélioration visible de la reconstruction sur une large majorité des tirages effectués.

### 2.5.4 Influence de l'hyper-paramètre

Comme nous l'avons mentionné précédemment, ces images ont été reconstruites avec un poids de régularisation réglé automatiquement. Pour mesurer l'influence de la régularisation sur l'AEQM, nous avons réalisé les même simulations (c'est-à-dire 100 reconstructions pour chaque méthode, pour chaque hyper paramètre $\mu$)

- pour un poids de régularisation 10 fois plus faible : $\mu = 0,1$ ;
- pour un poids de régularisation 10 fois plus fort : $\mu = 10$.

Le tableau (2.1) montre les moyennes de l'AEQM et leur écart-type sur 100 réalisations.

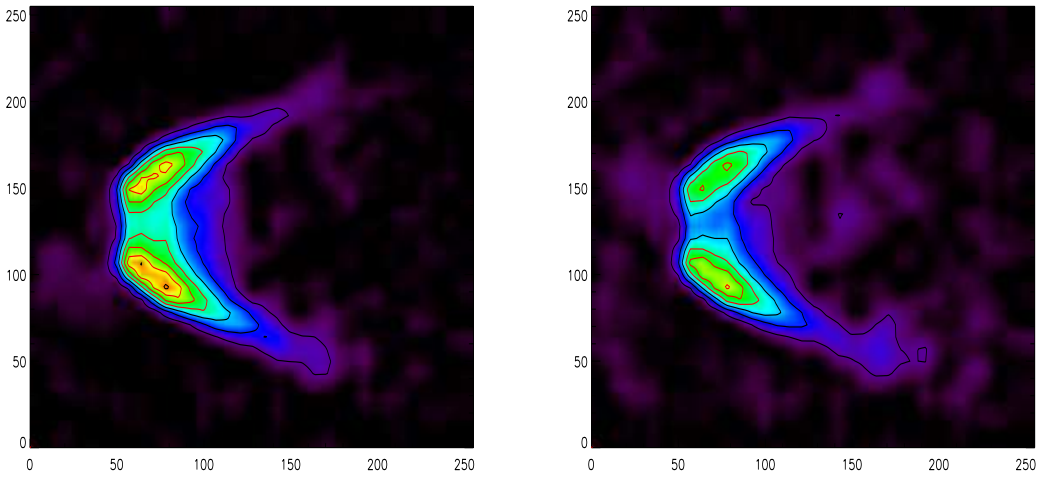
| $\mu$ | $\langle$AEQM$\rangle$ | Écart-type(AEQM) |
|---|---|---|
| 0.1 | 2.6 | 1.8 |
| 1 | 4.1 | 1.4 |
| 10 | 7.8 | 0.9 |

TAB. 2.1 – – Influence du paramètre de régularisation sur l'AEQM

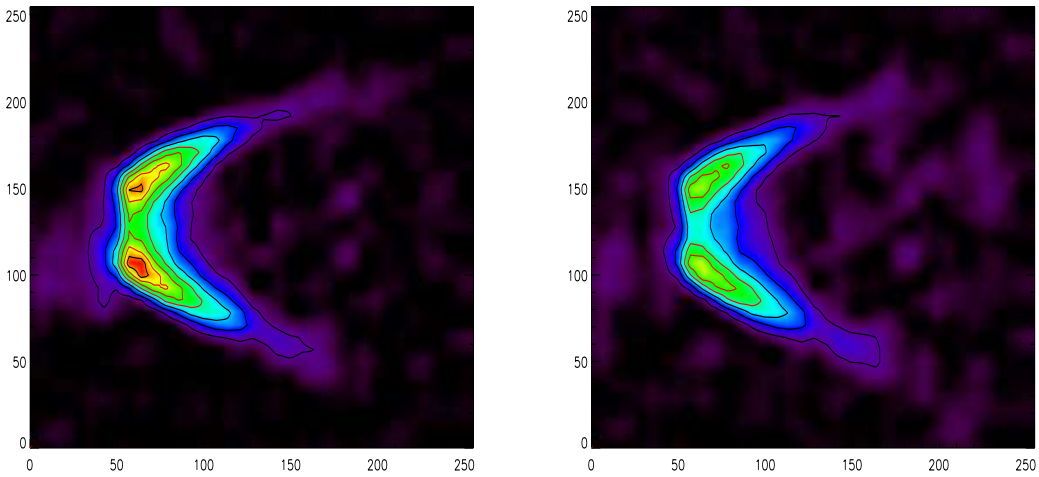
Pour une décade de part et d'autre de la valeur nominale de l'hyper paramètre $\mu$, on constate toujours une claire amélioration de la reconstruction avec notre méthode.

### 2.5.5 Conclusion des simulations

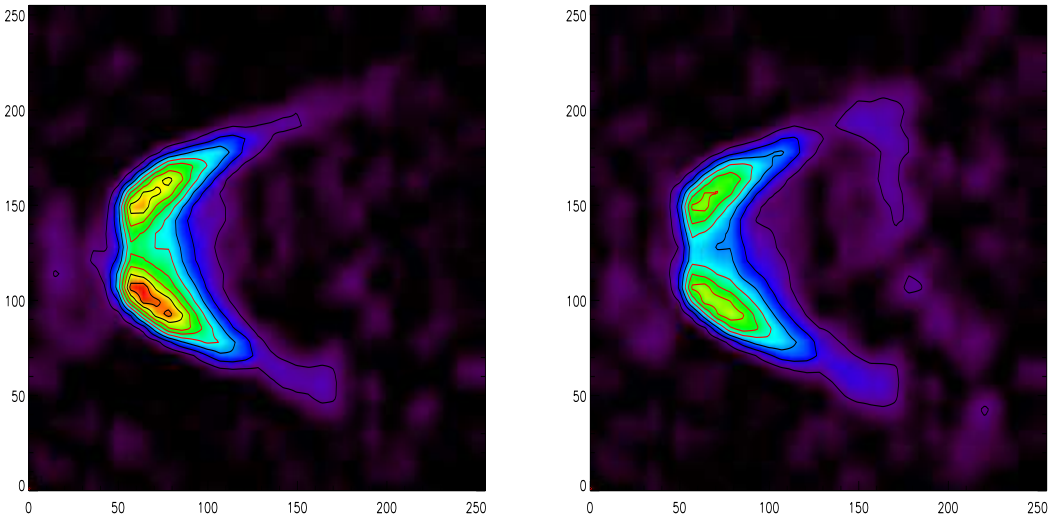
Nous avons prouvé l'intérêt de notre méthode sur simulations. Le modèle de bruit étant construit à partir des données en début de minimisation, la méthode elliptique n'est pas plus coûteuse en temps qu'une méthode utilisant une approximation circulaire. De plus, pour toutes les simulations réalisées, l'approximation elliptique s'est toujours comporté au moins aussi bien que l'approximation circulaire. Dans toute la suite, le modèle de bruit sera donc approché par approximation elliptique.



(a) EQM$= 19,3\%$ (b)EQM$= 33,0\%$



(c) EQM$= 32.6\%$ (d)EQM$= 33,7\%$



(e) EQM$= 27,7\%$ (f) EQM$= 35,2\%$

FIG. 2.10 – Résultats de simulation. A gauche, reconstruction avec approximation elliptique, à droite avec approximation circulaire. Meilleur cas (AEQM=9,35dB) : (a) et (b) ; Pire cas (AEQM=0,56dB) : (c) et (d) ; Cas moyen (AEQM=4,16dB) : (e) et (f). $256 \times 256$ pixels, de 0.2 mas. de côté.

## 2.6 Traitement des données issues de l'interféromètre fibré de l'IRCOM

Nous avons alors voulu confronter l'algorithme RIME, validé sur simulations, à des données interférométriques expérimentales. En effet, le modèle de bruit « vrai » que nous avons considéré jusqu'à présent est en pratique estimé à partir des données, et limité à des statistiques des premier et deuxième ordres. Il reste à vérifier que notre méthode est suffisamment robuste à des erreurs de modélisation du bruit.

### 2.6.1 Présentation du dispositif expérimental

Les données expérimentales auxquelles nous avons eu accès ont été recueillies par Delage *et al.* (1999) sur l'interféromètre fibré à 3 télescopes [Delage *et al.* (2000)] de l'Institut de Recherche en Communications Optiques et Microondes(IRCOM). Cet interféromètre est reconfigurable, de telle sorte qu'il permet de réaliser des bases multiples d'une base élémentaire de $3mm$ (Cf Figure 2.11 page 67). La rotation de l'interféromètre est simulée par rotations suc-

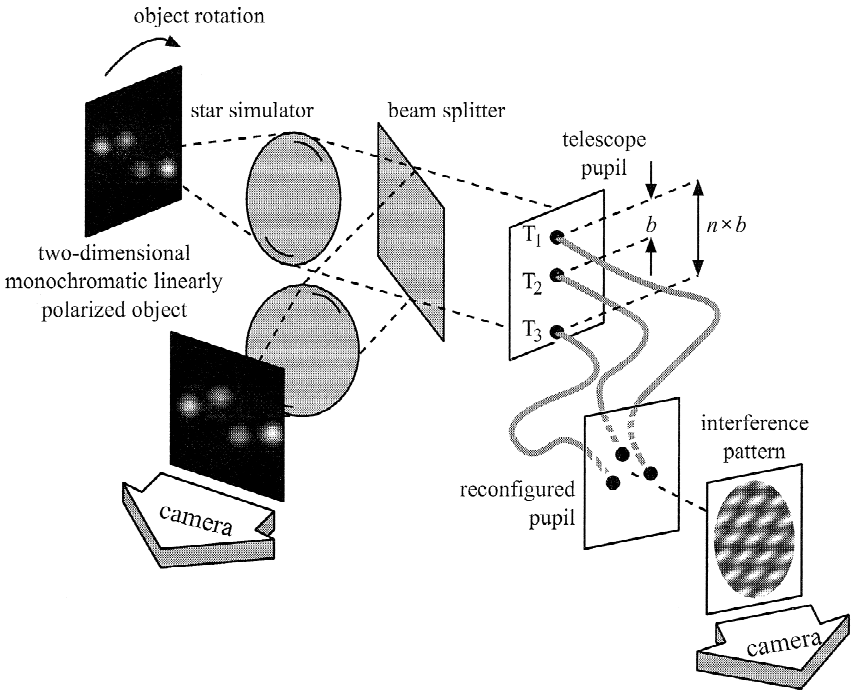



FIG. 2.11 – Dispositif expérimental. Delage *et al.* (1999)

cessives de l'objet, réalisé à l'aide de quatre fibres mono-modes éclairées par un laser He-Ne. La couverture $(u, v)$ ainsi synthétisée est présentée Figure 2.12.

Une voie d'imagerie permet d'enregistrer l'objet, auquel on comparera les reconstructions.

### 2.6.2 Justification du modèle de bruit proposé

Ce dispositif permet de déterminer les phases de visibilité à partir des clôtures de phase, selon la méthode évoquée section 1.2.4.1 page 27. En effet, si l'on considère une position an-

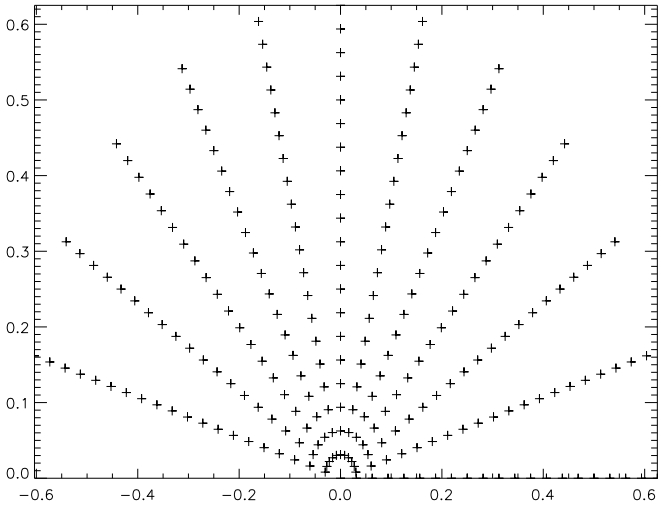



FIG. 2.12 – Couverture fréquentielle expérimentale.

gulaire de l'objet, les 3 bases obtenues avec l'interféromètre à l'instant $n$ sont (Cf. Figure 2.13) :

$$\begin{cases} \boldsymbol{u}_{12} \\ \boldsymbol{u}_{23} = n\,\boldsymbol{u}_{12} \\ \boldsymbol{u}_{13} = (n+1)\boldsymbol{u}_{12} \end{cases}$$

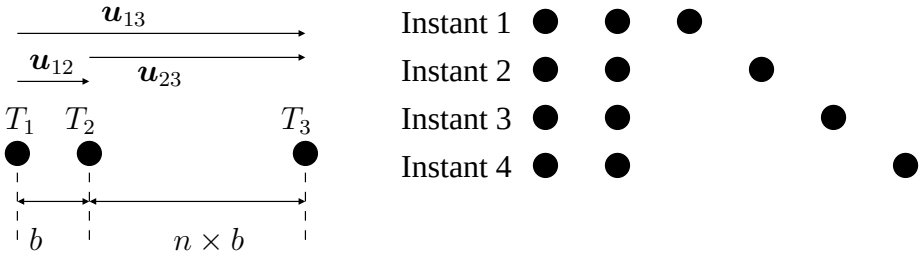



FIG. 2.13 – Configuration pupillaire.

A l'instant $n$, on mesure la clôture de phase

$$\begin{cases} \beta(1) = 2\phi(\boldsymbol{\nu}_{12}) + \phi(2\boldsymbol{\nu}_{12}) \\ \beta(2) = \phi(\boldsymbol{\nu}_{12}) + \phi(2\boldsymbol{\nu}_{12}) + \phi(3\boldsymbol{\nu}_{12}) \\ \beta(3) = \phi(\boldsymbol{\nu}_{12}) + \phi(3\boldsymbol{\nu}_{12}) + \phi(4\boldsymbol{\nu}_{12}) \\ \quad\vdots \\ \beta(n) = \phi(\boldsymbol{\nu}_{12}) + \phi(n\boldsymbol{\nu}_{12}) + \phi((n+1)\boldsymbol{\nu}_{12}). \end{cases}$$

Pour alléger les calculs, on pose $\phi_n = \phi(n\boldsymbol{\nu}_{12})$.L'équation précédente s'écrit avec cette notation

$\beta(n) = \phi_1 + \phi_n + \phi_{n+1}$. On a :

$$\begin{aligned}
\sum_{k=1}^{n}(-1)^{n-k}\beta(k) &= \underbrace{\sum_{k=1}^{n}(-1)^{n-k}}_{\epsilon}\phi_1 + \sum_{k=1}^{n}(-1)^{n-k}\phi_n + \sum_{k=1}^{n}(-1)^{n-k}\phi_{n+1} \\
&= \epsilon\phi_1 + \left((-1)^{n-1}\phi_1 + \sum_{k=2}^{n}(-1)^{n-k}\phi_n\right) + \left(\phi_{n+1} + \sum_{k=1}^{n-1}(-1)^{n-k}\phi_{n+1}\right) \\
&= \underbrace{\left(\epsilon + (-1)^{n-1}\right)}_{\epsilon'}\phi_1 + \sum_{k=2}^{n}(-1)^{n-k}\phi_n + \phi_{n+1} - \sum_{k=2}^{n}(-1)^{n-k}\phi_n \\
&= \epsilon'\phi_1 + \phi_{n+1}
\end{aligned}$$

On peut donc déduire les phases de visibilités à partir des clôtures selon :

$$\phi((n+1)\boldsymbol{\nu}_{12}) = \sum_{k=1}^{n}(-1)^{n-k}\beta(k) - \epsilon'\phi_1$$

On fait alors l'approximation $\phi(\boldsymbol{\nu}_{12}) \simeq 0$ (aux plus faibles bases, l'objet est très faiblement résolu, et sa phase est donc faible), et on suppose que le bruit sur chaque clôture est $\sigma_\beta$. On a alors :

$$\phi((n+1)\boldsymbol{\nu}_{12}) \simeq \sum_{k=1}^{n}(-1)^{n-k}\beta(k),$$

et l'écart-type sur $\phi((n+1)\boldsymbol{\nu}_{12})$ est :

$$\sigma_{\phi,n} = \sqrt{n}\sigma_\beta$$

Le modèle de formation des données utilisé est donc :

$$\begin{cases} \boldsymbol{a}^d = |\boldsymbol{H}\boldsymbol{x}| + \boldsymbol{a}_b \\ \boldsymbol{\phi}^d = \arg \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\phi}_b \end{cases}$$

où les deux distributions du bruit sont gaussiennes centrées et indépendantes, avec les propriétés suivantes :

$$\begin{cases} \boldsymbol{\sigma}_a = 0,11\boldsymbol{a}^d \\ \boldsymbol{\sigma}_\phi = 0,06\sqrt{\boldsymbol{n}}\ (\text{rad}) \end{cases}$$

où $\boldsymbol{n}$ est un vecteur contenant, pour la mesure courante, le nombre de translations d'une longueur égale à la plus petite base.

### 2.6.3 Analyse des reconstructions obtenues

Les figures (2.14) et (2.15) présentent une comparaison de l'objet « vrai » et de l'objet reconstruit avec RIME. Les 4 sources de l'objet ont bien été restituées. Néanmoins, les pics ont été trop lissés par la régularisation quadratique.

Le tableau (2.2) montre que la reconstruction est correcte en termes de position relative des sources.

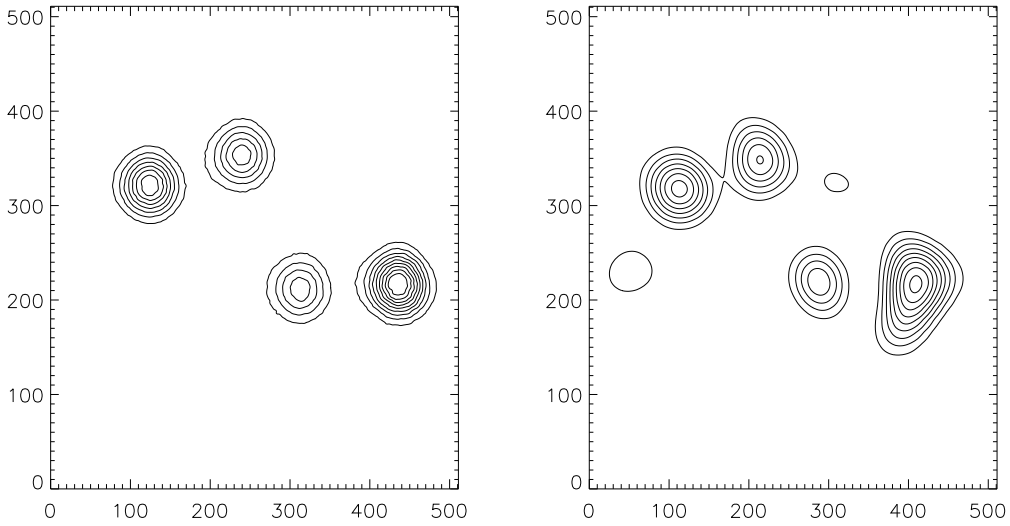



FIG. 2.14 – Objet vrai (gauche) et objet reconstruit (droite). Niveaux des contours : $10\%, 20\%, \ldots, 100\%$ du maximum.

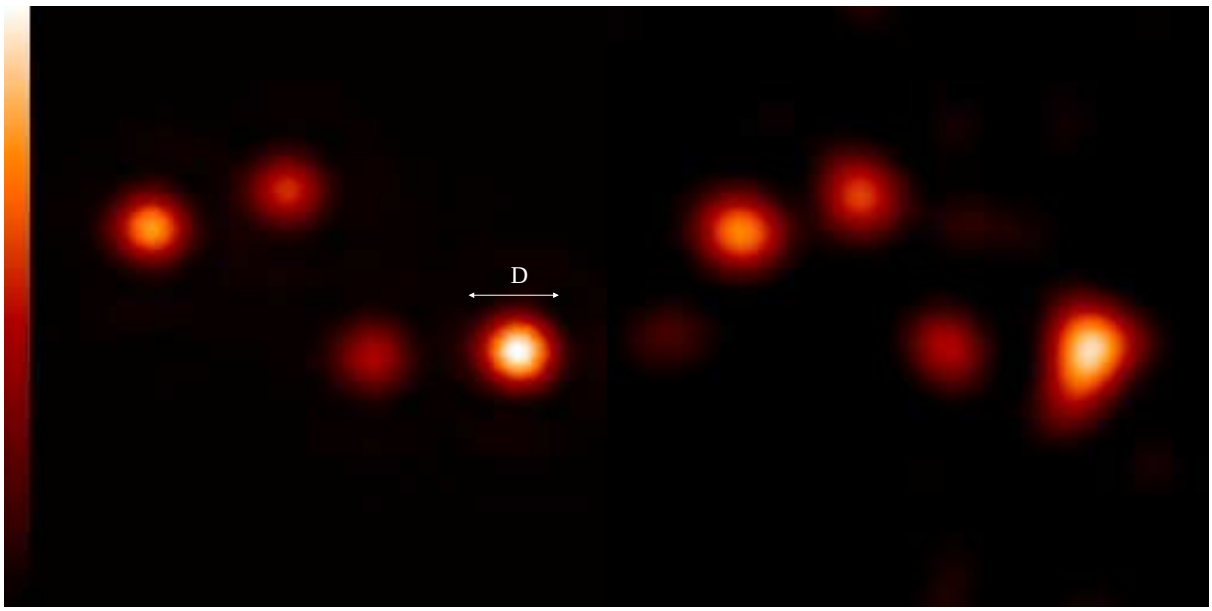



FIG. 2.15 – Objet vrai (gauche) et objet reconstruit (droite). D est le diamètre de la composante principale, utilisé dans le tableau 2.2.

## 2.7 Conclusion

Nous avons donc prouvé que notre méthode est plus précise qu'une approximation circulaire, que notre stratégie de reconstruction non supervisée est efficace, et que l'algorithme RIME permet des reconstructions satisfaisantes sur des données expérimentales.
Ces résultats ont été acceptés pour publication dans le Journal of the Optical Society of America A et figurent en Annexe F.

TAB. 2.2 – – Positions relatives et flux des 3 sources les plus faibles par rapport à la plus brillante.

| | | Erreur de position rapportée au diamètre $D$ de la source principale (figure 2.15) | | Intensité ($-2.5 \log_{10}$ flux) | |
|---|---|---|---|---|---|
| source # | (x,y) | $\Delta x/D$ | $\Delta y/D$ | vrai | reconstruit |
| 1 | $(0, 0)$ | 0 | 0 | 1.33 | $1.07 \pm 19\%$ |
| 2 | $(-237, 81)$ | 0.13 | 0.05 | 1.89 | $1.72 \pm 9\%$ |
| 3 | $(-152, 105)$ | 0.03 | 0.06 | 2.20 | $1.95 \pm 11\%$ |
| 4 | $(-94, -4)$ | 0.01 | 0.08 | 2.44 | $2.41 \pm 2\%$ |

# Chapitre 3

# Modèle direct des données interférométriques affectées par la turbulence

Dans ce chapitre et les suivants, nous nous plaçons dans le cas d'un interféromètre au sol, avec une turbulence non corrigée, et trop forte pour être assimilée à un bruit. C'est le cas d'intérêt pour les interféromètres VLTI, IOTA, NPOI, etc...
En présence de la turbulence atmosphérique, les observables interférométriques ne sont plus les *amplitudes* et *phases* de visibilités complexes, mais les *visibilités carrées* et les *clôtures de phase*.
Après avoir rappelé les raisons pour lesquelles on mesure des visibilités carrées plutôt que des visibilités en module, nous montrons section 3.1 que l'on peut se ramener du modèle de formation des données « vrai », c'est-à-dire en visibilités carrées, à un modèle de formation de données sur les visibilités en module.
Cependant, l'enjeu principal de la prise en compte de la turbulence est l'étude de la structure **en phase** du problème. Dans ce domaine, un travail exhaustif de mise à jour des structures des problèmes en phase en interférométrie a été effectué par Lannes (1990). Nous proposons dans ce chapitre une description de ces concepts appliqués au problème que nous traitons :

- La section 3.2 caractérise les informations de phase indépendantes de la turbulence (dont les clôtures sont un exemple particulier), et définit la structure des informations de phase manquantes ;
- la section 3.3 établit le critère d'attache aux données formé pour un modèle de données en clôtures de phase et modules de visibilités (ou visibilités carrées), et pointe les défauts d'un tel critère, qui masque les dégénérescences du problème en phase ;
- Dans la section 3.4, nous proposons un autre choix de modèle direct en phases, où le manque d'information de phase figure explicitement sous forme d'aberrations myopes. Nous appliquons ensuite ce choix de modèle à une configuration redondante à 3 télescopes ;
- Enfin, nous rassemblons (section 3.5) le modèle direct en phases de la section 3.4 et le modèle en modules de visibilités déduit du modèle en visibilités carrées (section 3.1). Nous formons ainsi un modèle de données myope sur les phases et amplitudes de visibilités complexes.

## 3.1 Des visibilités carrées aux visibilités en module

### 3.1.1 Pourquoi mesure-t-on des visibilités carrées ?

Le problème des visibilités en amplitude est que leur mesure est biaisée, même si le bruit en amplitude est à moyenne nulle. En effet

$$E\left\{|a(\boldsymbol{x}) + a_b|\right\} \neq |a(\boldsymbol{x})| + E\left\{a_b\right\}$$

On mesure donc plutôt des visibilités carrées. En admettant que le bruit en amplitude $a_b$ est à moyenne nulle, on a :

$$\begin{aligned} E\left\{|a(\boldsymbol{x}) + a_b|^2\right\} &= E\left\{a^2(\boldsymbol{x}) + a_b^2 + 2a(\boldsymbol{x})a_b\right\} \\ &= a^2(\boldsymbol{x}) + E\left\{a_b^2\right\} + 2a(\boldsymbol{x})E\left\{a_b\right\} \\ &= a^2(\boldsymbol{x}) + \text{Var}\left\{a_b\right\} \end{aligned}$$

Il est possible d'estimer la variance du bruit, en faisant l'hypothèse que celui-ci provient de deux sources d'erreur décorrelées :

$$\text{Var}\left\{a_b\right\} = \text{Var}\left\{a_b\right\}^{\text{détection}} + \text{Var}\left\{a_b\right\}^{\text{photon}}$$

- Le terme $\text{Var}\left\{a_b\right\}^{\text{détection}}$ est dû à un bruit instrumental. Il peut être mesuré en faisant un « dark field », c'est à dire en bouchant les ouvertures des télescopes. Les fluctuations d'amplitude sont alors dues uniquement au bruit instrumental ;
- Le terme $\text{Var}\left\{a_b\right\}^{\text{photon}}$ est dû à un bruit de photons de statistique poissonnienne, dont on estime théoriquement la variance.

$a^2$ est le module carré de la TF de l'objet, c'est-à-dire son spectre de puissance, que l'on notera $s$ :

$$s \triangleq a^2.$$

Avec ceci, on peut donc *débiaiser* la mesure des amplitudes carrées :

$$\boldsymbol{s}^{data}(t) = E\left\{|a(\boldsymbol{x}) + a_b|^2\right\} - \text{Var}\left\{a_b\right\}^{\text{détection}} - \text{Var}\left\{a_b\right\}^{\text{photon}} \tag{3.1}$$

La variance sur les amplitudes carrées $\text{Var}\left\{\boldsymbol{s}\right\} = \boldsymbol{\sigma}_s^2$ est estimée à l'aide de l'équation 3.1.

Les visibilités carrées constituent donc de meilleures observables que les visibilités en module. Dans la suite, nous montrons que l'on peut se ramener du modèle de formation des données en visibilités carrées à un modèle de formation de données sur les visibilités en module.

### 3.1.2 Déduire un modèle de données sur les amplitudes de visibilité

Soit $s$ une variable gaussienne d'écart-type $\sigma_s$ et de moyenne $\bar{s}$. Alors la loi de probabilité de $s$ s'écrit :

$$p(s) \propto \exp -\frac{1}{2}\frac{(s-\bar{s})^2}{\sigma_s^2}$$

Le fait d'avoir une loi gaussienne pour une visibilité carrée $s$, donc positive, est contradictoire. Cependant, cette approximation est acceptable sous l'hypothèse que la loi est « piquée », c'est-à-dire que la quasi totalité de l'énergie est concentrée dans les valeurs positives :

$$\bar{s} \gg \sigma_s \tag{3.2}$$

On souhaite faire un changement de variable dans cette distribution de probabilités, pour passer d'une loi sur $s$ à une loi sur $a$. Pour cela, on écrit :

$$p(s)\mathrm{d}s = p(a)\mathrm{d}a,$$

avec $a$ la racine carrée de $s$ :

$$a \triangleq \begin{cases} \sqrt{s} \text{ si } s > 0; \\ 0 \text{ si } s \leq 0 \end{cases}$$

On obtient donc :

$$\begin{aligned} p(a) &\propto p(s)\frac{\mathrm{d}s}{\mathrm{d}a} \\ &\propto \exp -\frac{1}{2}\frac{(s-\bar{s})^2}{\sigma_s^2}\frac{\mathrm{d}a^2}{\mathrm{d}a} \qquad\qquad \propto 2a\,\exp -\frac{1}{2}\frac{(s-\bar{s})^2}{\sigma_s^2} \end{aligned}$$

d'où

$$p(a) \propto a \exp -\frac{(a^2-\bar{s})^2}{2\sigma_s^2} \tag{3.3}$$

On souhaite approcher cette loi par une loi gaussienne :

$$p_g(a) \propto \exp -\frac{1}{2}\frac{(a-\bar{a})^2}{\sigma_a^2}$$

Pour cela, on considère une visibilité en amplitude $a$ proche du pic de l'exponentielle, c'est-à-dire en $\sqrt{\bar{s}}$ :

$$a = \sqrt{\bar{s}}(1+\delta) \tag{3.4}$$

On a :

$$\begin{aligned} a \exp -\frac{(a^2-\bar{s})^2}{\sigma_s^2} &= \sqrt{\bar{s}}(1+\delta)\exp -\frac{((\sqrt{\bar{s}}(1+\delta))^2-\bar{s})^2}{2\sigma_s^2} \\ &= \sqrt{\bar{s}}(1+\delta)\exp -\frac{(\bar{s}(1+\delta)^2-\bar{s})^2}{2\sigma_s^2} \\ &= \sqrt{\bar{s}}(1+\delta)\exp -\frac{(\bar{s}(1+2\delta+\delta^2)-\bar{s})^2}{2\sigma_s^2} \\ &= \sqrt{\bar{s}}(1+\delta)\exp -\frac{(\bar{s}(2\delta+\delta^2))^2}{2\sigma_s^2} \\ &\simeq \sqrt{\bar{s}}(1+\delta)\exp -\frac{(\bar{s}(2\delta))^2}{2\sigma_s^2} \\ &\simeq \sqrt{\bar{s}}(1+\delta)\exp -\frac{(2\bar{s})^2}{2\sigma_s^2}\delta^2 \end{aligned}$$

Avec l'équation 3.4, il vient :

$$\delta^2 = \frac{(a-\sqrt{\bar{s}})^2}{\bar{s}}$$

On a donc :

$$\begin{aligned} a \exp -\frac{(a^2-\bar{s})^2}{2\sigma_s^2} &\simeq \sqrt{\bar{s}}(1+\delta)\exp -\frac{(2\bar{s})^2}{2\sigma_s^2}\frac{(a-\sqrt{\bar{s}})^2}{\bar{s}} \\ &\simeq \sqrt{\bar{s}}\exp -\frac{4\bar{s}}{2\sigma_s^2}(a-\sqrt{\bar{s}})^2 \end{aligned}$$

Avec l'équation 3.3, on voit donc, que autour de $\sqrt{\bar{s}}$, $a$ suit la loi approchée :

$$\begin{aligned} p_g(a) &\propto \exp -\frac{4\bar{s}}{2\sigma_s^2}(a-\sqrt{\bar{s}})^2 \\ &\propto \exp -\frac{1}{2}\left(\frac{\sigma_s}{2\sqrt{\bar{s}}}\right)^{-2}(a-\sqrt{\bar{s}})^2 \\ &\propto \exp -\frac{(a-\bar{a})^2}{2\sigma_a^2} \qquad (3.5) \\ \text{avec } \bar{a} = \sqrt{\bar{s}} \text{ et } \sigma_a &= \frac{\sigma_s}{2\bar{a}} \qquad (3.6) \end{aligned}$$

La figure 3.1 illustre le fait que le choix

$$\bar{a} = \sqrt{\bar{s}}, \quad \sigma_a = \frac{\sigma_s}{2\bar{a}} \qquad (3.7)$$

est satisfaisant. Comme $a$ est un module de visibilité complexe, nous ne comparons $p_g$ et $p$ que pour des valeurs de $a$ positives.

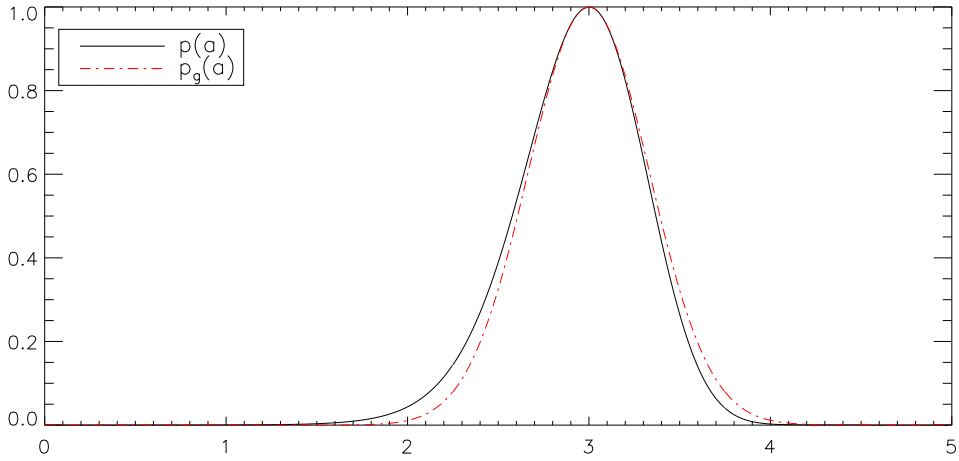



FIG. 3.1 – Statistique $p$ d'une variable aléatoire $a$ de carré $s$ gaussien et approximation gaussienne $p_g$.

Dans le cas qui nous occupe, on cherche à construire des pseudo mesures d'amplitude $\boldsymbol{a}^{data}$ et l'écart-type du bruit correspondant $\boldsymbol{\sigma}_a$, à partir de mesures d'amplitudes carrées $\boldsymbol{s}^{data}$ et de l'écart-type du bruit correspondant $\boldsymbol{\sigma}_s$. On fera donc le choix :

$$\boldsymbol{a}^{data} = \sqrt{\boldsymbol{s}^{data}}, \quad \boldsymbol{\sigma}_a = \frac{\boldsymbol{\sigma}_s}{2\sqrt{\boldsymbol{s}^{data}}} \qquad (3.8)$$

La matrice de covariance des modules de visibilité s'écrit

$$\boldsymbol{R}_a = \boldsymbol{W}_a^{-1} = \text{Diag}\left\{\boldsymbol{\sigma}_a^2\right\}.$$

## 3.2 Structures de l'information de phase dans le cas fortement turbulent

### 3.2.1 Informations de phase et noyau de $\boldsymbol{B}^{\mathrm{T}}$

#### 3.2.1.1 Informations de clôture

Dans le cas d'une turbulence non corrigée, et trop forte pour être assimilée à un bruit, cette dernière affecte les phases des interférogrammes de la façon suivante :

$$\boldsymbol{\phi}^{data}(t) = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}(t) + \boldsymbol{\phi}_b(t)$$

où $\boldsymbol{B}$ est l'opérateur permettant de passer des positions des télescopes aux bases de l'interféromètre (Eq. 1.12 page 24).
Nous avons vu que l'on peut former des clôtures de phase (section 1.2.3 page 26), c'est-à-dire des combinaisons linéaires des phases telles que l'influence des pistons turbulents $\boldsymbol{\varphi}(t)$ est réduite à 0 (Eq. 1.18 page 30) :

$$\boldsymbol{\beta}^{data}(t) = \boldsymbol{C}\boldsymbol{\phi}^{data}(t) = \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x}) + \underbrace{\boldsymbol{C}\boldsymbol{B}}_{=\boldsymbol{0}}\boldsymbol{\varphi}(t) + \boldsymbol{C}\boldsymbol{\phi}_b(t)$$

$\boldsymbol{C}$ est une matrice, dont les lignes correspondent aux combinaisons linéaires annulant l'influence des pistons turbulents. En d'autres termes, on a :

$$\boldsymbol{C}\boldsymbol{B} = \boldsymbol{0} \Rightarrow \boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{0} \Rightarrow \boldsymbol{C}^{\mathrm{T}} \in \ker \boldsymbol{B}^{\mathrm{T}}$$

Les lignes de $\boldsymbol{C}$ sont des éléments du **noyau de** $\boldsymbol{B}^{\mathrm{T}}$. On démontre (Cf Annexe B page 157) que l'ensemble des clôtures à 3 télescopes incluant le télescope 1 (ou tout autre télescope arbitrairement fixé) constitue une base de ce noyau, que l'on note $\boldsymbol{\beta}_{ind}$. **Toute combinaison linéaire annulant l'effet des pistons turbulents est une combinaison linéaire de clôtures à 3 télescopes incluant le télescope 1.** Donc le nombre d'informations indépendantes de la turbulence obtenues par clôture de phase est égal à la dimension du noyau de $\boldsymbol{B}^{\mathrm{T}}$. Nous démontrons dans l'annexe B qu'une base de ce noyau est l'ensemble des clôtures indépendantes à 3 télescopes $\boldsymbol{\beta}_{ind}$, et que son cardinal est $\binom{N_t-1}{2}$, d'où :

$$N_o = \dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \underbrace{\dim \boldsymbol{F}_{bases}}_{N_b} - \underbrace{\mathrm{rg}(\boldsymbol{B}^{\mathrm{T}})}_{infos manquantes} = \binom{N_t-1}{2} = \frac{(N_t-1)(N_t-2)}{2}. \tag{3.9}$$

#### 3.2.1.2 Informations de phase quelconques

Nous nous sommes posé la question de l'existence d'autres informations sur les phases indépendantes de la turbulence.

On a $\boldsymbol{\phi}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi} + \boldsymbol{\phi}_b$ et on cherche des fonctions $f$ telles que

$$f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\phi}_b + \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\phi}_b),\ \forall \boldsymbol{\varphi}.$$

Pour ces calculs, on peut ne pas tenir compte de $\boldsymbol{\phi}_b$. On cherche donc des fonctions $f$ telles que

$$f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}) = f(\boldsymbol{\phi}(\boldsymbol{x}),\ \forall \boldsymbol{\varphi}.$$

Si $f$ est linéaire, alors c'est une fonction de type clôture, donc générée par $\boldsymbol{\beta}_{ind}$. Puisqu'on veut moyenner les informations obtenues, il est raisonnable de nous limiter à des fonctions $f$

continues et dérivables.
on pose donc $g(\boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\varphi}) = f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi})$, avec $g$ continue et différentiable. Soient $N_b$ le nombre de bases. On a :

$$\begin{aligned} f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}) &= f(\boldsymbol{\phi}(\boldsymbol{x})), \forall \boldsymbol{\varphi} \\ \Rightarrow g(\boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\varphi}) &= g(\boldsymbol{\phi}(\boldsymbol{x}), \mathbf{0}_{N_b}) \\ \Rightarrow \frac{\partial g(\boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} &= \mathbf{0}_{N_b} \\ \Rightarrow \frac{\partial f(\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} &= \mathbf{0}_{N_b} \\ \Rightarrow \boldsymbol{B}^{\mathrm{T}} f' &= \mathbf{0}_{N_b} \end{aligned}$$

Donc $f'$ est dans le noyau de $\boldsymbol{B}^{\mathrm{T}}$. Soient $\{c_{1ij}\}_{2 \leq i < j \leq n}$ une base de $\boldsymbol{B}^{\mathrm{T}}$ formée des clôtures à trois télescopes contenant le télescope 1 (Cf. Annexe B). Alors $f'$ est de la forme :

$$f' = f'_{123} \circ c_{123} + f'_{134} \circ c_{134} + \ldots + f'_{1ij} \circ c_{1ij} + \ldots$$

En intégrant cette équation, on obtient :

$$f(\boldsymbol{\phi}) = f_{123}(c_{123}^{\mathrm{T}}\boldsymbol{\phi}) + f_{134}(c_{134}^{\mathrm{T}}\boldsymbol{\phi}) + \ldots + f_{1ij}(c_{1ij}^{\mathrm{T}}\boldsymbol{\phi})$$

**Conclusion :**

- Les informations que l'on peut obtenir à partir des phases perturbées par la turbulence sont des fonctions des clôtures indépendantes. En pratique, au lieu de mesurer la clôture $\phi_{12}^{data} + \phi_{23}^{data} - \phi_{13}^{data}$, on mesure le *bispectre* correspondant $a_{12}^{data} a_{23}^{data} a_{13}^{data} \exp \boldsymbol{i}[\phi_{12}^{data} + \phi_{23}^{data} - \phi_{13}^{data}]$.
- Le nombre d'informations de phase indépendantes de la turbulence est (cf. Eq. 3.9) :

$$\boxed{N_o = \dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \frac{(N_t - 1)(N_t - 2)}{2}}$$

où $N_t$ est le nombre de télescopes de l'interféromètre.

#### 3.2.1.3 Degré d'information de phase

Considérons un interféromètre à $N_t$ télescopes.

Il y a autant de bases que de façons de choisir 2 points parmi $N_t$, soit

$$N_b = \mathrm{card}(\boldsymbol{\phi}(\boldsymbol{x})) = \binom{N_t}{2} = \frac{N_t(N_t - 1)}{2}.$$

Le nombre d'informations de phase indépendantes mesurable correspond à la dimension du noyau de $\boldsymbol{B}^{\mathrm{T}}$(Eq. 3.9) :

$$N_o = \dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \frac{(N_t - 1)(N_t - 2)}{2}$$

. On définit le degré d'information de phase (Lannes, 1998b) par

$$d = \frac{N_o}{N_b} = \frac{N_t - 2}{N_t} \tag{3.10}$$

Le nombre d'informations manquantes est $N_m = N_b - N_o = N_t - 1$. Ces résultats sont regroupés dans la table 3.1.

| $N_t$ | $N_b$ | $N_o$ | $d$ | $N_m$ |
|---|---|---|---|---|
| 2 | 1 | 0 | 0 | 1 |
| 3 | 3 | 1 | 33% | 2 |
| 4 | 6 | 3 | 50% | 3 |
| 5 | 10 | 6 | 60% | 4 |
| 10 | 45 | 36 | 80% | 9 |
| 20 | 190 | 171 | 90% | 19 |
| $n$ | $\frac{n(n-1)}{2}$ | $\frac{(n-1)(n-2)}{2}$ | $\frac{n-2}{n}$ | $(n-1)$ |

TAB. 3.1 – – Nombre de bases, d'observables de phase indépendants, degré d'information de phase et nombre minimal d'informations manquantes

#### 3.2.1.4 Le cas « Weak-phase » : les interféromètres à 3 et 4 télescopes

On parle de configuration « Weak phase » (Lannes, 1998b) quand le degré d'information de phase est inférieur ou égal à $50\%$, c'est-à-dire quand on dispose d'un interféromètre à 3 ou 4 télescopes. Pour des raisons techniques, il est difficile de combiner interférométriquement plus de 4 télescopes, c'est-à-dire de former simultanément plus de 6 clôtures de phase. Dans les années à venir, les problèmes de reconstruction d'image à partir de modules de visibilité et de clôtures de phase se placeront majoritairement dans ce cadre.

### 3.2.2 Informations manquantes, pistons turbulents et rang de $\boldsymbol{B}$

Les clôtures mesurées (longue pose) sont les phases de bispectres instantanés (courte pose) moyennés (cf. section 1.2.3 page 26). La phase instantanée d'un bispectre est la somme de trois phases courte pose extraites des interférogrammes correspondant. Ces phases $\boldsymbol{\phi}^{data}(t)$ affectées par les pistons turbulents $\boldsymbol{\varphi}(t)$ et par un bruit en phase $\boldsymbol{\phi}_b(t)$ s'écrivent :

$$\boldsymbol{\phi}^{data}(t) = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}(t) + \boldsymbol{\phi}_b(t).$$

Pour $N_{tels} = 4$, l'opérateur $\boldsymbol{B}$ s'écrit :

$$\boldsymbol{B} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Avec la structure de $\boldsymbol{B}$, on retrouve le constat qu'un piston global sur tous les $\boldsymbol{\varphi}(t)$ ne modifie pas la phase mesurée $\boldsymbol{\phi}^{data}(t)$. En effet, ce piston global $\varphi^0[1,\ldots,1]^{\mathrm{T}}$ est dans le noyau de l'opérateur $\boldsymbol{B}$ :

$$\boldsymbol{B}(\boldsymbol{\varphi}(t) + \varphi^0[1,\ldots,1]^{\mathrm{T}}) = \boldsymbol{B}\boldsymbol{\varphi}(t) + \varphi^0\boldsymbol{B}[1,\ldots,1] = \boldsymbol{B}\boldsymbol{\varphi}(t) \tag{3.11}$$

Pour corriger la turbulence, il suffit de connaître les pistons turbulents $\boldsymbol{\varphi}(t)$ à un piston global près. Il y a un piston turbulent à considérer par télescope, soit $N_t$ inconnues, à un piston global près, ce qui diminue de 1 le nombre d'informations sur la turbulence qu'il faudrait mesurer. On retrouve le fait que le nombre d'informations de phase manquantes est égal au rang de l'opérateur $\boldsymbol{B}$ (ou $\boldsymbol{B}^{\mathrm{T}}$, dans l'équation 3.9 page 77), c'est-à-dire $N_t - 1$.

## 3.3 Modèle de données et critère d'attache aux données exact

### 3.3.1 Critère d'attache aux données sur les visibilités carrées et les clôtures

A chaque instant, les fréquences spatiales couvertes sont les $\boldsymbol{\nu}(t)$, et on calcule $\boldsymbol{H}(t) = \boldsymbol{H}(\boldsymbol{\nu}(t))$. Les observables à chaque instant $t$ sont

– Les **visibilités carrées**, affectées d'un bruit $\boldsymbol{s}_b$ :

$$\begin{aligned}\boldsymbol{s}^{data}(t) &= \boldsymbol{s}(\boldsymbol{x},t) + \boldsymbol{s}_b(t)\\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}|^2 + \boldsymbol{s}_b(t).\end{aligned}$$

– Les **clôtures** de phase, affectées d'un bruit $\boldsymbol{\beta}_b$ :

$$\begin{aligned}\boldsymbol{\beta}^{data}(t) &= \boldsymbol{\beta}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \boldsymbol{\beta}_b(t)\end{aligned}$$

Les bruits $\boldsymbol{s}^b(t)$ et $\boldsymbol{\beta}_b(t)$ sont considérés gaussiens à moyenne nulle, d'écarts-type respectifs $\boldsymbol{\sigma}_{\boldsymbol{s}}(t)$ et $\boldsymbol{\sigma}_{\boldsymbol{\beta}}(t)$.
C'est l'hypothèse implicitement faite par les auteurs du format d'échange de données « OIFITS »[Pauls *et al.* (2004)] de l'International Astronomical Union (IAU). En l'absence d'autres données statistiques (intercorrélations, ou statistiques d'ordre supérieur), c'est la seule hypothèse acceptable.
Le critère d'attache aux données s'écrit comme un terme de $\chi^2$ :

$$\begin{aligned}\mathcal{J}^{data}(\boldsymbol{x},t) &= {\boldsymbol{s}_b}^{\mathrm{T}}(t)\boldsymbol{W}_s(t)\boldsymbol{s}_b(t) + {\boldsymbol{\beta}_b}^{\mathrm{T}}(t)\boldsymbol{W}_\beta(t)\boldsymbol{\beta}_b(t)\\ &= \left[\boldsymbol{s}_{\mathrm{data}}(t) - |H(t)\boldsymbol{x}|^2\right]^{\mathrm{T}}\boldsymbol{W}_s(t)\left[\boldsymbol{s}_{\mathrm{data}}(t) - |H(t)\boldsymbol{x}|^2\right] +\\ &\quad \left[\boldsymbol{\beta}_{\mathrm{data}}(t) - \boldsymbol{C}\arg(H(t)\boldsymbol{x})\right]^{\mathrm{T}}\boldsymbol{W}_\beta(t)\left[\boldsymbol{\beta}_{\mathrm{data}}(t) - \boldsymbol{C}\arg(H(t)\boldsymbol{x})\right]\\ &= \chi^2_s(\boldsymbol{x},t) + \chi^2_\beta(\boldsymbol{x},t)\end{aligned} \tag{3.12}$$

$\mathcal{J}^{data}(\boldsymbol{x})$ est la somme des termes $\mathcal{J}^{data}(\boldsymbol{x},t)$ formés pour chaque instant de mesure :

$$\mathcal{J}^{data}(\boldsymbol{x}) = \sum_t \mathcal{J}^{data}(\boldsymbol{x},t)$$

### 3.3.2 Terme d'amplitude de visibilités

Le premier terme du critère d'attache aux données de l'équation 3.12 peut être remplacé sans pertes par un terme en amplitude, si l'on prend soin de choisir un modèle de « pseudo-données » en amplitudes de visibilités compatibles avec le modèle de données en visibilités carrées, c'est-à-dire (cf.eq. 3.8) :

$$\bar{a} = \sqrt{\bar{s}}, \quad \sigma_a = \frac{\sigma_s}{2\bar{a}}$$

Lannes (1998a) suppose implicitement que les observables sont les clôtures de phase et les visibilités en module, puisque le passage de visibilités carrées à des visibilités en module n'est pas abordé.

Pour un tel modèle, les observables à chaque instant $t$ sont

– Les **amplitudes** de visibilités, affectées d'un bruit $\boldsymbol{a}_b$ :

$$\begin{aligned}\boldsymbol{a}^{data}(t) &= \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}_b(t)\\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}| + \boldsymbol{a}_b(t)\end{aligned}$$

– Les **clôtures** de phase, affectées d'un bruit $\boldsymbol{\beta}_b$ :

$$\begin{aligned}\boldsymbol{\beta}^{data}(t) &= \boldsymbol{\beta}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t)\\ &= \boldsymbol{C}\,\arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \boldsymbol{\beta}_b(t)\end{aligned}$$

Pour ce modèle, le critère d'attache aux données s'écrit :

$$\begin{aligned}\mathcal{J}^{data}(\boldsymbol{x},t) &= \boldsymbol{a}_b{}^{\mathrm{T}}(t)\boldsymbol{W}_a(t)\boldsymbol{a}_b(t) + \boldsymbol{\beta}_b{}^{\mathrm{T}}(t)\boldsymbol{W}_\beta(t)\boldsymbol{\beta}_b(t)\\ &= [\boldsymbol{a}_{\mathrm{data}}(t) - |H(t)\boldsymbol{x}|]^{\mathrm{T}}\boldsymbol{W}_a(t)\,[\boldsymbol{a}_{\mathrm{data}}(t) - |H(t)\boldsymbol{x}|] +\\ &\quad [\boldsymbol{\beta}_{\mathrm{data}}(t) - \boldsymbol{C}\arg(H(t)\boldsymbol{x})]^{\mathrm{T}}\,\boldsymbol{W}_\beta(t)\,[\boldsymbol{\beta}_{\mathrm{data}}(t) - \boldsymbol{C}\arg(H(t)\boldsymbol{x})]\\ &= \chi^2_a(\boldsymbol{x},t) + \chi^2_\beta(\boldsymbol{x},t)\end{aligned} \tag{3.13}$$

$\mathcal{J}^{data}(\boldsymbol{x})$ est la somme des termes $\mathcal{J}^{data}(\boldsymbol{x},t)$ formés pour chaque instant de mesure :

$$\mathcal{J}^{data}(\boldsymbol{x}) = \sum_t \mathcal{J}^{data}(\boldsymbol{x},t)$$

### 3.3.3 Dégénérescences du problème

Le terme en module du critère de l'équation 3.13 est non convexe, comme nous l'avons montré au chapitre 2 (de façon évidente, cette non convexité est vérifiée aussi par le terme en visibilités carrées du critère de l'équation 3.12).
Ces deux critères masquent le caractère sous-déterminé du problème en phase, sous-jacent dans le terme $\boldsymbol{C}\arg(H(t)\boldsymbol{x})$. L'incomplétude des données en phase entraîne l'apparition de minima locaux, qui peuvent être classées en deux groupes :

– Des minima locaux quelconques : deux objets $\boldsymbol{x}_1$ et $\boldsymbol{x}_2$ tels que $\arg(H(t)\boldsymbol{x}_1 - H(t)\boldsymbol{x}_2)$ est dans le noyau de $\boldsymbol{C}$ conduisent à un critère en phase de même valeur. Déterminer $\boldsymbol{x}$ implique un choix de la composante de $\arg(H(t)\boldsymbol{x})$ située dans le noyau de $\boldsymbol{C}$ ;
– Des minima locaux périodiques : il apparaît clairement que deux objets $\boldsymbol{x}_1$ et $\boldsymbol{x}_2$ tels que $\arg(H(t)\boldsymbol{x}_1) = \arg(H(t)\boldsymbol{x}_2) + 2k\pi$, avec $k$ entier, conduisent à un critère en phase de même valeur. Mais d'autres minima locaux périodiques sont induits par l'opérateur $\boldsymbol{C}$. Considérons une clôture à 3 télescopes, reliée à l'objet par

$$\beta(\boldsymbol{x}) = \phi_1(\boldsymbol{x}) + \phi_2(\boldsymbol{x}) + \phi_3(\boldsymbol{x})$$

Alors deux objets $\boldsymbol{x}_1$ et $\boldsymbol{x}_2$ tels que $\phi_i(\boldsymbol{x}_1) = \phi_i(\boldsymbol{x}_2) + \frac{2k\pi}{3}$, avec $k$ entier, conduisent eux aussi à un critère en phase de même valeur. Ces « ambiguïtés entières » ont été identifiées par Lannes (2001).

Nous avons cherché à établir un modèle direct des données équivalent au modèle vrai, mais faisant apparaître explicitement les inconnues supplémentaires à déterminer lors de la reconstruction, et correspondant à des éléments du noyau de l'opérateur de clôture $\boldsymbol{C}$.

## 3.4 Modélisation myope des données

Nous considérons dans cette section, que les observables à chaque instant $t$ sont
– les amplitudes de visibilités, affectées d'un bruit $\boldsymbol{a}_b$ :

$$\begin{aligned} \boldsymbol{a}^{data}(t) &= \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}_b(t) \\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}| + \boldsymbol{a}_b(t) \end{aligned} \tag{3.14}$$

– les **clôtures** de phase, affectées d'un bruit $\boldsymbol{\beta}_b$ :

$$\begin{aligned} \boldsymbol{\beta}^{data}(t) &= \boldsymbol{\beta}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t) \\ &= \boldsymbol{C}\,\boldsymbol{\phi}(\boldsymbol{x},t) + \boldsymbol{\beta}_b(t) \\ &= \boldsymbol{C}\,\arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \boldsymbol{\beta}_b(t) \end{aligned} \tag{3.15}$$

### 3.4.1 Informations manquantes et aberrations du système

Dans cette section, nous modifions l'équation 3.15 de façon à faire apparaître les informations manquantes. Nous présentons l'approche intuitive qui nous a conduits à l'équation 3.26, mais une interprétation algébrique simple, présentée section 3.4.1.4 permet de retrouver ce résultat en quelques lignes.

#### 3.4.1.1 Introduction d'informations supplémentaires virtuelles

Une façon alternative de décrire la formation des données est d'introduire les informations de phase manquantes dans le problème, en les assimilant à une connaissance imparfaite du système, ou "myopie".
L'équation d'observation est la suivante :

$$\begin{aligned} \boldsymbol{\beta}^{data} &= \boldsymbol{\beta}(\boldsymbol{x}) + \boldsymbol{\beta}^b \\ &= \boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{\beta}^b \end{aligned}$$

Nous avons cherché à introduire d'introduire des informations de phase supplémentaires $\boldsymbol{\alpha}$ de façon à exprimer les phases objet $\boldsymbol{\phi}(\boldsymbol{x})$ en fonction des mesures $\boldsymbol{\beta}^{data}$ et du bruit $\boldsymbol{\beta}^b$ :

$$\boldsymbol{\phi}(\boldsymbol{x}) = f(\boldsymbol{\alpha}, \boldsymbol{\beta}^{data}, \boldsymbol{\beta}^b)$$

Pour cela, il faudrait inverser l'opérateur $\boldsymbol{C}$ dans l'équation d'observation. Cependant, l'opérateur $\boldsymbol{C}$ n'est pas inversible, ce qui traduit le fait qu'il manque des informations de phase. Supposons qu'on ait mesuré (virtuellement) d'autres informations de phase $\boldsymbol{\alpha}$, complémentaires des clôtures, et définies par :

$$\boldsymbol{\alpha} = \boldsymbol{M}\boldsymbol{\phi}(\boldsymbol{x}) \tag{3.16}$$

On suppose ces mesures virtuelles parfaites, c'est-à-dire sans bruit. L'équation d'observation peut alors s'étendre en :

$$\left[\frac{\boldsymbol{\beta}^{data}}{\boldsymbol{\alpha}}\right] = \left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]\boldsymbol{\phi}(\boldsymbol{x}) + \left[\frac{\boldsymbol{C}}{\boldsymbol{0}}\right]\boldsymbol{\phi}^b \tag{3.17}$$

Si le nombre d'informations supplémentaires $\boldsymbol{\alpha}$ est exactement le nombre d'informations manquantes, et qu'elles sont indépendantes des clôtures, alors la matrice $\left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]$ est carrée et inversible.
Pour obtenir une expression de la forme

$$\boldsymbol{\phi}(\boldsymbol{x}) = f(\boldsymbol{\alpha}, \boldsymbol{\beta}^{data}, \boldsymbol{\beta}^b),$$

il est nécessaire de définir l'inverse de $\left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]$ :

$$\left[\begin{array}{c|c} \boldsymbol{M}_1 & \boldsymbol{M}_2 \end{array}\right] \triangleq \left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]^{-1} \tag{3.18}$$

On multiplie donc l'équation 3.17 par $\left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]^{-1}$ :

$$\left[\begin{array}{c|c} \boldsymbol{M}_1 & \boldsymbol{M}_2 \end{array}\right] \left[\frac{\boldsymbol{C}\boldsymbol{\phi}^{data}}{\boldsymbol{\alpha}}\right] = \boldsymbol{\phi}(\boldsymbol{x}) + \left[\begin{array}{c|c} \boldsymbol{M}_1 & \boldsymbol{M}_2 \end{array}\right] \left[\frac{\boldsymbol{C}}{\boldsymbol{0}}\right] \boldsymbol{\phi}^b$$
$$\boldsymbol{M}_1\boldsymbol{\beta}^{data} + \boldsymbol{M}_2\boldsymbol{\alpha} = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{M}_1\boldsymbol{\beta}^b$$

On obtient :

$$\boldsymbol{M}_1\boldsymbol{\beta}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) - \boldsymbol{M}_2\boldsymbol{\alpha} + \boldsymbol{M}_1\boldsymbol{\beta}^b$$

d'où :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{M}_1\boldsymbol{\beta}^{data} + \boldsymbol{M}_2\boldsymbol{\alpha} - \boldsymbol{M}_1\boldsymbol{\beta}^b \tag{3.19}$$

#### 3.4.1.2 Choix des informations supplémentaires à introduire

Il reste à définir quelles sont les informations supplémentaires $\boldsymbol{\alpha}$ à introduire, c'est-à-dire préciser les matrices $\boldsymbol{M}$, $\boldsymbol{M}_1$ et $\boldsymbol{M}_2$. Nous avons procédé par analogie avec les phases instantanées, affectées par la turbulence, que l'on mesurerait sur un interférogramme.
En l'absence de bruit, celles-ci s'écrivent (Eq. 1.17 page 29) :

$$\boldsymbol{\phi}^{data}(t) = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}(t),$$

soit avec un bruit de mesure

$$\boldsymbol{\phi}^{\mathrm{data}}(t) = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}(t) + \boldsymbol{\phi}_b(t)$$

Les phases objet vérifient donc :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}^{\mathrm{data}}(t) - \boldsymbol{B}\boldsymbol{\varphi}(t) - \boldsymbol{\phi}_b(t) \tag{3.20}$$

Nous avons vu (Eq. 3.11 page 79) que l'influence de la turbulence sur les phases est indépendante d'un piston global $\varphi^0$ :

$$\boldsymbol{B}(\boldsymbol{\varphi}(t) + \varphi^0[1,\ldots,1]^{\mathrm{T}}) = \boldsymbol{B}\boldsymbol{\varphi}(t) + \varphi^0\boldsymbol{B}[1,\ldots,1] = \boldsymbol{B}\boldsymbol{\varphi}(t)$$

Quitte à retrancher un piston global, on peut, sans perte de généralité, considérer que le piston sur le télescope 1 est nul, ce qui revient à fixer à 0 le premier élément $\varphi_1(t)$ du vecteur $\boldsymbol{\varphi}(t)$.
On définit donc $\bar{\boldsymbol{B}}$, obtenu en retirant la première colonne de $\boldsymbol{B}$. Cette colonne traduit l'influence du piston turbulent $\varphi_1$ sur le premier télescope. Pour $N_t = 4$, on a :

$$\bar{\boldsymbol{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

Comme on a (Eq. 1.18 page 30)

$$\boldsymbol{C}\boldsymbol{B} = \boldsymbol{0},$$

on conserve toujours le fait que

$$\boldsymbol{C}\bar{\boldsymbol{B}} = \boldsymbol{0}. \tag{3.21}$$

On a donc :

$$\boldsymbol{B}\boldsymbol{\varphi}(t) = \bar{\boldsymbol{B}}\boldsymbol{\varphi}(t)$$

et l'équation 3.20 devient :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}^{\text{data}}(t) - \bar{\boldsymbol{B}}\boldsymbol{\varphi}(t) - \boldsymbol{\phi}_b(t)$$

Le nombre d'informations manquantes est précisément le nombre de pistons turbulents sur chacun des $N_t$ télescopes, à un piston global près, soit $N_t - 1$. **Nous avons donc défini les $N_t - 1$ informations supplémentaires $\boldsymbol{\alpha}(t)$ comme l'analogue mathématique des pistons turbulents sur chacun des $N_t$ télescopes, à un piston global près.** La comparaison des équations 3.20 :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}^{\text{data}}(t) - \bar{\boldsymbol{B}}\boldsymbol{\varphi}(t) - \boldsymbol{\phi}_b(t)$$

et 3.19 :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{M}_1\boldsymbol{\beta}^{data} + \boldsymbol{M}_2\boldsymbol{\alpha} - \boldsymbol{M}_1\boldsymbol{\beta}^{b}$$

conduit donc à choisir :

$$\boxed{\boldsymbol{M}_2 = -\bar{\boldsymbol{B}}}$$

#### 3.4.1.3 Calcul de la matrice $\boldsymbol{M}_1$

L'équation 3.18 et le choix $\boldsymbol{M}_2 = -\bar{\boldsymbol{B}}$ imposent :

$$\begin{aligned}
\left[\begin{array}{c|c} \boldsymbol{M}_1 & \boldsymbol{M}_2 \end{array}\right] &= \left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]^{-1} \\
\Rightarrow \left[\frac{\boldsymbol{C}}{\boldsymbol{M}}\right]\left[\begin{array}{c|c} \boldsymbol{M}_1 & -\bar{\boldsymbol{B}} \end{array}\right] &= \left[\begin{array}{c|c} \mathbf{Id} & 0 \\ \hline 0 & \mathbf{Id} \end{array}\right] \\
\Rightarrow \left[\begin{array}{c|c} \boldsymbol{C}\boldsymbol{M}_1 & -\boldsymbol{C}\bar{\boldsymbol{B}} \\ \hline \boldsymbol{M}\boldsymbol{M}_1 & -\boldsymbol{M}\boldsymbol{B} \end{array}\right] &= \left[\begin{array}{c|c} \mathbf{Id} & 0 \\ \hline 0 & \mathbf{Id} \end{array}\right] \\
\Rightarrow \boldsymbol{C}\boldsymbol{M}_1 = \mathbf{Id}\,(i);\ \boldsymbol{C}\bar{\boldsymbol{B}} = \boldsymbol{0}\,(ii);\ \boldsymbol{M}\boldsymbol{M}_1 = \boldsymbol{0}\,(iii);\ -\boldsymbol{M}\bar{\boldsymbol{B}} = \mathbf{Id}\,(iv).
\end{aligned}$$

L'équation 3.21 assure déjà la condition (ii).
Les conditions (i) et (iv) impliquent :

$$\boxed{\boldsymbol{M}_1 = \boldsymbol{C}^\dagger, \boldsymbol{M} = -\boldsymbol{B}^\dagger}$$

avec :

$$\boldsymbol{C}^\dagger = \boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1} \tag{3.22}$$

$$\boldsymbol{B}^\dagger = [\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]^{-1}\boldsymbol{B}^{\mathrm{T}} \tag{3.23}$$

Lannes (1990) note par ailleurs que ces opérateurs s'écrivent simplement :

$$\boldsymbol{C}^\dagger = \boldsymbol{C}^{\mathrm{T}}/n \tag{3.24}$$

$$\boldsymbol{B}^\dagger = \boldsymbol{B}^{\mathrm{T}}/n \tag{3.25}$$

où $n$ est le nombre de télescopes. Cependant, l'auteur ne considère pas la même matrice $\boldsymbol{C}$ que nous, **la relation 3.24 n'est donc pas vérifiée.**

La condition (iii) est bien, elle aussi, vérifiée :

$$\begin{aligned}\boldsymbol{M}\boldsymbol{M}_1 &= -\boldsymbol{B}^\dagger\boldsymbol{C}^\dagger \\ &= -[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]^{-1}\boldsymbol{B}^{\mathrm{T}}\,\boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1} \\ &= -[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]^{-1}\,(\boldsymbol{C}\boldsymbol{B})^{\mathrm{T}}\,[\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}]^{-1}\end{aligned}$$

et l'équation 3.21 donne $\boldsymbol{C}\boldsymbol{B} = 0$.

L'équation 3.19

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{M}_1\boldsymbol{\beta}^{data} + \boldsymbol{M}_2\boldsymbol{\alpha} - \boldsymbol{M}_1\boldsymbol{\beta}^b$$

s'écrit donc :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{C}^\dagger\boldsymbol{\beta}^{data} - \bar{\boldsymbol{B}}\boldsymbol{\alpha} - \boldsymbol{C}^\dagger\boldsymbol{\beta}^b. \tag{3.26}$$

#### 3.4.1.4 Interprétation algébrique

L'équation 3.26 s'interprète assez simplement en terme de projections. On a un modèle de formation des données en phase de la forme (cf. eq. 3.15)

$$\boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\beta}^{data} - \boldsymbol{\beta}_b(t)$$

Toute solution $\boldsymbol{\phi}(\boldsymbol{x})$ de cette équation se décompose en une solution particulière $\boldsymbol{\phi}_0$ et un élément quelconque $\boldsymbol{\phi}_1$ du noyau de $\boldsymbol{C}$ :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1$$

Il est facile de voir que $\boldsymbol{\phi}_0 = \boldsymbol{C}^\dagger\boldsymbol{\beta}^{data} - \boldsymbol{C}^\dagger\boldsymbol{\beta}^b$ est une solution particulière de cette équation ; d'autre part, le noyau de $\boldsymbol{C}$ est exactement l'image de $\bar{\boldsymbol{B}}$ (cf eq. B.9. On peut donc paramétrer l'ensemble des $\boldsymbol{\phi}_1$ par :

$$\boldsymbol{\phi}_1 = \bar{\boldsymbol{B}}\boldsymbol{\alpha}$$

Avec ceci, on a donc :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1 = \boldsymbol{C}^\dagger\boldsymbol{\beta}^{data} - \boldsymbol{C}^\dagger\boldsymbol{\beta}^b + \bar{\boldsymbol{B}}\boldsymbol{\alpha}$$

### 3.4.2 Ambiguïtés entières

#### 3.4.2.1 Structure des ambiguïtés entières

Les ambiguïtés entières mentionnées section 3.3.3 correspondent à une indétermination de $2\pi$ sur les clôtures. Si l'on impose que les phases $\boldsymbol{\phi}(\boldsymbol{x})$ de la TF de l'objet sont définies entre $-\pi$ et $\pi$, alors chaque clôture à 3 télescopes a une valeur entre $-3\pi$ et $3\pi$. Or ces clôtures sont mesurées ***modulo*** $2\pi$, c'est-à-dire entre $-\pi$ et $\pi$. Le modèle de formation des données en phase de l'équation 3.15 est modifié en :

$$\boldsymbol{C}\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\beta}^{data} - \boldsymbol{\beta}_b(t) + 2\pi\boldsymbol{\kappa},$$

où $\boldsymbol{\kappa}$ est un vecteur dont chaque coefficient a pour valeur 0, 1 ou -1.
On note que si toutes les phases $\boldsymbol{\phi}(\boldsymbol{x})$ de la TF de l'objet sont comprises entre $-\pi/3$ et $\pi/3$, alors chaque clôture à 3 télescopes a une valeur entre $-\pi$ et $\pi$, et $\boldsymbol{\kappa}$ est alors le vecteur nul :

$$|\boldsymbol{\phi}_i(\boldsymbol{x})| < \pi/3,\ \forall i \Rightarrow \boldsymbol{\kappa} = \boldsymbol{0}. \tag{3.27}$$

#### 3.4.2.2 Influence sur le modèle en phases

En reprenant l'approche de la section 3.4.1.4, on écrit que toute solution $\boldsymbol{\phi}(\boldsymbol{x})$ de cette équation se décompose en une solution particulière $\boldsymbol{\phi}_0(\kappa)$ et un élément quelconque $\boldsymbol{\phi}_1$ du noyau de $\boldsymbol{C}$. On peut toujours écrire $\boldsymbol{\phi}_1 = \bar{\boldsymbol{B}}\boldsymbol{\alpha}$. Par contre, une solution particulière s'écrit :

$$\boldsymbol{\phi}_0(\kappa) = \boldsymbol{\phi}_0(0) + 2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa} = \boldsymbol{C}^\dagger\boldsymbol{\beta}^{data} - \boldsymbol{C}^\dagger\boldsymbol{\beta}^b + 2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa}$$

Avec ceci, on a donc :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}_0(0) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} + 2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa}$$

On peut interpréter le terme $\bar{\boldsymbol{B}}\boldsymbol{\alpha} + 2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa}$ comme un paramétrage du noyau de l'opérateur $\boldsymbol{C} \mod 2\pi$

On ne s'intéresse aux phases de la TF de l'objet $\boldsymbol{\phi}(\boldsymbol{x})$ que modulo $2\pi$. Cependant, les coefficients du vecteur $\boldsymbol{C}^\dagger\boldsymbol{\kappa}$ ne sont pas nécessairement entiers. A titre d'exemple, pour une seule clôture, sur un interféromètre à 3 télescopes, on a :

$$2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa} = 2\pi\boldsymbol{\kappa}/3$$

Le terme $2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa}$ ne s'annule donc pas, dans le cas général, même lorsqu'on considère des phases objet modulo $2\pi$.

#### 3.4.2.3 L'hypothèse $\boldsymbol{\kappa} = \mathbf{0}$

Le traitement du problème à $\boldsymbol{\kappa}$ quelconque est un point dur, car ce vecteur prend par exemple $3^{N_c}$ valeurs possibles, où $N_C$ est le nombre total de mesures de clôtures du problème. Cependant, nous avons vu (eq. 3.27) que si toutes les phases $\boldsymbol{\phi}(\boldsymbol{x})$ de la TF de l'objet sont comprises entre $-\pi/3$ et $\pi/3$, alors $\boldsymbol{\kappa}$ est le vecteur nul.
C'est le cas pour un objet non résolu, pour lequel toutes les phases de la TF sont nulles, c'est-à-dire pour les composantes basse fréquence d'un objet quelconque. L'hypothèse $\boldsymbol{\kappa} = \mathbf{0}$ pénalise donc moins les basses fréquences.

Cette remarque n'est valable que pour des **objets centrés**. En effet, comme nous l'avons déjà mentionné (section 3.3.3), une translation de l'objet induit l'addition d'un plan de phase sur les phases de la TF. Pour un objet non résolu mais non centré, les phases basse fréquence ne sont pas nulles, qui plus est elles ne sont pas comprises entre $\pi/3$ et $-\pi/3$.

Nous sommes conscients qu'attribuer la valeur 0 aux coefficients de $\boldsymbol{\kappa}$ n'est pas satisfaisant. Cependant, l'étude précise des ambiguïtés entières n'est pas triviale. Nous avons fait le choix de la laisser de côté, et de privilégier l'étude du problème à $\boldsymbol{\kappa} = \mathbf{0}$. Les simulations proposées au section 5.1 page 115 laissent penser que l'hypothèse $\boldsymbol{\kappa} = \mathbf{0}$ n'est pas préjudiciable à la reconstruction. Le traitement propre des ambiguïtés entières constitue une suite envisagée à ces travaux.

#### 3.4.2.4 Justification de l'hypothèse $\boldsymbol{\kappa} = 0$

Cette partie a été ajoutée après envoi aux rapporteurs, et correspond à des travaux menés entre la redaction du manuscrit et la soutenance de thèse.
L'equation d'inversion s'écrit :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}_0(0) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} + 2\pi\boldsymbol{C}^\dagger\boldsymbol{\kappa}$$

Or, $\boldsymbol{C}$ vérifie(Eq. B.8 page 160) :

$$\boldsymbol{C} = \left[ \; * \; \middle| \; \mathbf{Id} \; \right]$$

On pose $\boldsymbol{\mu} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix}$. $\boldsymbol{\mu}$ est un vecteur d'entiers, qui vérifie :

$$\boldsymbol{C\mu} = \left[ \; * \; \middle| \; \mathbf{Id} \; \right] \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\kappa} \end{bmatrix} = \boldsymbol{\kappa}$$

On a donc :

$$\begin{aligned} \boldsymbol{C\mu} = \boldsymbol{\kappa} &\Rightarrow \boldsymbol{C\mu} = \boldsymbol{CC^\dagger\kappa} \\ &\Rightarrow \boldsymbol{C}(\boldsymbol{C^\dagger\kappa} - \boldsymbol{\mu}) = 0 \\ &\Rightarrow \boldsymbol{C^\dagger\kappa} - \boldsymbol{\mu} = \bar{\boldsymbol{B}}\boldsymbol{\alpha}' \\ &\Rightarrow \boldsymbol{C^\dagger\kappa} = \bar{\boldsymbol{B}}\boldsymbol{\alpha}' + \boldsymbol{\mu} \end{aligned}$$

L'équation d'inversion s'écrit donc :

$$\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}_0(0) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} + 2\pi\boldsymbol{C^\dagger\kappa} = \boldsymbol{\phi}_0(0) + \bar{\boldsymbol{B}}(\boldsymbol{\alpha} + \boldsymbol{\alpha}') + 2\pi\boldsymbol{\mu}$$

La phase objet $\boldsymbol{\phi}(\boldsymbol{x})$ n'a de sens que modulo $2\pi$. Le terme $2\pi\boldsymbol{\mu}$ n'a donc pas d'effet.
Nous avons montré que toute phase $\boldsymbol{\phi}(\boldsymbol{x})$ solution se met sous la forme $\boldsymbol{\phi}_0(0) + \bar{\boldsymbol{B}}(\boldsymbol{\alpha})$ à $2\pi$ près. Le fait de prendre $\boldsymbol{\kappa} = \mathbf{0}$ ne restreint pas l'ensemble des phases objet solutions.

### 3.4.3 Construction du modèle myope

L'équation 3.17 peut donc se réécrire :

$$\boldsymbol{C^\dagger\beta}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} + \boldsymbol{C^\dagger\beta}^{b}$$

Les clôtures mesurées $\boldsymbol{\beta}^{data}$ ne sont pas obtenues en faisant la clôture des phases mesurées :

$$\boldsymbol{\beta}^{data} \neq \boldsymbol{C\phi}^{data}$$

En effet, une clôture mesurée est la phase de la *moyenne temporelle du bispectre* (Cf. Section 1.2.3 page 26) :

$$Cl_{ijk}^{data} = \arg \left\langle a_i^{data} a_j^{data} a_k^{data} \mathrm{e}^{\mathrm{i}\boldsymbol{C\phi}^{data}} \right\rangle_t$$

Cependant, si l'on faisait une mesure sans bruit, *i. e.* stable au cours du temps, la clôture mesurée serait bien égale à la clôture des phases mesurées. Dans ce cas, on aurait :

$$\boldsymbol{C^\dagger\beta}^{data} \equiv \boldsymbol{C^\dagger C\phi}^{data}$$

L'opérateur $\boldsymbol{C^\dagger C}$ est un projecteur :

$$\boldsymbol{C^\dagger CC^\dagger C} = \boldsymbol{C^\dagger C}$$

Donc $\boldsymbol{C^\dagger\beta}^{data}$ est analogue à une mesure de phase, d'où

$$\boldsymbol{C^\dagger\beta}^{data} \equiv \boldsymbol{C^\dagger C\phi}^{data} \equiv \boldsymbol{\phi}^{data}$$

Lorsqu'on recombine les mesures de clôtures $\boldsymbol{\beta}^{data}$, affectées par un bruit $\boldsymbol{\beta}_b$, en calculant des pseudo-mesures en phase

$$\boldsymbol{\phi}^{data} \triangleq \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^{data},$$

ces dernières sont affectées par un pseudo-bruit en phase $\boldsymbol{\phi}_b$ :

$$\boldsymbol{\phi}_b = \boldsymbol{C}^{\dagger}\boldsymbol{\beta}_b \tag{3.28}$$

Si $\boldsymbol{R}_\beta$ est la matrice de covariance du bruit de clôtures, la matrice de covariance du pseudo bruit de phase $\boldsymbol{\phi}^b = \boldsymbol{C}^{\dagger}\boldsymbol{\beta}^b$ est donnée par :

$$\boxed{\boldsymbol{R}_\phi = \boldsymbol{C}^{\dagger}\boldsymbol{R}_\beta\boldsymbol{C}^{\dagger T}} \tag{3.29}$$

Le modèle myope des données obtenu à chaque instant $t$ de mesure est donc de la forme :

– Pour les amplitudes de visibilités :

$$\begin{aligned}\boldsymbol{a}^{data}(t) &= \boldsymbol{a}(\boldsymbol{x}) + \boldsymbol{a}_b(t)\\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}| + \boldsymbol{a}_b(t)\end{aligned}$$

– Pour les pseudo phases de visibilités :

$$\begin{aligned}\boldsymbol{\phi}^{\mathrm{data}}(t) &= \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}_b(t)\\ &= \arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}_b(t)\end{aligned}$$

où $\boldsymbol{\phi}_b(t)$ est un bruit gaussien de matrice de covariance donnée par l'équation 3.29.
Les dégénérescences du problème évoquées section 3.3.3 figurent à présent explicitement sous la forme des $\boldsymbol{\alpha}$.

### 3.4.4 Application à une couverture fréquentielle redondante avec un interféromètre à 3 télescopes

Nous avons cherché à déterminer quelles sont les informations virtuelles supplémentaires à introduire lorsque l'on utilise un interféromètre à 3 télescopes en faisant des mesures redondantes (Cf. Section 1.2.4.1 page 27).

Supposons que l'on dispose uniquement d'un IO3T, et qu'on l'utilise à 3 instants différents, dans des configurations correspondant aux positions $123$ puis $124$ puis $134$ d'un IO4T. Le nombre de phases objet à mesurer est toujours 6 : $\mathrm{card}\,\boldsymbol{\phi}(\boldsymbol{x}) = 6$.

Les mesures courte pose des phases fourniraient :

– Pour l'instant 1, en configuration $123$,

$$\begin{cases}\phi_{12}^{data}(1) = \phi(\boldsymbol{x})_{12} + \varphi_2(1) - \varphi_1^{turb,1}(1)\\ \phi_{13}^{data}(1) = \phi(\boldsymbol{x})_{13} + \varphi_3(1) - \varphi_1(1)\\ \phi_{23}^{data}(1) = \phi(\boldsymbol{x})_{23} + \varphi_3(1) - \varphi_2(1)\end{cases} \tag{3.30}$$

– Pour l'instant 2, en configuration $124$,

$$\begin{cases}\phi_{12}^{data}(2) = \phi(\boldsymbol{x})_{12} + \varphi_2(2) - \varphi_1(2)\\ \phi_{14}^{data}(2) = \phi(\boldsymbol{x})_{14} + \varphi_4(2) - \varphi_1(2)\\ \phi_{24}^{data}(2) = \phi(\boldsymbol{x})_{24} + \varphi_4(2) - \varphi_2(2)\end{cases} \tag{3.31}$$

– Pour l'instant 3, en configuration $134$,

$$\begin{cases} \phi_{13}^{data}(3) = \phi(\boldsymbol{x})_{13} + \varphi_3(3) - \varphi_1(3) \\ \phi_{14}^{data}(3) = \phi(\boldsymbol{x})_{14} + \varphi_4(3) - \varphi_1(3) \\ \phi_{34}^{data}(3) = \phi(\boldsymbol{x})_{34} + \varphi_4(3) - \varphi_3(3) \end{cases} \tag{3.32}$$

On pose :

$$\boldsymbol{\phi}^{data} = \begin{bmatrix} \phi_{12}^{data}(1) \\ \phi_{13}^{data}(1) \\ \phi_{23}^{data}(1) \\ \phi_{12}^{data}(2) \\ \phi_{14}^{data}(2) \\ \phi_{24}^{data}(2) \\ \phi_{13}^{data}(3) \\ \phi_{14}^{data}(3) \\ \phi_{34}^{data}(3) \end{bmatrix}, \ \boldsymbol{\varphi} = \begin{bmatrix} \varphi_1(1) \\ \varphi_2(1) \\ \varphi_3(1) \\ \varphi_1(2) \\ \varphi_2(2) \\ \varphi_4(2) \\ \varphi_1(3) \\ \varphi_3(3) \\ \varphi_4(3) \end{bmatrix}$$

De cette façon, on peut reformer une équation de la forme

$$\boldsymbol{\phi}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}$$

avec

$$\boldsymbol{B} = \left[\begin{array}{c|c|c} \boldsymbol{B}_i & 0 & 0 \\ \hline 0 & \boldsymbol{B}_i & 0 \\ \hline 0 & 0 & \boldsymbol{B}_i \end{array}\right]$$

et

$$\boldsymbol{B}_i = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

Cette modélisation conduirait à introduire 6 aberrations myopes, soit 2 par instant. Cependant, les clôtures formées sont :

$$\boldsymbol{C}\phi(\boldsymbol{x}) = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 1 \end{bmatrix}$$

Le nombre d'informations virtuelles à introduire est donc :

$$\mathrm{card}(\phi(x)) - \mathrm{card}(\boldsymbol{\beta}^{\mathrm{data}}) = 6 - 3 = 3$$

On n'introduit que 3 aberrations myopes $\alpha_2$, $\alpha_3$, et $\alpha_4$. La modélisation myope

$$\boldsymbol{\phi}^{data} = \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}'\boldsymbol{\alpha}$$

avec

$$\boldsymbol{B}' = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

et

$$\bar{\boldsymbol{B}}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

est compatible avec les données de clôture.
Le nombre d'aberrations myopes à introduire n'est **pas** dans ce cas le rang de $\boldsymbol{B}$ ou de $\boldsymbol{B}'$, mais seulement le nombre d'informations manquantes, soit

$$\mathrm{card}(\phi(x)) - \mathrm{card}(\boldsymbol{\beta}^{\mathrm{data}}) = 6 - 3 = 3$$

Il suffit donc d'introduire 3 informations manquantes, alors que si l'on traitait le problème comme trois triplets de télescopes indépendants, on introduirait 6 informations manquantes. Le problème est donc « moins myope » qu'il n'y parait. On verra au chapitre 4 (section 4.5.3 page 112) qu'il est effectivement bien moins difficile de reconstruire un objet avec un interféromètre à 6 télescopes (ou avec les données équivalentes fournies par un interféromètre à 3 télescopes avec une couverture fréquentielle redondante équivalente) qu'avec un interféromètre à 3 télescopes.
On remarque de plus que **la situation est exactement équivalente à celle que l'on aurait avec un interféromètre à 4 télescopes**. On peut donc avec cette technique obtenir avec un interféromètre à 3 télescopes les données que fournirait un interféromètre à n télescopes, pour tout n.

### 3.4.5 Un autre traitement des données IRCOM

Il est possible d'utiliser cette modélisation myope « minimale » à la configuration de l'interféromètre fibré de l'IRCOM. En effet, ce dispositif est un exemple de redondance fréquentielle avec un interféromètre à 3 télescopes. Cependant, nous n'avons pas eu le temps suffisant pour retraiter ces données. Nous ne faisons donc que décrire dans cette section la modélisation myope que nous aurions utilisée.
A l'instant $n$, on mesure la clôture de phase

$$\begin{cases} \beta(1) = 2\phi(\boldsymbol{\nu}_{12}) + \phi(2\boldsymbol{\nu}_{12}) \\ \beta(2) = \phi(\boldsymbol{\nu}_{12}) + \phi(2\boldsymbol{\nu}_{12}) + \phi(3\boldsymbol{\nu}_{12}) \\ \quad \vdots \\ \beta(n) = \phi(\boldsymbol{\nu}_{12}) + \phi(n\boldsymbol{\nu}_{12}) + \phi((n+1)\boldsymbol{\nu}_{12}). \end{cases}$$

On mesure donc $n$ clôtures pour $n+1$ phases objet recherchées. Il suffit donc d'introduire une seule aberration myope par direction de mesure. L'inclusion de cette inconnue est une alternative à la propagation de l'erreur en phase le long de la direction de mesure, évoquée section 2.6.2 page 67.

## 3.5 Bilan : Un modèle myope de données sur des phases et des amplitudes

A ce stade, nous avons donc

- construit des pseudo mesures de phases $\boldsymbol{\phi}^{data}$, construites à partir des mesures de clôtures $\boldsymbol{\beta}^{data}$, en introduisant les informations de phase manquantes sous forme d'aberrations myopes du système $\boldsymbol{\alpha}$ (Section 3.4 page 82) ;
- construit des pseudo mesures de modules de visibilité $\boldsymbol{a}^{data}$, construites à partir des mesures de visibilité carrée $\boldsymbol{s}^{data}$ (Section 3.1 page 74) ;
- proposé une loi de probabilité des bruits sur ces pseudo mesures compatibles avec les statistiques disponibles.

Le modèle myope des données obtenu à chaque instant $t$ de mesure est donc de la forme :

- Pour les pseudo mesures d'amplitudes de visibilités :

$$\begin{aligned} \boldsymbol{a}^{data}(t) &= \boldsymbol{a}(\boldsymbol{x}) + \boldsymbol{a}_b(t) \\ &= |\boldsymbol{H}(t)\,\boldsymbol{x}| + \boldsymbol{a}_b(t) \end{aligned} \tag{3.33}$$

- Pour les pseudo mesures de phases de visibilités :

$$\begin{aligned} \boldsymbol{\phi}^{\mathrm{data}}(t) &= \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}_b(t) \\ &= \arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}_b(t) \end{aligned} \tag{3.34}$$

On considère généralement que les bruits en amplitude carrées $\boldsymbol{s}_b(t)$ et en clôture de phase $\boldsymbol{\beta}_b(t)$ sont blancs, indépendants et gaussiens à moyenne nulle, d'écart-types respectifs $\boldsymbol{\sigma}_s(t)$ et $\boldsymbol{\sigma}_\beta(t)$.

$$\begin{aligned} \bar{\boldsymbol{s}}_b &= \mathbf{0} \\ \bar{\boldsymbol{\beta}}_b &= \mathbf{0} \\ \boldsymbol{R}_s(t) &= \boldsymbol{W}_s^{-1}(t) = \mathrm{Diag}\left\{\boldsymbol{\sigma}_s^2(t)\right\} \\ \boldsymbol{R}_\beta(t) &= \boldsymbol{W}_\beta^{-1}(t) = \mathrm{Diag}\left\{\boldsymbol{\sigma}_\beta^2(t)\right\} \end{aligned}$$

Les lois de probabilité des bruits sur les pseudo mesures des équations 3.33 et 3.34 sont construites de la façon suivante :

- Les bruits en amplitude $\boldsymbol{a}^b(t)$ et en phase $\boldsymbol{\phi}^b(t)$ sont considérés gaussiens.
- La moyenne des bruits en phase vaut (3.28 page 88) :

$$\bar{\boldsymbol{\phi}}_b = \boldsymbol{C}^\dagger \bar{\boldsymbol{\beta}}_b = \mathbf{0};$$

- La matrice de covariance des bruits en phase $\boldsymbol{R}_\phi(t) = \boldsymbol{W}_\phi^{-1}(t)$ se déduit de la matrice de covariance des bruits en clôture $\boldsymbol{R}_\beta(t)$ selon l'équation 3.29 page 88 ;
- La moyenne des bruits en amplitude vaut (3.6 page 76) :

$$\bar{\boldsymbol{a}}_b = \sqrt{\bar{\boldsymbol{s}}_b} = \mathbf{0};$$

- La matrice de covariance des bruits en amplitude $\boldsymbol{R}_a(t) = \boldsymbol{W}_a^{-1}(t)$ se déduit des écarts-type en visibilités carrées $\boldsymbol{\sigma}_s(t)$ selon l'équation 3.8 page 76.

# Chapitre 4

# Reconstruire des images à partir de visibilités carrées et de clôtures de phase

La reconstruction d'images en interférométrie optique butte sur au moins trois obstacles :

1. comme nous l'avons déjà mentionné dans le cas non-turbulent (cf. section 1.3.1), le faible nombre de données nécessite l'introduction d'informations *a priori* sur l'objet à reconstruire. Nous avons proposé au chapitre 2 la méthode de reconstruction RIME, validée sur simulations, et testée avec succès sur données expérimentales ;
2. s'y ajoute la sous-détermination en phase, étudiée au chapitre précédent ;
3. même à aberrations fixées, le critère d'attache aux données est non convexe par rapport à l'objet. Nous avons proposé au chapitre 2, dans le cadre de RIME, une approximation quadratique de ce critère.

Nous avons déjà partiellement répondu aux points 1 et 3. En revanche, la prise en compte du manque d'information de phase est l'enjeu principal de la reconstruction d'image en interférométrie optique en présence de turbulence.
En effet, **l'information en phase est cruciale** en synthèse de Fourier, car elle contient la majorité des structures de l'objet, et de fait, des méthodes de reconstruction à partir de la phase uniquement ont prouvé leur efficacité (Voir Oppenheim *et al.* (1980) et Oppenheim & Lim (1981)). Cependant, il existe des problèmes pour lesquels l'information en phase est partielle ou même absente, comme en cristallographie [Navaza (1985)]. L'efficacité de méthodes exploitant uniquement que les informations en module reste à démontrer en interférométrie optique. Ceci nous a conduit à prendre en compte le manque d'information en phase du à la turbulence par le biais de l'estimation conjointe de l'objet et d'aberrations en phase. Suivant cette approche, les principaux travaux concernent des données radio-interférométriques, qui diffèrent du cadre optique par leur nombre, la structure du bruit (cf. section 2.3.1) et l'existence d'*a priori* sur les aberrations de phase. Ces méthodes, dites **d'auto-calibration** (section 4.1.2 à suivre), restent une source d'inspiration privilégiée en optique, comme en témoignent les travaux de Lannes (1998a). A ce titre, l'algorithme MIRA, développé par Thiébaut *et al.* (2003) fait figure d'exception (section 4.1).

De façon générale, toutes ces méthodes s'appuient sur la minimisation d'un critère qui, étant donné le manque d'information de phase, présente de nombreux minima locaux. Une recherche globale du minimum s'impose logiquement, bien qu'à notre connaissance l'utilisation d'une méthode d'optimisation globale n'ait pas été envisagée, vraisemblablement pour des raisons de temps de calcul.

Si l'on se restreint à une optimisation **locale** du problème, elle est heuristique, et vise à parcourir un « bon » chemin (qui dépend, entre autres, des *a priori* introduits et de la méthode d'optimisation), à partir d'un « bon » point de départ (l'initialisation), pour espérer arriver à un « bon » minimum. **L'ensemble des méthodes de reconstruction d'image en interférométrie, y compris l'algorithme WISARD présenté dans ce chapitre, ne diffèrent que par le chemin suivi et l'initialisation utilisée.**
Comme il n'est pas possible de démontrer théoriquement que l'on atteint dans tous les cas le minimum global, seule une validation *a posteriori* sur simulations permet de mesurer l'efficacité de ces méthodes.

La section 4.1 de ce chapitre présente rapidement les techniques existantes, en identifiant les manques que nous avons tenté de combler. Nous établissons ensuite un modèle de formation de données de visibilités complexes affectées par un bruit complexe, et dépendant des aberrations $\boldsymbol{\alpha}$, de façon à en déduire un critère d'attache aux données dépendant de l'objet et des aberrations $\mathcal{J}_{\text{data}}(\boldsymbol{x}, \boldsymbol{\alpha})$ (Section 4.2).
Puis nous proposons (Section 4.4) l'algorithme WISARD (Weak-phase Interferometric Sample Alternating Reconstruction Device), qui s'appuie notamment sur une méthode originale de détermination des aberrations (section 4.3).

## 4.1 Méthodes de reconstruction d'image en interférométrie à clôtures

### 4.1.1 Inférence bayésienne classique

Dans le cadre d'une inférence bayésienne classique (cf. section 1.4 page 35), on est amené à minimiser un critère composite constitué d'un terme d'*a priori* $\mathcal{J}_{\text{prior}}$ et du critère d'attache aux données déduit directement du modèle de données en visibilités carrées et clôture de phase :

$$\begin{aligned}\mathcal{J}^{data}(\boldsymbol{x}, t) &= \boldsymbol{s}_b{}^{\mathrm{T}}(t)\boldsymbol{W}_s(t)\boldsymbol{s}_b(t) + \boldsymbol{\beta}_b{}^{\mathrm{T}}(t)\boldsymbol{W}_\beta(t)\boldsymbol{\beta}_b(t) \\ &= \chi^2_s(\boldsymbol{x}, t) + \chi^2_\beta(\boldsymbol{x}, t)\end{aligned}$$

C'est le critère d'attache aux données exact décrit section 3.3.1 (eq. 3.12 page 80).

En utilisant cette méthode, on espère que le terme d'*a priori* seul permet de régler les trois points durs évoqués en introduction de ce chapitre :
- lever l'indétermination en l'objet due à la sous-détermination du problème à phases connues ;
- lever les indéterminations en phase ;
- rendre le critère composite « suffisamment convexe » pour qu'une optimisation locale soit possible

C'est l'approche choisie par Éric Thiébaut [Thiébaut *et al.* (2003)], qui propose l'algorithme MIRA(Multi-aperture Image Reconstruction Algorithm)[1]. L'*a priori* utilisé correspond à la régularisation quadratique à pixels indépendants décrite section 1.4.2.2 page 39, alliée à une contrainte de positivité. L'approche bayésienne utilisée dans MIRA conduit à un algorithme modulaire, pour lequel il est aisé de tester différents types de régularisation.

[1] Dans cet algorithme, le terme $\chi^2_\beta(\boldsymbol{x}, t)$ est remplacé par un $\chi^2$ sur des phaseurs, c'est-à-dire sur des termes $\mathrm{e}^{\mathrm{i}\beta}$, pour eviter des problèmes d'enroulement de phase.

La seule critique que l'on peut formuler à l'égard de la base théorique de MIRA est que le critère d'attache aux données exact, que cet algorithme minimise, masque les structures en phase du problème.

### 4.1.2 Méthodes d'auto-calibration en radio-astronomie

En radio astronomie, les visibilités complexes recueillies sur des radio-interféromètres « instables » sont affectées par des erreurs instrumentales en phase et en module. Ce contexte est suffisamment proche de l'interférométrie optique en présence de turbulence pour que nous puissions nous inspirer des méthodes dites d'« auto-calibration », décrites par Cornwell et Wilkinson[Cornwell & Wilkinson (1981)].

#### 4.1.2.1 Modèle de données

Cornwell et Wilkinson proposent le modèle de données suivant :

– Les visibilités complexes mesurées $y_{j,k}^{\mathrm{data}}$ sont affectées par des erreurs en module $\rho$, des erreurs en phase $\varphi$, et un bruit additif $y^{\mathrm{b}}$ :

$$y_{j,k}^{\mathrm{data}} = y_{j,k}(\boldsymbol{x}, \boldsymbol{\rho}, \boldsymbol{\varphi}) + y_{j,k}^{\mathrm{b}} = y_{j,k}(\boldsymbol{x})(1+\rho_j)(1+\rho_k)e^{i(\varphi_k - \varphi_j)} + y_{j,k}^{\mathrm{b}}$$

L'hypothèse fondamentale dans ce modèle est que les erreurs en phase et en module sont « attribuables » à une antenne (ou un télescope). En effet, dans le cas de radio-interféromètres instables, chaque antenne est à la source d'erreurs en phase et modules.

– Ces trois sources d'erreur sont supposées décorrelées et gaussiennes :
  – Bruit additif : $p(y_{j,k}^{\mathrm{b}}) \propto \exp\left(-\frac{1}{2}\frac{|y_{j,k}^{\mathrm{b}}|^2}{\sigma_{j,k}^2}\right)$. C'est un modèle de bruit « circulaire » ;
  – « Aberrations de module » :$\rho_j \sim \mathcal{N}\left(\rho_{j_m}, \sigma_{\rho_j}\right)$
  – « Aberrations de phase » :$\varphi_j \sim \mathcal{N}\left(\varphi_{j_m}, \sigma_{\varphi_j}\right)$

La modélisation de Cornwell et Wilkinson propose une distribution gaussienne des aberrations en phase et en module. C'est l'une des principales différences avec l'interférométrie optique, où l'on ne considère pas d'aberrations en module, et, surtout, où l'on ne dispose d'aucun *a priori* sur les aberrations en phase.

#### 4.1.2.2 Une minimisation alternée

Le principe de la méthode d'auto-calibration proposée par Cornwell et Wilkinson consiste à alterner des étapes d'estimation de l'objet – à aberrations fixées – avec des étapes d'estimation des aberrations – à objet fixé (figure 4.1 page 96).

L'optimisation de l'objet à aberrations connues est réalisée à l'aide de l'algorithme CLEAN. La détermination des aberrations à objet fixé fait l'objet du paragraphe suivant.

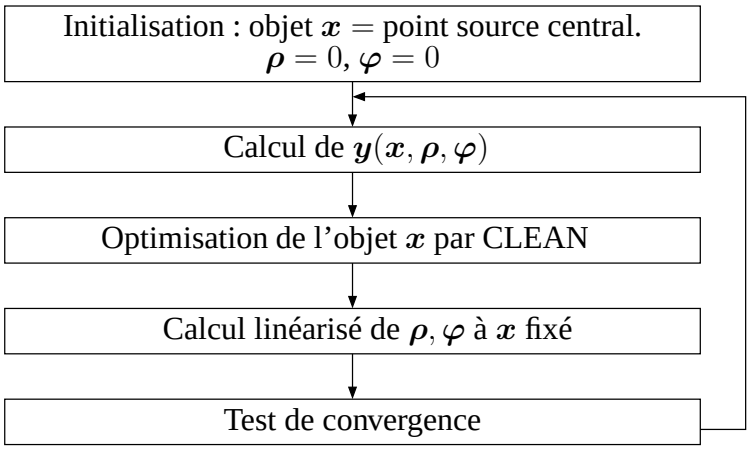



FIG. 4.1 – Boucle d'auto-calibration

#### 4.1.2.3 Détermination des aberrations à objet fixé

Les auteurs proposent de déterminer les aberrations à objet $\boldsymbol{x}$ fixé en maximisant la probabilité jointe du bruit et des aberrations :

$$\begin{aligned}
\mathcal{J}(\boldsymbol{x}, \boldsymbol{\rho}, \boldsymbol{\varphi}) &= -2 \log p(\boldsymbol{\rho}, \boldsymbol{\varphi}, \boldsymbol{y}^{\mathrm{b}}) \\
&= \sum_{j,k} \frac{|y^{\mathrm{b}}_{j,k}(\boldsymbol{x})|^2}{\sigma^2_{j,k}} + \sum_j \frac{\rho_j^2}{\sigma^2_{\rho_j}} + \sum_j \frac{\varphi_j^2}{\sigma^2_{\varphi_j}} \\
&= \underbrace{\sum_{j,k} \frac{1}{\sigma^2_{j,k}} |y^{\mathrm{data}}_{j,k} - y_{j,k}(\boldsymbol{x})(1+\rho_j)(1+\rho_k) e^{i(\varphi_k - \varphi_j)}|^2}_{\mathcal{J}_{data}(\boldsymbol{x},\boldsymbol{\rho},\boldsymbol{\varphi})} \\
&\quad + \underbrace{\sum_j \frac{\rho_j^2}{\sigma^2_{\rho_j}}}_{\mathcal{J}_{Prior}(\boldsymbol{\rho})} + \underbrace{\sum_j \frac{\varphi_j^2}{\sigma^2_{\varphi_j}}}_{\mathcal{J}_{Prior}(\boldsymbol{\varphi})}
\end{aligned}$$

Ce critère n'est pas à proprement parler de type maximum *a posteriori*, car il n'y a aucun *a priori* sur l'objet, mais une vraisemblance généralisée.

La détermination des aberrations est réalisée séparément, c'est-à-dire
- détermination de $\boldsymbol{\rho}$ à $\boldsymbol{x}$ et $\boldsymbol{\varphi}$ connus ;
- détermination de $\boldsymbol{\varphi}$ à $\boldsymbol{x}$ et $\boldsymbol{\rho}$ connus.

Nous ne présentons ici que cette deuxième étape, la seule d'intérêt en interférométrie optique.

Quitte à considérer $\boldsymbol{x}$ et $\boldsymbol{\rho}$ connus, on peut remplacer $y_{j,k}(\boldsymbol{x})$ par $y_{j,k}(\boldsymbol{x})(1+\rho_j)(1+\rho_k)$ et obtenir sans perte de généralité une expression du critère à minimiser indépendante des aberrations en amplitude :

$$\mathcal{J}(\boldsymbol{\varphi}) = \sum_{j,k} \frac{1}{\sigma^2_{j,k}} |y^{\mathrm{data}}_{j,k} - y_{j,k} e^{i(\varphi_k - \varphi_j)}|^2 + \sum_j \frac{\varphi_j^2}{\sigma^2_{\varphi_j}}$$

L'annulation de la dérivée de $\mathcal{J}$ par rapport à chaque $\varphi_j$ donne :

$$\sum_{k \neq j} \frac{1}{\sigma^2_{j,k}} \left[ |y^{\mathrm{data}}_{j,k}||y^{o*}_{j,k})| \sin(\phi_{j,k}(\boldsymbol{x}) - \phi^{\mathrm{data}}_{j,k} - \varphi_k + \varphi_j) \right] = \frac{\rho_j^2}{\sigma^2_{\rho_j}}$$

En considérant l'argument des sinus $\sin(\phi_{j,k}(\boldsymbol{x}) - \phi_{j,k}^{\text{data}} - \varphi_k + \varphi_j)$ petits, et en linéarisant les sinus, on peut déterminer les aberrations en phase en résolvant une équation linéaire.

#### 4.1.2.4 Une méthode de type Gerchberg-Saxton

Comme le note Le Besnerais (1993b), les méthodes standard d'auto-calibration reposent sur un mécanisme de double projection, décrit par Gerchberg & Saxton (1972) dans un cadre particulier[2] puis modifié et utilisé par Fienup (1982) en reconstruction d'image. Cet algorithme, dit de Gerchberg-Saxton, consiste à « coincer » l'objet entre deux ensembles de contraintes dans des espaces duaux :

- des contraintes dans le domaine de Fourier, correspondant aux mesures, d'une part ;
- d'autre part, des contraintes dans l'espace dual, c'est-à-dire l'espace objet, correspondant aux informations *a priori*, telles que la positivité des pixels, ou un « ressort de rappel » ramenant l'objet $\boldsymbol{x}$ vers l'objet *a priori* $\boldsymbol{x}^0$ (cf. section 1.4.2.4)

#### 4.1.2.5 Un algorithme spécifique à la radio-astronomie

La méthode de Cornwell et Wilkinson nous semble intéressante parcequ'elle propose un modèle de données faisant apparaître explicitement les aberrations, et une algorithmie permettant d'estimer conjointement l'objet d'intérêt et les aberrations.

Cependant, des différences majeures entre les contextes radio et optique nous empêchent d'utiliser directement cette méthode sur des données interférométriques optiques :

- Les données considérées par Cornwell et Wilkinson sont des visibilités complexes corrompues par les aberrations et le bruit, à partir desquelles on peut former des clôtures de phase. En optiques, les observables sont directement des visibilités carrées et des clôtures, à partir desquelles on peut éventuellement se ramener à un modèle myope de pseudo-données de visibilités complexes, comme nous l'avons montré au chapitre précédent ;
- Le modèle de données proposé par les auteurs correspond à des mesures de visibilités complexes affectées d'un bruit **circulaire**. En optique, les pseudo-données de visibilités complexes évoquées précédemment sont affectées d'un bruit complexe analogue à celui décrit au chapitre 2, et dont une approximation convexe optimale conduit à un modèle de bruit **elliptique** ;
- Dans l'étape d'optimisation à objet fixé, les auteurs utilisent un *a priori* sur les aberrations de phase. Comme nous l'avons vu section 1.2.2 page 25, de tels *a priori* n'existent pas dans le domaine optique.

#### 4.1.2.6 Les travaux d'André Lannes *et al.* : une adaptation à l'optique incomplète

Lannes (1998a) propose une méthode de reconstruction d'image en interférométrie optique, dédiée au cas « weak phase », c'est-à-dire à 3 ou 4 télescopes, et dont la différence majeure avec la méthode de Cornwell et Wilkinson consiste à remplacer l'algorithme CLEAN par WIPE [Lannes *et al.* (1997)], et à ne pas considérer d'*a priori* sur les aberrations de phase.
Cette adaptation au domaine optique est incomplète, car les données expérimentales considérées par Lannes (1998a) sont toujours des visibilités complexes, affectées par la turbulence, et non des clôtures et des visibilités carrées. De plus, le modèle de bruit sur ces visibilités complexes est un modèle gaussien circulaire, qui n'est pas adapté au domaine optique (cf. section 2.2 page 44).

[2] l'estimation de phase à partir de deux mesures d'intensité

### 4.1.3 Adapter ces méthodes au domaine Optique

Plutôt que d'utiliser le critère d'attache aux données exact, défini par (eq. 3.12 page 80)

$$\mathcal{J}^{data}(\boldsymbol{x}) = \chi^2_s(\boldsymbol{x}) + \chi^2_\beta(\boldsymbol{x}),$$

nous avons préféré un critère d'attache aux données déduit du modèle myope décrit dans le chapitre précédent : un tel modèle permet en effet de faire apparaître explicitement l'incomplétude des données en phase sous forme d'aberrations myopes à estimer conjointement avec l'objet. Nous démontrons l'équivalence de ce critère d'attache aux données myope et du critère d'attache aux données exact en section 4.2.2 page 99).

Pour faire cette estimation conjointe, les méthodes d'auto calibration sont évidemment une source d'inspiration privilégiée. Nous avons donc choisi de

**garder** le schéma d'estimation alternée de l'objet et des aberrations (figure 4.1), en minimisant un critère joint approché (section 4.2) ;

**modifier** l'étape de détermination des aberrations à objet connu, qui doit être elle aussi adaptée, en tenant compte de l'absence d'*a priori* sur les aberrations de phase (section 4.3) ;

**modifier** l'étape de détermination de l'objet à aberrations connues : CLEAN est certes un algorithme de référence en radio-astronomie, mais on peut le remplacer par un algorithme plus adapté au cadre optique du problème traité, c'est-à-dire utilisant des *a priori* plus appropriés et un meilleur contrôle de convergence. En effet, à aberrations fixées, le problème à résoudre est analogue à celui traité avec l'algorithme RIME, au chapitre 2 ;

**proposer** l'algorithme WISARD(Weak-phase Interferometric Sample Alternating Reconstruction Device), décrit section 4.4.

## 4.2 Le critère joint approché $\mathcal{J}_{WISARD}$

### 4.2.1 Construction du critère joint approché $\mathcal{J}_{WISARD}$

On définit la matrice $\boldsymbol{H}(\boldsymbol{\alpha})(t)$ par :

$$\boldsymbol{H}(\boldsymbol{\alpha}(t), t) \triangleq \text{Diag}(\exp \mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}(t))\boldsymbol{H}(t).$$

Avec cette définition, on a :

$$\begin{aligned} \boldsymbol{H}(t) &= \boldsymbol{H}(\boldsymbol{0}, t) \\ \arg(\boldsymbol{H}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}) &= \arg(\boldsymbol{H}(t)\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) \\ |\boldsymbol{H}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}| &= |\boldsymbol{H}(t)\boldsymbol{x}| \end{aligned}$$

Le modèle des équations 3.33 et 3.34 peut donc s'écrire :

$$\left\{ \begin{aligned} \boldsymbol{\phi}^{\text{data}}(t) &= \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}^{\text{b}}(t) \\ &= \arg(\boldsymbol{H}(t)\,\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}(t) + \boldsymbol{\phi}^{\text{b}}(t) \\ &= \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{\alpha}(t)) + \boldsymbol{\phi}^{\text{b}}(t) \\ &= \arg(\boldsymbol{H}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}) + \boldsymbol{\phi}^{b}(t) \end{aligned} \right. \tag{4.1}$$

$$\begin{cases} \boldsymbol{a}^{data}(t) = \boldsymbol{a}(\boldsymbol{x},t) + \boldsymbol{a}^b(t) \\ \qquad\quad = |\boldsymbol{H}(t)\boldsymbol{x}| + \boldsymbol{a}^b(t) \\ \qquad\quad = |\boldsymbol{H}(\boldsymbol{\alpha}(t),t)\boldsymbol{x}| + \boldsymbol{a}^b(t) \end{cases} \tag{4.2}$$

Cette modélisation est identique à la modélisation 2.1 et 2.2 page 43, exceptés
- l'influence de la turbulence, qui se caractérise par une incertitude sur la matrice $\boldsymbol{H}(\boldsymbol{\alpha})$ ;
- le fait que l'on sépare les données selon l'instant $t$ de mesure

En prenant $\boldsymbol{\alpha} = \boldsymbol{0}$, on retrouve bien le problème traité au chapitre 2.
Nous pouvons donc procéder de façon strictement analogue :
- On reforme des pseudo mesures de visibilités complexes $\boldsymbol{y}^{data}$ et les résidus correspondants $\boldsymbol{y}^{\mathrm{b}}$ selon la méthode présentée section 2.3.2.1 page 50 ;
- On fait une approximation elliptique de la distribution de bruit $\boldsymbol{y}^{\mathrm{b}}$ selon la méthode présentée section 2.3.2.4 page 52 ;
- On forme un critère d'attache aux données selon l'équation 2.21 page 56 :

$$\mathcal{J}_{ell}(\boldsymbol{x},\boldsymbol{\alpha}(t),t) = \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t),t)\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}}(t) \cdot \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t),t)\boldsymbol{x}\right)$$

Le critère final est la somme des termes $\mathcal{J}_{ell}(\boldsymbol{x},\boldsymbol{\alpha}(t),t)$ formés pour chaque instant de mesure :

$$\mathcal{J}_{WISARD}(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_t \mathcal{J}_{ell}(\boldsymbol{x},\boldsymbol{\alpha}(t),t)$$

Nous avons formé un critère joint, approché, sur des pseudo mesures de visibilités complexes. Nous nous sommes alors demandé si ce critère est « équivalent » au critère vrai (et comment définir l'équivalence entre deux critères d'attache aux données).

### 4.2.2 Équivalence de $\mathcal{J}_{WISARD}$ et du critère exact

Nous avons d'abord cherché à comparer le critère joint $\mathcal{J}_{WISARD}$ approché et le critère vrai. Il suffit de considérer un seul instant $t$ de mesure. On comparera donc $\mathcal{J}_{ell}(\boldsymbol{x},\boldsymbol{\alpha})$ au critère vrai $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x})$.
Le but de cette section est de montrer que, si l'on minimise le critère $\mathcal{J}_{ell}$ – c'est-à-dire si $\mathcal{J}_{ell}(\boldsymbol{x},\boldsymbol{\alpha}) < \epsilon$ – alors on minimise aussi le critère vrai – $\mathcal{J}^{\mathrm{data}}(\boldsymbol{x}) < \epsilon'$.

On a

$$\mathcal{J}_{ell} = \sum_i \mathcal{J}_{ell}^i$$

Supposons $\mathcal{J}_{ell} < \epsilon$. Alors $\mathcal{J}_{ell}^i < \epsilon$.
$\mathcal{J}_{ell}^i$ se décompose en (équation E.1 page 171, annexe E) :

$$\mathcal{J}_{ell}^i = \frac{(a_i(\boldsymbol{x})\cos\phi_{\mathrm{res},i} - a_i^{\mathrm{data}})^2}{\sigma_{\mathrm{rad},i}^2} + \frac{(a_i(\boldsymbol{x})\sin\phi_{\mathrm{res},i})^2}{\sigma_{\mathrm{tan},i}^2},$$

avec la notation, pour toute variable $v$

$$v_{\mathrm{res}} \triangleq -v_{\mathrm{b}} \tag{4.3}$$

En l'occurrence, on a (Eqs. 4.1 et 4.2) :

$$\begin{cases} \phi_{\mathrm{res},i} \triangleq -\phi_{b,i} = \phi_i(\boldsymbol{x},\boldsymbol{\alpha}) - \phi_i^{\mathrm{data}} \\ a_{\mathrm{res},i} \triangleq -a_{b,i} = a_i(\boldsymbol{x}) - a_i^{\mathrm{data}} \end{cases} \tag{4.4}$$

On a donc

$$\begin{aligned}
\mathcal{J}_{ell} < \epsilon &\Rightarrow \mathcal{J}_{ell}^{i} < \epsilon,\ \forall i \\
&\Rightarrow \frac{(a_i(\boldsymbol{x})\cos\phi_{\text{res},i} - a_i^{\text{data}})^2}{\sigma_{\text{rad},i}^2} + \frac{(a_i(\boldsymbol{x})\sin\phi_{\text{res},i})^2}{\sigma_{\text{tan},i}^2} < \epsilon,\ \forall i \\
&\Rightarrow \frac{(a_i(\boldsymbol{x})\cos\phi_{\text{res},i} - a_i^{\text{data}})^2}{\sigma_{\text{rad},i}^2} < \epsilon \quad \textbf{et} \quad \frac{(a_i(\boldsymbol{x})\sin\phi_{\text{res},i})^2}{\sigma_{\text{tan},i}^2} < \epsilon,\ \forall i \\
&\Rightarrow (a_i(\boldsymbol{x})\cos\phi_{\text{res},i} - a_i^{\text{data}})^2 < k_1\epsilon \quad \textbf{et} \quad (a_i(\boldsymbol{x})\sin\phi_{\text{res},i})^2 < k_1\epsilon,\ \forall i
\end{aligned} \tag{4.5}$$

avec $k_1 = \max_i\{\sigma_{\text{rad},i}^2, \sigma_{\text{tan},i}^2\}$. De plus, on montre aisément que

$$(a_i(\boldsymbol{x})\cos\phi_{\text{res},i} - a_i^{\text{data}})^2 = (a_i(\boldsymbol{x}) - a_i^{\text{data}})^2 + 2a_i(\boldsymbol{x})a_i^{\text{data}}(1-\cos\phi_{\text{res},i}) - a_i^2(\boldsymbol{x})\sin^2\phi_{\text{res},i} \tag{4.6}$$

En combinant les équations 4.5 et 4.6, on obtient :

$$\mathcal{J}_{ell} < \epsilon \Rightarrow (a_i(\boldsymbol{x}) - a_i^{\text{data}})^2 + 2\underbrace{a_i(\boldsymbol{x})a_i^{\text{data}}}_{>0}\underbrace{(1-\cos\phi_{\text{res},i})}_{\geq 0} < k_1\epsilon$$

Par ailleurs, on a toujours (Eq. 4.5) :

$$\sin^2\phi_{\text{res},i} < \frac{k_1\epsilon}{a_i^2(\boldsymbol{x})},\ \forall i$$

On peut donc écrire :

$$\begin{aligned}
\mathcal{J}_{ell} < \epsilon &\Rightarrow (a_i(\boldsymbol{x}) - a_i^{\text{data}})^2 < k_1\epsilon \quad \textbf{et} \quad |\sin\phi_{\text{res},i}| < \frac{\sqrt{k_1\epsilon}}{a_i(\boldsymbol{x})},\ \forall i \\
&\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad \textbf{et} \quad |\sin\phi_{\text{res},i}| < \frac{\sqrt{k_1\epsilon}}{a_i(\boldsymbol{x})},\ \forall i
\end{aligned}$$

On suppose $\phi_{\text{res},i} \ll \pi$. Avec cette hypothèse, et pour $\epsilon$ suffisamment petit, on a $|\sin\phi_{\text{res},i}|$ petit, et donc :

$$|\sin\phi_{\text{res},i}| = \sin|\phi_{\text{res},i}|$$

De plus, si $\sin x < \sqrt{3}/2$, alors $x \leq 2\sin x$ . On peut donc écrire :

$$\begin{aligned}
\mathcal{J}_{ell} < \epsilon &\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad \textbf{et} \quad |\phi_{\text{res},i}| < 2\sin|\phi_{\text{res},i}| = 2|\sin\phi_{\text{res},i}| < \frac{2\sqrt{k_1\epsilon}}{a_i(\boldsymbol{x})},\ \forall i \\
&\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad \textbf{et} \quad |\phi_{\text{res},i}|^2 < \frac{4k_1\epsilon}{a_i^2(\boldsymbol{x})},\ \forall i \\
&\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad \textbf{et} \quad |\phi_{\text{res},i}| < \frac{2\sqrt{k_1\epsilon}}{a_i(\boldsymbol{x})},\ \forall i \\
&\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad \textbf{et} \quad |\phi_{\text{res},i}| < \sqrt{k_2\epsilon},\ \forall i
\end{aligned}$$

avec $k_2 = \min_i\{\frac{4k_1}{a_i^2(\boldsymbol{x})}\}$. Une clôture de phase s'écrit $\boldsymbol{C\phi}$ et le résidu de clôture vérifie :

$$\boldsymbol{\beta}_{\text{res}} = \boldsymbol{C\phi}_{\text{res}}$$

De plus, l'opérateur $\boldsymbol{C}$ est une matrice contenant des 1, des $-1$ et des 0. Comme $|a+b| < |a| + |b|$, alors

$$|\boldsymbol{C\phi}| < n\ \max_i |\phi_i|$$

avec $n$ le nombre de coefficients de $\boldsymbol{C}$.

$$\begin{aligned}
\mathcal{J}_{ell} < \epsilon \Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad &\textbf{et} \quad |\phi_{\text{res},i}| < \sqrt{k_2\epsilon},\ \forall i \\
\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad &\textbf{et} \quad n \max_i |\phi_{\text{res},i}| < n\sqrt{k_2\epsilon} \\
\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad &\textbf{et} \quad |\boldsymbol{C}\boldsymbol{\phi}_{\text{res}}| < n\sqrt{k_2\epsilon} \\
\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad &\textbf{et} \quad |\beta_{\text{res},j}| < n\sqrt{k_2\epsilon},\ \forall i,j \\
\Rightarrow a_{\text{res},i}^2 < k_1\epsilon \quad &\textbf{et} \quad \beta_{\text{res},j}^2 < k_2 n^2\epsilon,\ \forall i,j \\
\Rightarrow \frac{a_{\text{res},i}^2}{\sigma_{a,i}^2} < k_3\epsilon \quad &\textbf{et} \quad \frac{\beta_{\text{res},j}^2}{\sigma_{\beta,i}^2} < k_4\epsilon,\ \forall i,j
\end{aligned}$$

avec $k_3 = \frac{k_1}{\min_i \sigma_{a,i}^2}$, $k_4 = \frac{k_2 n^2}{\min_i \sigma_{\beta,i}^2}$.

Le critère d'attache aux données exact s'écrit comme une somme de termes du type :

$$\mathcal{J}^{\text{data}} = \sum_i \frac{a_{\text{res},i}^2}{\sigma_{a,i}^2} + \sum_j \frac{\beta_{\text{res},j}^2}{\sigma_{\beta,i}^2}$$

On voit que, pour $\epsilon$ suffisamment petit, $\mathcal{J}^{\text{data}}$ peut être rendu aussi petit que l'on veut. En termes mathématiques, on a :

$$\forall \epsilon,\ \exists \epsilon',\ \mathcal{J}_{ell} < \epsilon' \Rightarrow \mathcal{J}^{\text{data}} < \epsilon \tag{4.7}$$

Le critère myope approché et le critère « vrai » sont donc équivalents[3]. En terme d'ensemble de niveaux (cf. eq. 2.5 page 45), cette propriété s'exprime ainsi :
Pour tout ensemble de niveau $\bar{\text{Iso}}_k \mathcal{J}^{\text{data}}$, il existe $k'$ tel que

$$\bar{\text{Iso}}_{k'} \mathcal{J}_{ell} \subset \bar{\text{Iso}}_k \mathcal{J}^{\text{data}}$$

### 4.2.3 Non convexité du critère $\mathcal{J}_{WISARD}$

Même si le critère $\mathcal{J}_{WISARD}$ est convexe par rapport à l'objet, il ne l'est pas par rapport à l'ensemble des variables aberrations et objet.

En particulier, il est connu que les données de clôture et de visibilités carrées ne sont pas sensibles à une translation de l'objet. Le critère $\mathcal{J}^{\text{data}}(\boldsymbol{x})$ est donc invariant par translation de $\boldsymbol{x}$. Or ce n'est pas le cas du critère $\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha})$ qui, à aberrations $\boldsymbol{\alpha}$ fixées, dépend de la position de l'objet $\boldsymbol{x}$. Pourtant, les modèles de données myope et exact sont équivalents, et cette propriété d'invariance par translation est donc elle aussi vérifiée par le critère myope. Il est facile (mais fastidieux) de montrer que pour chaque translation $T(\boldsymbol{x})$ de l'objet $x$, il existe une pente de phase $\boldsymbol{k}_T \cdot \boldsymbol{r}$, telles que :

$$\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha}) = \mathcal{J}_{WISARD}(T(\boldsymbol{x}), \boldsymbol{\alpha} + \boldsymbol{k}_T \cdot \boldsymbol{r})$$

Le critère **joint** $\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha})$ est donc lui aussi invariant par l'addition d'une translation de l'objet et d'une pente de phase correspondante sur les aberrations.

De façon à apprécier l'influence d'une variation globale de tous les pistons, différentiels on va visualiser l'influence d'une pente globale des phases (tip en abscisse, tilt en ordonnée) sur le critère (fig 4.2) La conséquence est que le centrage de l'objet reconstruit dépend de l'objet

[3] Le critère myope approché considéré dans cette démonstration est le développement limité de l'approximation gaussienne optimale décrit section 2.3.2.4. Le critère myope approché est de même minimum que le critère vrai, ce qui n'est pas le cas pour l'approximation optimale. Pour cette dernière, l'équivalence avec le critère vrai n'est pas vérifiée.

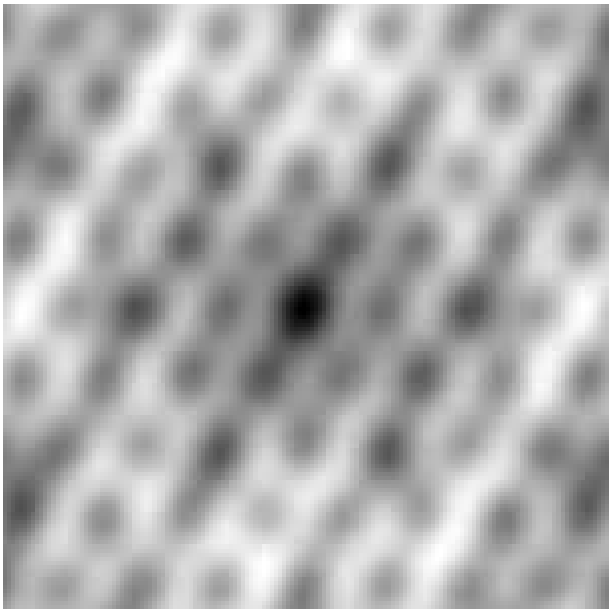


FIG. 4.2 – Influence sur le critère d'un tip-tilt global

d'initialisation.

Ces dégénérescences, qui visent à illustrer la non convexité du critère $\mathcal{J}_{WISARD}$, ne sont pas problématiques, puisque l'on cherche à reconstruire un objet à une translation près. Par contre, la non-convexité du critère due à l'incomplétude des données en phase, donne lieu à des minima locaux. Comme nous l'avons mentionné en introduction de ce chapitre, l'enjeu de l'optimisation est de parcourir un « bon » chemin, à partir d'un « bon » point de départ (l'initialisation), pour espérer arriver à un « bon » minimum. L'adjonction d'informations *a priori* pertinentes est un moyen d'atteindre ces objectifs. Il est cependant aussi nécessaire de disposer d'une stratégie de minimisation performante.

### 4.2.4 Choix d'une stratégie de minimisation alternée

Nous disposons d'un critère joint dépendant des aberrations et de l'objet. Il a été construit de telle façon qu'à aberrations fixées, le problème à résoudre soit convexe. Nous avons donc opté pour une stratégie de minimisation **alternée**, de type auto-calibration (Cf Section 4.1.2 page 95) :

1. On propose un objet d'initialisation $\boldsymbol{x} = \boldsymbol{x}_0$ ;
2. Pour l'objet courant $\hat{\boldsymbol{x}}$, on trouve les aberrations optimales, c'est-à-dire minimisant le critère joint :
$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \mathcal{J}_{WISARD}(\hat{\boldsymbol{x}}, \boldsymbol{\alpha})$$
Cette étape fait l'objet de la section 4.3, et utilise une méthode de descente par gradients conjugués ;
3. Pour les aberrations courantes $\hat{\boldsymbol{\alpha}}$, on trouve l'objet optimal, c'est-à-dire minimisant le critère joint :
$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \mathcal{J}_{WISARD}(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$$
Cette étape est strictement équivalente au problème traité chapitre 2. La méthode que nous avions utilisée incluait une méthode de descente utilisant le calcul du gradient du critère ;

4. On revient à l'étape 2, jusqu'à convergence.

Il a donc été nécessaire de calculer les gradients de ce critère joint par rapport à $(\boldsymbol{x}, \boldsymbol{\alpha})$. On note
- $\mathcal{J}_{WISARD}(\boldsymbol{x}|\boldsymbol{\alpha})$ le critère à $\boldsymbol{\alpha}$ fixés ;
- $\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})$ le critère à $\boldsymbol{x}$ fixé.

Le calcul de ces gradients fait l'objet de l'annexe E.

## 4.3 Détermination des aberrations optimales par auto calibration

### 4.3.1 Séparation du problème d'optimisation en aberrations

A $\boldsymbol{x}$ fixé, on recherche les aberrations

$$\hat{\boldsymbol{\alpha}} = [\hat{\boldsymbol{\alpha}}(t_1), \hat{\boldsymbol{\alpha}}(t_2), \ldots]$$

telles que

$$\frac{\partial \mathcal{J}_{WISARD}(\hat{\boldsymbol{\alpha}}|\boldsymbol{x})}{\partial \boldsymbol{\alpha}} = 0$$

Comme ce gradient se divise en termes de la forme

$$\frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \boldsymbol{\alpha}(t)},$$

le problème de recherche des aberrations optimales peut donc se découper en plusieurs problèmes, un pour chaque instant de mesure.

**A chaque fois, on n'a à déterminer que** $\operatorname{card} \boldsymbol{\alpha}(t) = N_t - 1$ **aberrations.**

Dans la suite de la section 4.3, on ne considérera donc que les données prises à un instant $t$, c'est-à-dire qu'on ne cherche à résoudre que le problème :

$$\frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \hat{\boldsymbol{\alpha}}(t)} = 0$$

On ne tient pas compte de redondances à deux instants différents, c'est-à-dire qu'une même base intervient dans plusieurs clôtures mesurées au même instant (pour un interféromètre à plus de 3 télescopes), mais n'intervient pas dans deux clôtures mesurées à deux instants différents (contrairement au cas de l'interféromètre de l'IRCOM traité section 3.4.5). Puisqu'il n'y a pas d'ambiguïtés, on omettra la dépendance en $t$ :

$$\frac{\partial \mathcal{J}_{ell}(\hat{\boldsymbol{\alpha}}|\boldsymbol{x})}{\partial \boldsymbol{\alpha}} = 0$$

Pour chaque instant de mesure, il faut résoudre une équation du type $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = 0$, soit, d'après l'équation E.5 page 172

$$\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = \bar{\boldsymbol{B}}^{\mathrm{T}} \cdot \boldsymbol{W}_1 \sin 2\boldsymbol{\phi}_{\mathrm{res}} + \bar{\boldsymbol{B}}^{\mathrm{T}} \cdot \boldsymbol{W}_2 \sin \boldsymbol{\phi}_{\mathrm{res}} = 0$$

avec

$$\boldsymbol{\phi}_{\mathrm{res}} = \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} - \boldsymbol{\phi}^{\mathrm{data}}$$

$$\boldsymbol{W}_1 = \mathrm{Diag}\left\{w_{1,i}\right\} = \mathrm{Diag}\left\{\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\mathrm{tan},i}} - \frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\mathrm{rad},i}}\right\}$$

$$\boldsymbol{W}_2 = \mathrm{Diag}\left\{w_{2,i}\right\} = \mathrm{Diag}\left\{2\frac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma^2_{\mathrm{rad},i}}\right\}$$

### 4.3.2 Recherche exhaustive des aberrations optimales dans le cas d'un interféromètre à 3 télescopes

A 3 télescopes, il ne faut déterminer pour chaque instant que $N_t - 1 = 2$ aberrations myopes $\alpha_1, \alpha_2$. Une descente locale n'est donc pas nécessaire : il suffit de tracer une carte 2D de la valeur du critère $\mathcal{J}_{ell}(\alpha_1, \alpha_2|\boldsymbol{x})$, et de sélectionner les valeurs de $(\alpha_1, \alpha_2)$ pour lesquelles le critère est minimal. Si le critère à minimiser a plusieurs minima, cette méthode permet de se placer autour du minimum global.
En pratique, on calcule la valeur du critère sur une grille de valeurs de $(\alpha_1, \alpha_2)$. Si la grille de valeurs pour lesquelles on a calculé le critère est suffisamment fine (une grille de $100 \times 100$ valeurs sur $[0, 2\pi] \times [0, 2\pi]$ est généralement suffisante), l'estimée $\hat{\boldsymbol{\alpha}}_{\text{grille}}$ sera suffisamment proche du minimum global pour qu'il soit atteint par une descente locale initialisée à $\hat{\boldsymbol{\alpha}}_{\text{grille}}$.
On a (Eq. E.1 page 171) :

$$\mathcal{J}_{ell}^{i} = \frac{a_i^2(\boldsymbol{x})}{\sigma_{\text{rad},i}^2} \cos^2 \phi_{\text{res},i} + \frac{a_{\text{data},i}^2}{\sigma_{\text{rad},i}^2} - 2\frac{a_i(\boldsymbol{x}) a_i^{\text{data}}}{\sigma_{\text{rad},i}^2} \cos \phi_{\text{res},i} + \frac{a_i^2(\boldsymbol{x})}{\sigma_{\text{tan},i}^2} \sin^2 \phi_{\text{res},i}$$

avec

$$\phi_{\text{res},i} = \bar{\boldsymbol{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

Les fonctions $\sin^2$ et $\cos^2$ sont de période $\pi$ et introduisent des dégénérescences dans le critère à minimiser. La carte 2D de la valeur du critère $\mathcal{J}_{ell}(\alpha_1, \alpha_2|\boldsymbol{x})$ peut avoir l'allure de la figure 4.3. Le tracé de cette carte pour $100 \times 100$ valeurs permet de déterminer le minimum global. Cependant, ce tracé n'est réalisable que dans le cas particulier $N_{tels} = 3$, et notre objectif est de construire une méthode de reconstruction la plus générale possible. Nous avons cherché à trouver une méthode de calcul des aberrations optimales pour un interféromètre à plus de 3 telescopes.

### 4.3.3 Solution linéarisée en $\boldsymbol{\phi}_{\text{res}}$

On ne peut résoudre de façon analytique l'equation

$$\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_1 \sin 2\boldsymbol{\phi}_{\text{res}} + \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_2 \sin \boldsymbol{\phi}_{\text{res}} = 0$$

Une première façon de résoudre cette equation est de considérer les résidus en phase suffisamment petits pour pouvoir faire l'approximation $\sin \boldsymbol{\phi}_{\text{res}} \simeq \boldsymbol{\phi}_{\text{res}}$ (c'est l'approche proposée par Cornwell & Wilkinson (1981) :

$$\begin{aligned} \frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = 0 &\Rightarrow \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_1 \sin 2\boldsymbol{\phi}_{\text{res}} + \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_2 \sin \boldsymbol{\phi}_{\text{res}} = 0 \\ &\Rightarrow 2\bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_1 \boldsymbol{\phi}_{\text{res}} + \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_2 \boldsymbol{\phi}_{\text{res}} = 0 \\ &\Rightarrow 2\bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_1 \boldsymbol{\phi}_{\text{res}} + \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_2 \boldsymbol{\phi}_{\text{res}} = 0 \\ &\Rightarrow \bar{\boldsymbol{B}}^{\mathrm{T}} \underbrace{[2\boldsymbol{W}_1 + \boldsymbol{W}_2]}_{\boldsymbol{W}} \boldsymbol{\phi}_{\text{res}} = 0 \end{aligned}$$

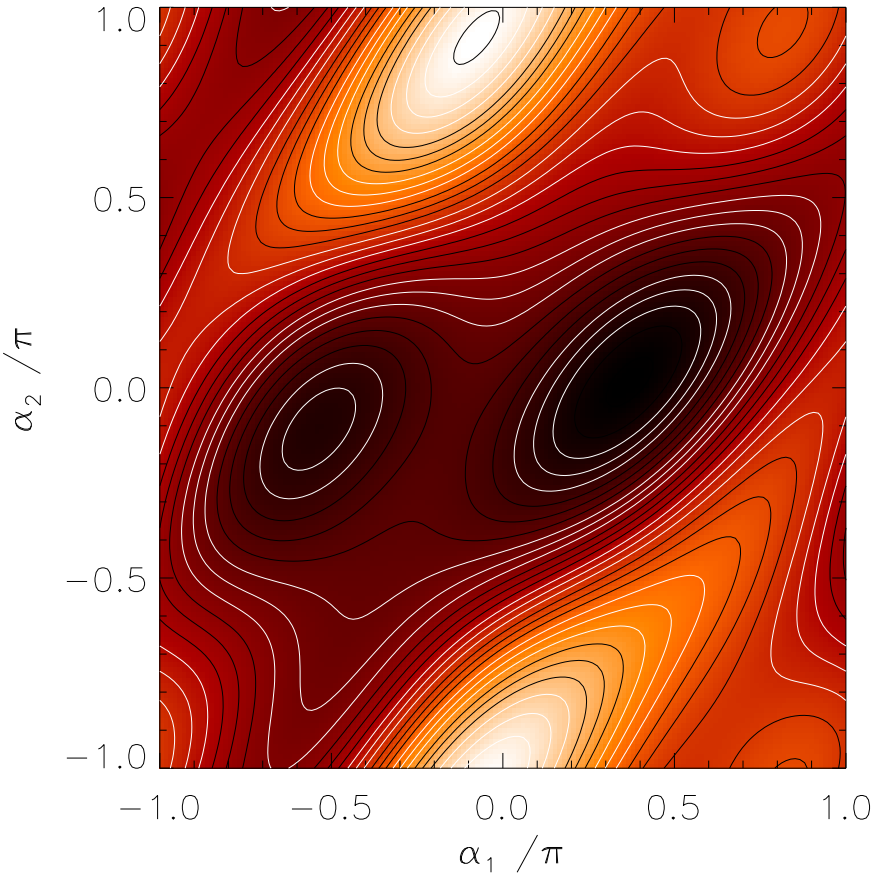



FIG. 4.3 – Critère $\mathcal{J}_{ell}$ en fonction des deux aberrations myopes à déterminer.

Comme $\boldsymbol{\phi}_{\text{res}} = \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} - \boldsymbol{\phi}^{\text{data}}$ (Eq. 4.4), les aberrations optimales $\hat{\boldsymbol{\alpha}}_{\text{lin}}$ vérifient :

$$\begin{aligned}\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = 0 &\Rightarrow \bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\left(\boldsymbol{\phi}_{\text{res}}\right) = 0 \\ &\Rightarrow \bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\left(\boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\hat{\boldsymbol{\alpha}}_{\text{lin}} - \boldsymbol{\phi}^{\text{data}}\right) = 0 \\ &\Rightarrow \bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\bar{\boldsymbol{B}}\hat{\boldsymbol{\alpha}}_{\text{lin}} = \bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\left(\boldsymbol{\phi}^{\text{data}} - \boldsymbol{\phi}(\boldsymbol{x})\right)\end{aligned}$$

D'où

$$\boxed{\hat{\boldsymbol{\alpha}}_{\text{lin}} = \left[\bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\bar{\boldsymbol{B}}\right]^{-1}\bar{\boldsymbol{B}}^{\text{T}}\boldsymbol{W}\left(\boldsymbol{\phi}^{\text{data}} - \boldsymbol{\phi}(\boldsymbol{x})\right)} \tag{4.8}$$

avec

$$\boldsymbol{W} \triangleq \text{Diag}\left\{w_i\right\} \triangleq 2\text{Diag}\left\{\frac{a_i^2(\boldsymbol{x})}{\sigma_{\text{tan},i}^2} - \frac{a_i^2(\boldsymbol{x})}{\sigma_{\text{rad},i}^2} + \frac{a_i(\boldsymbol{x})a_i^{\text{data}}}{\sigma_{\text{rad},i}^2}\right\} \tag{4.9}$$

$\hat{\boldsymbol{\alpha}}_{\text{lin}}$ est fonction de l'écart entre les phases $\boldsymbol{\phi}^{\text{data}}$ et $\boldsymbol{\phi}(\boldsymbol{x})$, et non d'une fonction périodique de cet écart. On observe une discontinuité lorsque cet écart est dans un voisinage de $2\pi$. On est donc soumis à des problèmes d'enroulement de phase. Cependant, cette solution linéarisée peut servir d'initialisation pour une descente locale du critère en aberrations.

### 4.3.4 Solution linéarisée en $\boldsymbol{\phi}_{\text{res}}$ et $\boldsymbol{a}_{\text{res}}$

Nous avons cherché à analyser la matrice $\boldsymbol{W}$ pour vérifier que le terme $\left[\bar{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{W}\bar{\boldsymbol{B}}\right]$ de l'équation 4.8 est bien inversible.

Si l'on suppose que le résidu en amplitude $\boldsymbol{a}_{\text{res}} = \boldsymbol{a}(\boldsymbol{x}) - \boldsymbol{a}^{\text{data}}$ est petit, le terme $w_i$ de l'équation 4.9 s'écrit :

$$\begin{aligned} w_i &= 2\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{tan},i}} - 2\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{rad},i}} + 2\frac{a_i(\boldsymbol{x})a_i^{\text{data}}}{\sigma^2_{\text{rad},i}} \\ &\simeq 2\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{tan},i}} - \frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{rad},i}} + 2\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{rad},i}} \\ &\simeq 2\frac{a_i^2(\boldsymbol{x})}{\sigma^2_{\text{tan},i}} \\ &\simeq 2\left(\frac{a_i^{\text{data}}(\boldsymbol{x})}{\sigma_{\text{tan},i}}\right)^2 \end{aligned}$$

Or on a construit le modèle de bruit approché avec (Eq. 2.17 page 52) :

$$\frac{\sigma^2_{\text{tan},i} \simeq a_i^{\text{data}}(\boldsymbol{x})^2\sigma^2_{\phi,i}}{2}$$

On a donc :

$$\boldsymbol{W} \simeq \text{Diag}\left\{\frac{2}{\boldsymbol{\sigma}^2_{\phi}}\right\} \simeq 2\boldsymbol{R}_{\phi}^{-1}$$

La matrice $\boldsymbol{W}$ est donc inversible. Comme $\bar{\boldsymbol{B}}^{\mathrm{T}}$ est de rang $N_t - 1$ et qu'elle a $N_t - 1$ lignes (cf. Annexe B), elle est surjective. On peut alors montrer que le terme $\left[\bar{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{W}\bar{\boldsymbol{B}}\right]$ de l'équation 4.8 est bien inversible.
Ces calculs permettent de surcroît de proposer une solution linéarisée à la fois en $\boldsymbol{\phi}_{\text{res}}$ et $\boldsymbol{a}_{\text{res}}$ :

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{\text{lin2}} &= \left[\bar{\boldsymbol{B}}^{\mathrm{T}}2\boldsymbol{R}_{\phi}^{-1}\bar{\boldsymbol{B}}\right]^{-1}\bar{\boldsymbol{B}}^{\mathrm{T}}2\boldsymbol{R}_{\phi}^{-1}\left(\boldsymbol{\phi}^{\text{data}} - \boldsymbol{\phi}(\boldsymbol{x})\right) \\ &= \left[\bar{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{R}_{\phi}^{-1}\bar{\boldsymbol{B}}\right]^{-1}\bar{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{R}_{\phi}^{-1}\left(\boldsymbol{\phi}^{\text{data}} - \boldsymbol{\phi}(\boldsymbol{x})\right) \end{aligned} \tag{4.10}$$

Cette solution présente néanmoins la même sensibilité aux enroulements de phase que la solution $\hat{\boldsymbol{\alpha}}_{\text{lin}}$.

### 4.3.5 Solution itérative

Pour s'abstraire des problèmes d'enroulement de phase, on prend une approche de type méthode de Newton, c'est-à-dire itérative, pour déterminer $\hat{\boldsymbol{\alpha}}$ annulant (Eq. E.5 page 172) :

$$\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}) = \bar{\boldsymbol{B}}^{\mathrm{T}}\,\boldsymbol{W}_1 \sin 2\boldsymbol{\phi}_{\text{res}}(\boldsymbol{\alpha}) + \bar{\boldsymbol{B}}^{\mathrm{T}}\,\boldsymbol{W}_2 \sin \boldsymbol{\phi}_{\text{res}}(\boldsymbol{\alpha}) = 0$$

Supposons une petite variation $\delta\boldsymbol{\alpha}$ de $\boldsymbol{\alpha}$. Alors,pour $\delta\boldsymbol{\alpha}$ est suffisamment petit, on peut faire le developpement limité au premier ordre suivant :

$$\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) \simeq \frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}) + \frac{\partial^2 \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}^2}(\boldsymbol{\alpha})\,\delta\boldsymbol{\alpha} \tag{4.11}$$

avec

$$\frac{\partial^2 \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}^2}(\boldsymbol{\alpha}) = 2\bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_1 \bar{\boldsymbol{B}} \cos\left[2\boldsymbol{\phi}_{\mathrm{res}}(\boldsymbol{\alpha})\right] + \bar{\boldsymbol{B}}^{\mathrm{T}} \boldsymbol{W}_2 \bar{\boldsymbol{B}} \cos\left[\boldsymbol{\phi}_{\mathrm{res}}(\boldsymbol{\alpha})\right] \tag{4.12}$$

On peut donc procéder de façon itérative :

1. On choisit une initialisation $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ la meilleure possible, par exemple la solution de l'equation linéarisée (4.8) ;
2. On calcule $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha})$ ;
3. On cherche à annuler $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}$. Comme $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) \simeq \frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}) + \frac{\partial^2 \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}^2}(\boldsymbol{\alpha})\,\delta\boldsymbol{\alpha}$, on cherche la variation de $\boldsymbol{\alpha}$ telle que $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) = 0$, soit :

$$\delta\boldsymbol{\alpha} = -\frac{\partial^2 \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}^2}(\boldsymbol{\alpha})^{-1} \frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha});$$

4. On pose $\boldsymbol{\alpha} = \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}$ et on revient à l'etape 2.

Si la convergence d'un tel algorithme n'est pas garantie, nous avons constaté son efficacité en pratique. La méthode décrite au paragraphe suivant permet de ne pas être pénalisé dans le cas où cette méthode divergerait.

### 4.3.6 Une méthode robuste de détermination des aberrations

Nous avons proposé 4 façons de determiner les aberrations optimales à objet connu :
- $\hat{\boldsymbol{\alpha}_{\mathrm{grille}}}$ obtenues par recherche exhaustive sur une grille (Section 4.3.2 page 104), dans le cas d'un interféromètre à 3 telescopes ;
- $\hat{\boldsymbol{\alpha}_{\mathrm{lin}}}$ obtenues par linéarisation de l'équation $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = 0$ (Section 4.3.3 page 104) ;
- $\hat{\boldsymbol{\alpha}_{\mathrm{iter}}}$ obtenues par recherche itérative d'une solution de l'équation $\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = 0$ (Section 4.3.5 page 106) ;
- $\hat{\boldsymbol{\alpha}_{\mathrm{ante}}}$ correspondant aux aberrations calculées au cycle d'auto-calibration précédent.

Notre méthode consiste à calculer les valeurs du critère $\mathcal{J}_{WISARD}$ pour ces quatre estimations, choisir la meilleure (celle pour laquelle le critère est minimum) comme initialisation, puis minimiser localement le critère à l'aide de la méthode de descente décrite section 2.4.2 page 57, utilisant le calcul du gradient de $\mathcal{J}_{WISARD}$ à objet connu (Section E.2 page 169).

## 4.4 L'algorithme WISARD : Weak-phase Interferometric Sample Alternating Reconstruction Device

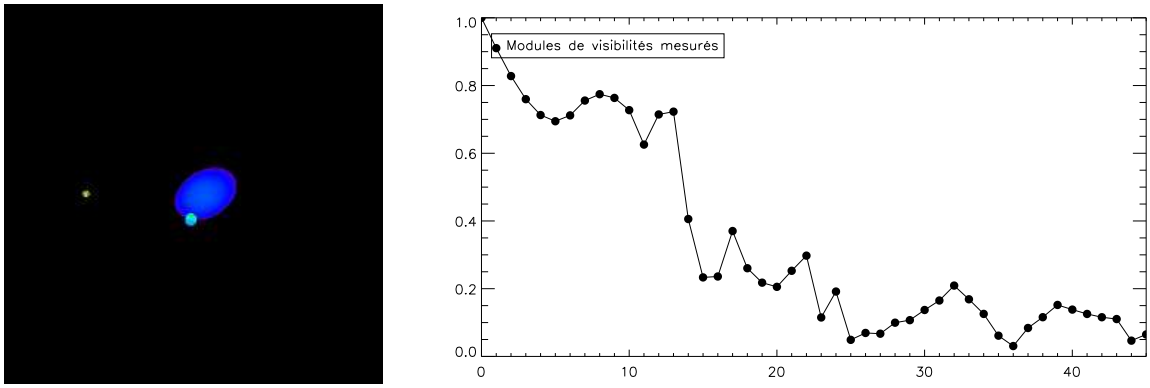
L'étape de détermination de l'objet à aberrations connues dont nous disposions à ce stade de l'étude avait été construite puis validée sur un problème de reconstruction non myope. Cependant, le caractère myope du problème de reconstruction d'images à partir de clôtures et de visibilités carrées d'une part, et la confrontation de nos méthodes à de nouveaux types d'objets d'autre part, nous ont conduits à faire évoluer la méthode décrite au chapitre 2 dans deux directions :
- Le critère joint $\mathcal{J}_{WISARD}$ n'est pas convexe, même si le critère que l'on minimise à aberrations connues l'est, et même si l'on peut déterminer les aberrations optimales à coup sûr (dans le cas à 3 télescopes). La solution obtenue par minimisation de ce critère est donc dépendante de l'initialisation choisie. Nous avons donc cherché à proposer une méthode de détermination d'un « bon » objet d'initialisation ;

– Les objets traités au chapitre 2 sont plutôt basse fréquence, c'est-à-dire sans bords francs ni pics isolés. Afin de reconstruire correctement des objets comportant des composantes haute fréquence, nous avons cherché à élargir la gamme des régularisations utilisées à des *a priori* permettant de préserver les eventuelles hautes fréquences de l'objet à reconstruire, tels que les régularisations L1L2 (section 1.4.2.3) et L1L2$_w$ (section 1.4.2.4).

Nos travaux sont essentiellement empiriques, car il n'existe pas aujourd'hui de caractérisation exhaustive des minima locaux. Les propositions décrites dans cette section font suite à de nombreux essais infructueux.

### 4.4.1 Objet d'initialisation

Nous avons cherché à construire un objet d'initialisation décrit par peu de paramètres, lesquels seraient ajustés pour être compatibles avec les données. L'interférométrie optique est souvent utilisé pour mesurer le diamètre de sources que l'on assimile à un disque uniforme. Cependant, il est possible de construire un objet d'initialisation plus complexe, c'est-à-dire avec un nombre de paramètres intermédiaire entre le modèle de disque uniforme (un seul paramètre de diamètre) et la carte finale (de quelques centaines de pixels). Nous avons opté pour un modèle d'objet invariant par rotation. La TF d'un tel objet est aussi à symétrie circulaire.




Objet visibilités absolues mesurées

FIG. 4.4 – Visibilités absolues en fonction de la fréquence radiale.

La figure 4.4 présente un exemple d'objet et la courbe correspondante de visibilités radiales. On forme un objet d'initialisation en faisant la transformée de Fourier inverse de la carte de visibilités complexes invariante par rotation que nous venons de décrire. Cet objet n'est *a priori* pas réel et encore moins positif. On en extrait donc la partie réelle positive. Des exemples d'objets d'initialisation ainsi calculés sont présentés figure 4.5. Cette méthode permet notamment d'identifier les objets constitués de plusieurs sources non superposées.

### 4.4.2 Une reconstruction par fréquences croissantes

Les tests préliminaires de notre méthode, sur des données simulant un interféromètre à 3 télescopes, ont abouti à des reconstructions où l'objet original n'etait pas reconnaissable. Ces échecs ont été interprétés comme la conséquence de la non convexité du critère à minimiser, particulièrement handicapante dans le cas d'un interféromètre à 3 télescopes. Nous avons alors

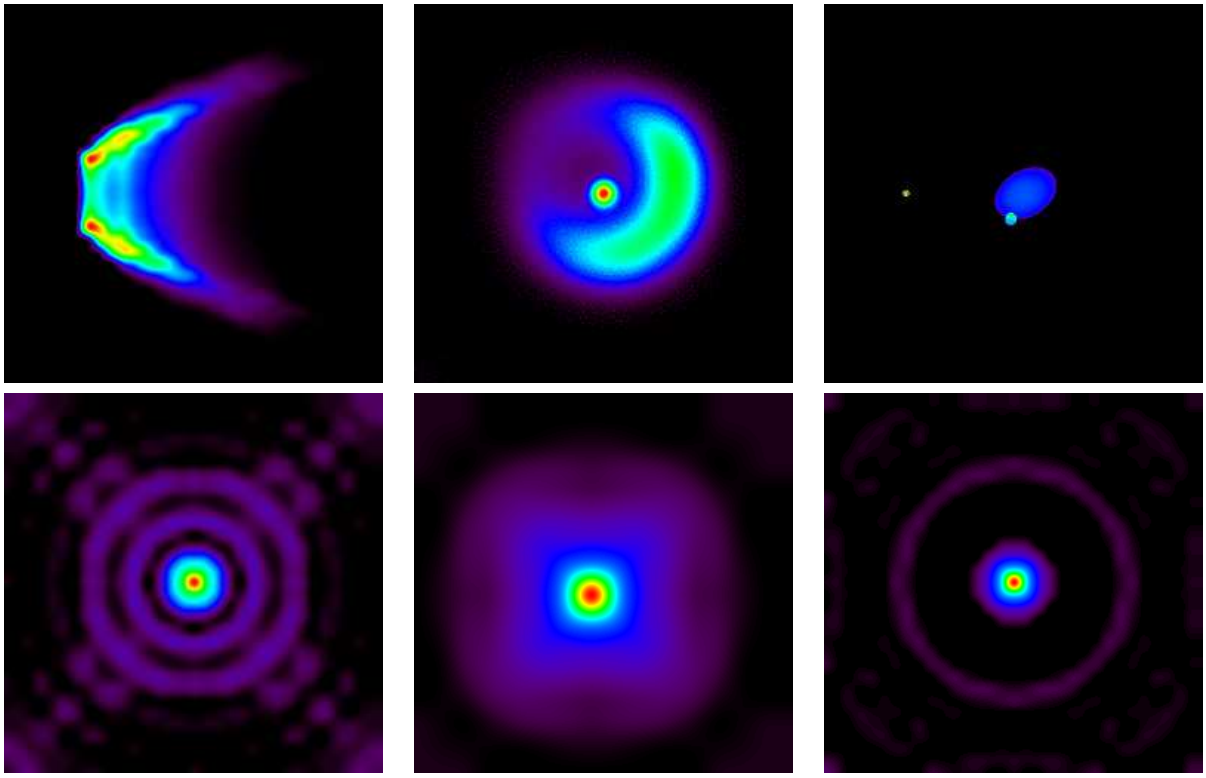


FIG. 4.5 – Initialisation circulaire.

pris le parti de privilégier les structures globales de l'objet, *i. e.* les basses fréquences. Le principe, évoqué en 1994 par Buscher (1994), est de ne sélectionner qu'un ensemble de mesures et de reconstruire un objet à partir de ces seules données, de façon à réduire le nombre de dégénérescences du critère. Nous proposons de ne retenir au départ que les données basses fréquences. Cette reconstuction sert d'initialisation pour une nouvelle reconstruction pour laquelle on rajoute de nouvelles mesures (Figure 4.6).

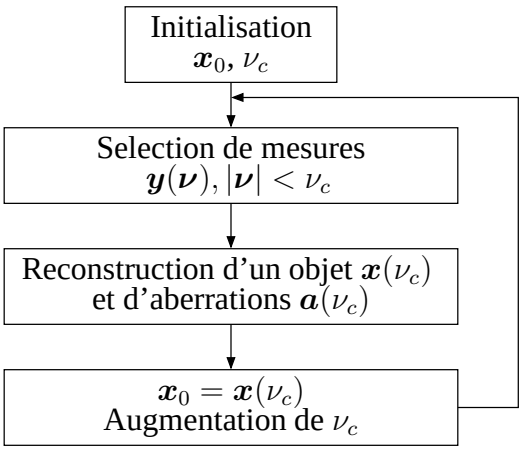



FIG. 4.6 – Reconstruction par fréquences croissantes.

#### 4.4.2.1 Incorporation progressive des données

Dans un premier temps, nous avons effectué cette selection fréquentielle en ne considérant qu'un sous-ensemble d'instants de mesure. En effet, notre méthode ne permet pas de traiter isolément une clôture de phase et les visibilités carrées correspondantes. Nous avons donc ordonné les « paquets » de mesures par la base maximale couverte à cet instant de mesure. L'augmentation de la fréquence de coupure considérée revient alors à « incorporer » un nouveau paquet de mesures.

#### 4.4.2.2 Apodisation variable des données

Cependant, nous avons été confrontés à des configurations interférométriques pour lesquelles, à chaque instant, on mesure à la fois des hautes et des basses fréquences. Il n'est pas possible dans ces cas là de separer les paquets de données en paquets basse fréquence et paquets haute fréquence.
Une autre façon de procéder consiste à apodiser les mesures de visibilité carrées au-delà de la fréquence de coupure. On multiplie donc les données par une fonction d'apodisation $f_{\text{apo}}$, dont on fait varier la largeur de façon à faire croitre la fréquence de coupure. Cette fonction doit être non nulle sur tout le domaine, de façon à préserver les mesures de phases. Une reconstruction parfaite conduirait à reconstruire un objet dont la TF est multipliée par la fonction d'apodisation :

$$\hat{\boldsymbol{x}} = \mathrm{TF}^{-1}\left[\widetilde{\boldsymbol{x}}_{\text{vrai}} f_{\text{apo}}\right] = \boldsymbol{x} * \mathrm{TF}^{-1}\left[f_{\text{apo}}\right]$$

L'objet idéalement reconstruit serait la convolution de l'objet vrai avec la TF inverse de la fenêtre d'apodisation. Celle ci a donc été choisie telle que sa TF inverse soit réelle et positive. Nous avons utilisé une fonction « cloche » dont le profil radial est une fenêtre de Hamming de largeur variable, au centre de laquelle on intercale un palier de hauteur 1 sur la zone de fréquence que l'on ne souhaite pas apodiser.

### 4.4.3 Post-traitement

Il est apparu durant les phases de test de la méthode que dans le cas d'interféromètres à 3 télescopes, où la determination des aberrations est particulièrement difficile, que l'on obtient de meilleures estimations des aberrations en ne régularisant pas l'objet (à part la contrainte de positivité). Pour améliorer la qualité de l'objet, il suffit alors de faire une dernière optimisation de l'objet, régularisée cette fois-ci, sans modifier les aberrations estimées précedemment. Nous avons donc adjoint à notre méthode un dernier module de « post-traitement », qui consiste à faire une minimisation régularisée en objet, à aberrations fixées. Cette méthode, bien qu'heuristique, est très similaire à ce que préconisent Blanc (2002) et Idier *et al.* (2005) dans le cadre de l'estimation d'aberrations par diversité de phase sur un objet étendu.

### 4.4.4 Structure de l'algorithme WISARD

## 4.5 Les versions successives de l'algorithme

L'algorithme ainsi constitué a cependant largement évolué au cours des trois années dont rend compte ce manuscrit. Certaines erreurs importantes n'ont été corrigées que tardivement. J'ai souhaité présenter les resultats obtenus avec l'algorithme tel qu'il était au moment des

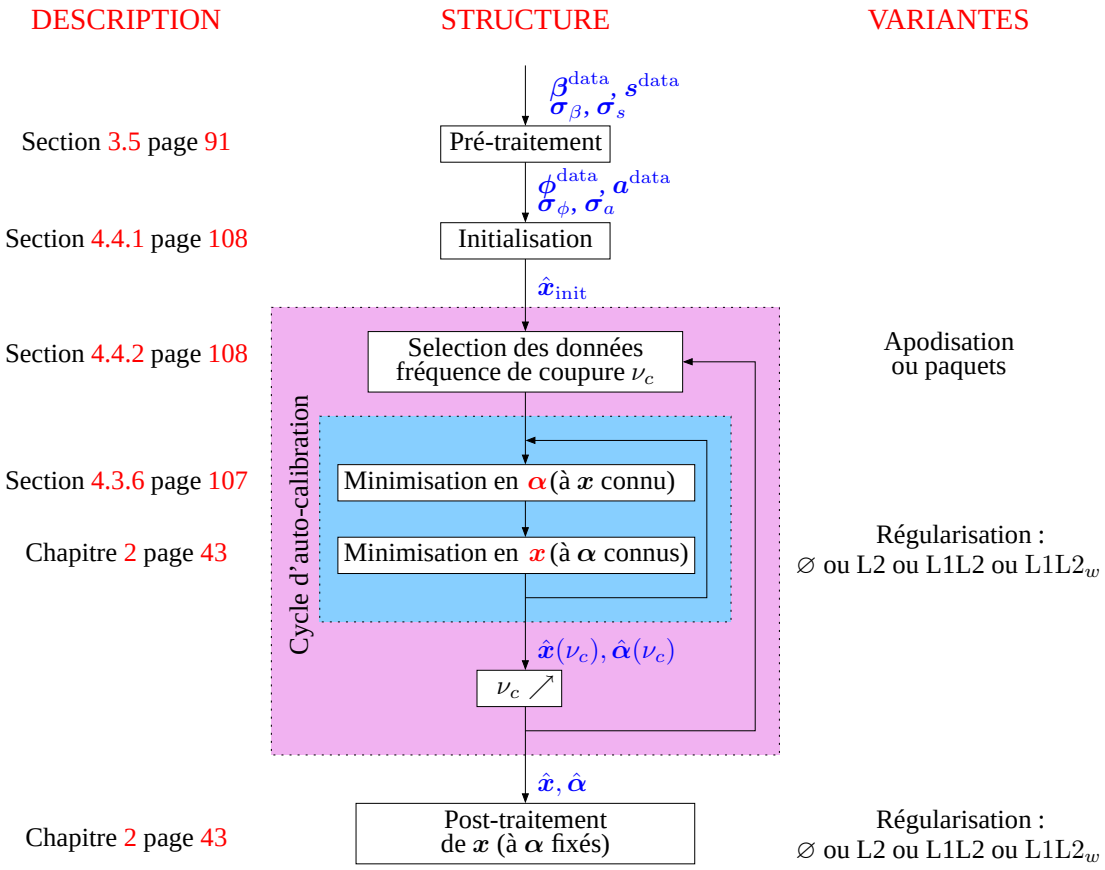



FIG. 4.7 – Structure de l'algorithme WISARD.

simulations, et non uniquement ceux que fournit l'algorithme à son stade actuel. Les simulations et traitements expérimentaux dans le chapitre suivant ont été effectués avec 3 versions différentes de l'algorithme WISARD.

### 4.5.1 WISARD2

La version actuelle de l'algorithme, dite « WISARD3 », est celle qui est décrite dans ce chapitre. Contrairement aux précédentes,

- elle s'appuie sur un modèle de formation des pseudo-données myopes cohérent avec celui des données réelles ;
- elle permet de traiter des données issues d'un interféromètre à n télescopes avec n quelconque ;
- elle propose une large gamme de termes d'*a priori* possibles (blancs ou non, quadratiques ou non).

Les deux versions précédentes sont présentées dans les sections à suivre, la version intermédiaire, dite « WISARD2 », d'abord, la plus ancienne, dite « WISARD1 », ensuite.

### 4.5.2 WISARD2

**Une statistique des visibilités en module erronée** Dans cette version, l'écart-type $\boldsymbol{\sigma}_a$ sur les amplitudes de visibilité était calculé à partir de l'écart-type $\boldsymbol{\sigma}_s$ sur les visibilités carrées par :

$$\boldsymbol{\sigma}_a = \sqrt{\boldsymbol{\sigma}_s}$$

Nous avons vu (Section 3.1 page 74) qu'une modélisation correcte est

$$\sigma_a = \frac{\sigma_s}{2a^{\text{data}}}$$

Cette erreur de modélisation conduit notamment à surestimer l'écart-type du bruit sur les composantes principales de l'objet (les fréquences pour lesquelles la visibilité carrée $s$ est forte), en particulier les basses fréquences (le spectre des objets considérés est en général décroissant).

**Une estimation des aberrations dégénérée** Les méthode d'estimation des aberrations par linéarisation, ou par une méthode itérative, n'était pas disponible à ce stade. Une descente locale était réalisée à partir de l'estimée précedente des aberrations, sauf dans le cas 3 télescopes, ou une recherche exhaustive des aberrations optimales sur une grille était disponnible.

### 4.5.3 WISARD1

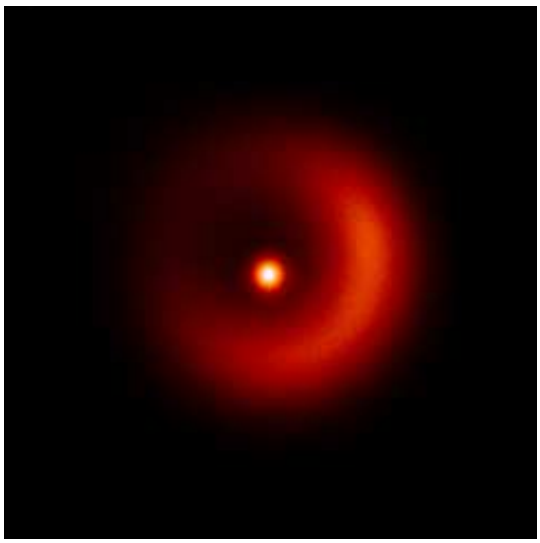
En plus des limitations et erreurs de la version 2 de l'algorithme, cette version présentait les défauts suivants :

**Un modèle de données défavorable** Cette première version de l'algorithme ne traitait que des cas à 3 télescopes. Nous avons néanmoins été conduit à traiter des données simulant un interféromètre à 6 télescopes. En effet, nous avons montré précedemment (Section 3.4.4 page 88) que l'on pourrait théoriquement obtenir avec un interféromètre à 3 telescopes les mêmes données. Nous avons donc traité ces données à 6 télescopes comme si elles étaient issues d'un interféromètre à 3 télescopes.

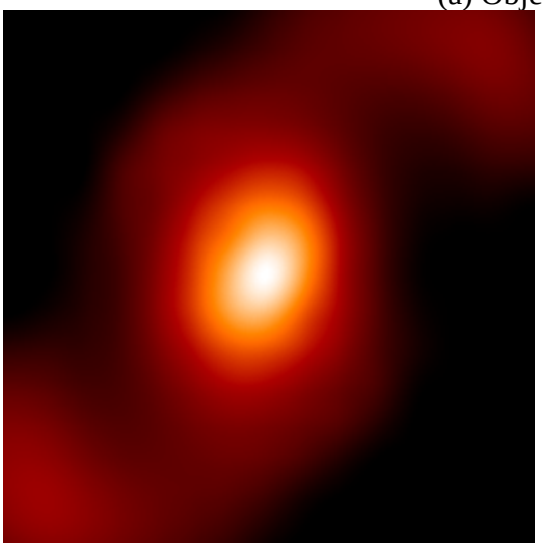
Pour chaque triplet de bases, nous avons donc construit des pseudo-visibilités complexes compatibles avec les 3 mesures de visibilités carrées et la mesure de clôture de phase **du triplet** considéré. Or il y a redondance, c'est à dire que plusieur clôtures « partagent » la même base. Pour les bases communes à deux triplets, les deux pseudo-mesures complexes ainsi construites sont généralement « incohérentes » entre elles, c'est-à-dire que la norme de leur différence est au-delà de quelques ecarts-type.

Cette erreur de synthétisation de pseudo mesures est illustrée figure 4.8. Nous avons déterminé la solution aux moindres carrés $\hat{\boldsymbol{x}}_{MC}$ obtenue à partir des pseudo mesures construites avec un modèle à 3 télescopes d'une part, puis avec la méthode préconisée dans ce manuscrit (Section 3.5 page 91), et qui prend en compte le fait qu'il y a 6 telescopes et non 3. Ceci illustre qu'indépendamment de la méthode de reconstruction, le modèle de données considéré à ce stade de l'évolution de WISARD était erronné.

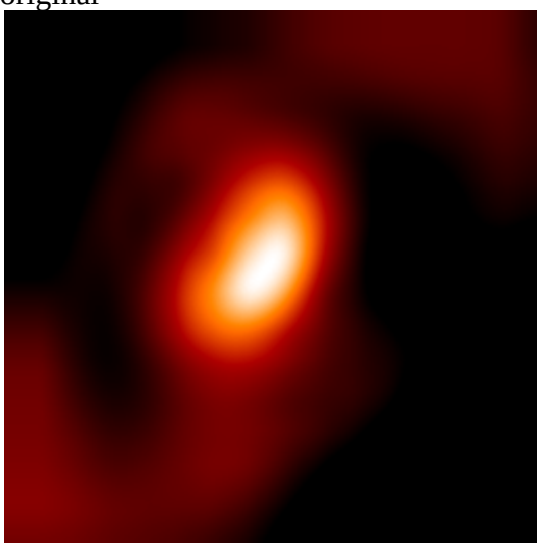
**Un nombre d'aberrations myopes trop important** D'autre part, cette modélisation à 3 télescopes introduit de plus 2 aberrations myopes par clôture et par instant de mesure. Un interféromètre à 6 télescopes permet de former à chaque instant 10 clôtures. Nous avons donc introduit 20 aberrations myopes par instant de mesure, au lieu des 5 nécessaires, soit 4 fois plus d'inconnues à estimer que nécessaire. L'absence à ce stade des méthodes d'estimation des aberrations par linéarisation, ou par méthode itérative, était peu handicapante, puisqu'avec une modélisation à 3 télescopes, une recherche exhaustive des aberrations optimales est possible (Cf. section 4.3.2 page 104).



(a) Objet original



(b)$\hat{\boldsymbol{x}}_{MC}$, Modèle 3 Tels



(c)$\hat{\boldsymbol{x}}_{MC}$, Modèle 6 Tels

FIG. 4.8 – Estimations aux moindres carrés avec un modèle équivalent à 3 télescopes.

**Une régularisation quadratique ou L1L2** A ce stade de l'étude, la régularisation L1L2$_w$ (section 1.4.2.4) n'était pas encore implémentée. Seules les régularisations quadratique et L1L2 sur les gradients étaient disponibles.

Malgré toutes ces limitations, les reconstructions, obtenues avec la version WISARD1 de l'algorithme, que nous avons présentées lors d'un concours internationnal de reconstruction en aveuge (présentées au chapitre suivant, section 5.1 page 115) nous ont permis d'occuper le second rand dans ce concours[Lawson *et al.* (2004)].

## 4.6 Conclusion

Nous avons élaboré une méthode originale de traitement de données interférométriques optiques, où les données d'entrée sont des clôtures de phases et des visibilités carrées. Cependant, étant donnée l'existence de minima locaux, en majorité dûs à l'incomplétude des données en phase, nous n'avons pas d'assurance que notre méthode converge vers une « bonne » solution. Seule la confontation à des données réelles ou simulées permet de mesurer les performances de

WISARD.

# Chapitre 5

# Valider WISARD par simulation puis expérimentalement

Le traitement de données simulées par une équipe indépendante est une forme de test intermédiaire, entre des simulations « maison » et le traitement de données réelles. Devant le développement rapide d'interféromètres optiques à plus de 2 télescopes, fournissant des clôtures de phase en plus des visibilités carrées, l'International Astronomical Union (IAU) a motivé l'organisation d'un concours de reconstruction d'images « en aveugle », organisé par Peter Lawson. Les résultats que nous avons présentés à ce concours sont présentés dans la section 5.1. Ces données sont simulées pour un interféromètre à 6 télescopes, ce qui nous a permis de tester différentes stratégies de choix d'initialisation et d'*a priori*. Cependant, sur des données issues d'interféromètres à plus de 5 télescopes, l'incomplétude des données en phase devient moins critique (cf. section 3.2.1.3).

L'enjeu principal des algorithmes comme WISARD consiste à traiter des données réelles issues d'interféromètres à 3 ou 4 télescopes. Bénéficiant des liens étroits entre l'ONERA et le Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique (LESIA), j'ai pu participer à deux missions d'observation sur l'interféromètre à 3 télescopes IOTA (Infrared Optical Telescope Array), situé au Fred Lawrence Whipple Observatory, en Arizona (cf. section 5.2). Ces missions ont permis l'acquisition de données sur les étoiles $\chi$ du Cygne et $\alpha$ du Bouvier, suffisamment riches pour être traitées par WISARD. Seules les données concernant l'étoile $\chi$ du Cygne ont pour l'instant été traitées. Les résultats obtenus constituent la section 5.3.

## 5.1 Validation de WISARD sur simulations : « An Imaging Beauty Contest »

### 5.1.1 Deux tests préliminaires

Une phase préliminaire du concours de reconstruction d'images de l'IAU, appelé Imaging Beauty Contest, était destinée notamment à vérifier que les participants parvenaient à lire correctement le format de données OIDATA[Pauls *et al.* (2004)] utilisé pour ce concours. Elle consistait à reconstruire deux cartes d'étoile double, dont les paramètres (séparation, angle, diamètre des deux composantes) étaient connus.
Ces tests n'ont pas présenté de difficultés particulières, mais ont permis d'identifier un point dur dans le cas d'une reconstruction en aveugle : la détermination du champ à reconstruire. En effet, si l'on ne dispose pas d'une taille approximative de l'objet à reconstruire, il est difficile

de l'extrapoler des données. L'objet à reconstruire dans l'une des deux phases préliminaires (Figure 5.1) était une étoile double dont les deux composantes étaient séparées de 22 millisecondes d'angle (mas).

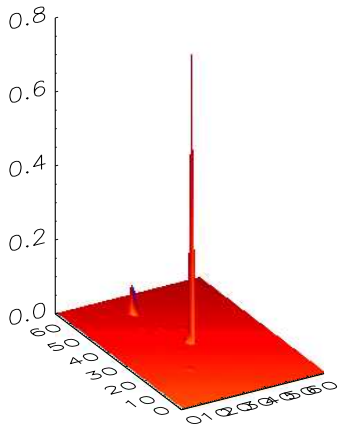



FIG. 5.1 – Reconstruction avec WISARD pour la phase préliminaire de l'Imaging Beauty Contest. $60mas \times 60mas$.

La figure 5.2 montre le profil de modules de visibilités obtenu avec des champs reconstruits de respectivement 60 mas et 30 mas de côté. On constate qu'il est possible de diagnostiquer une sous-estimation du champ reconstruit par l'inadéquation des modules de visibilités de l'objet reconstruit par rapport aux mesures. Ceci se traduit par ailleurs par une valeur trop grande du critère d'attache aux données en fin de convergence. En effet, pour cet exemple comprenant 105 fréquences de mesure, le terme d'attache aux données est un $\chi^2$ à 210 termes (pour chaque mesure complexe, il y a un terme radial et un terme tangentiel). Le $\chi^2$ réduit, c'est-à-dire le terme d'attache aux données divisé par 210, devrait être proche de 1. Il est en effet de $1{,}82$ pour un champ reconstruit suffisant ($60 \times 60$ millisecondes d'angle), or il est de plus de 27 avec un champ reconstruit trop petit ($20 \times 20$ millisecondes d'angle). C'est donc que des composantes de l'objet d'énergie non négligeable sont en dehors du champ reconstruit.

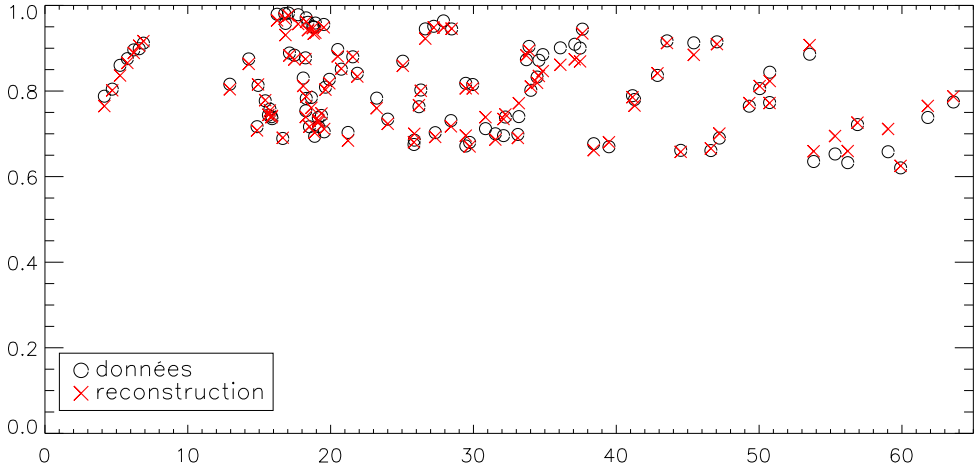
Cette phase préliminaire nous a permis de tester notre algorithme sur des données simulées par une équipe indépendante, et de vérifier que la valeur de $\chi^2$ réduit était un bon diagnostic d'erreur sur le champ reconstruit.

### 5.1.2 Un concours de reconstruction myope en aveugle

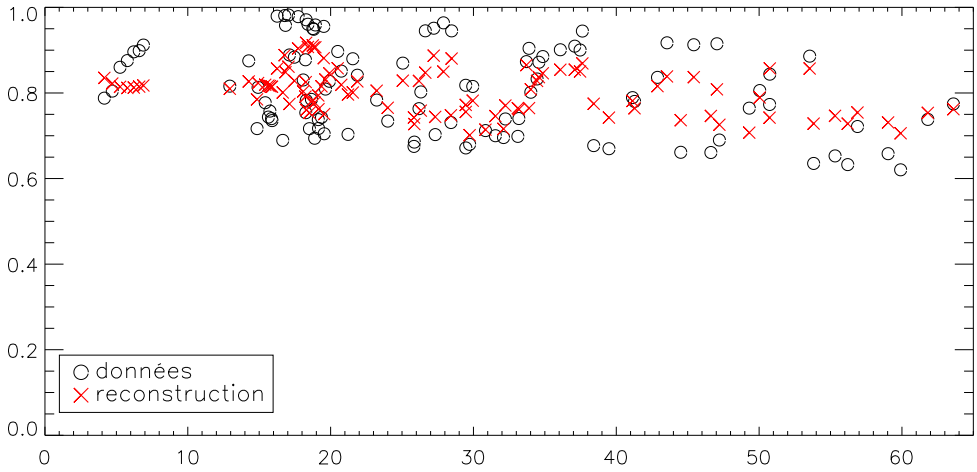
Nous avons alors participé au concours proprement dit. Nous avons donc tenté de reconstruire des objets sans aucune information autre que les données interférométriques, en particulier sans information sur le contraste ou la dimension de ces objets.

Les deux objets de test et les couvertures du plan $(u, v)$ correspondantes sont présentés figures 5.3 et 5.4.

Les reconstructions que nous avons soumises pour le concours ont été obtenues avec la version 1 de l'algorithme WISARD (Cf. Section 4.5.3 page 112). Les meilleurs résultats ont été obtenus en « incorporant » les fréquences par triplets (Cf. Section 4.4.2 page 108), sans régularisation autre que la positivité, puis en faisant un post-traitement avec une régularisation quadratique (la seule disponible à ce stade de l'évolution de l'algorithme) dont les paramètres




Champ reconstruit : $60mas \times 60mas$. $\chi^2$ réduit : $1.83$




Champ reconstruit : $20mas \times 20mas$. $\chi^2$ réduit : $27.56$

FIG. 5.2 – Erreur de reconstruction due au champ reconstruit.

(DSP de l'objet et niveau de bruit) ont été réglés automatiquement (Cf. Section 2.4.3 page 58).

Comme le montre la figure 5.5, nous avons retrouvé correctement les structures globales de l'objet. Le point central sur le premier fichier de test a cependant été « étalé » dans une direction diagonale. Cette direction correspond aux « trous » du plan $u - v$. En effet, le manque de mesure haute fréquence dans une direction rend difficile la reconstruction d'un point source. La régularisation L2 a tendance à ramener vers 0 les visibilités aux fréquences pour lesquelles il n'y a pas de mesure, et donc à étaler le point source dans les directions où il y a peu de données. Ces résultats, s'ils ont permis de mesurer la qualité de notre méthode, et de la comparer favorablement à ses concurrentes au niveau mondial [Meimon *et al.* (2004)] – puisque nous nous sommes placés au second rang –, nous ont poussés à rechercher notamment d'autres types de régularisation, préservant les pics.

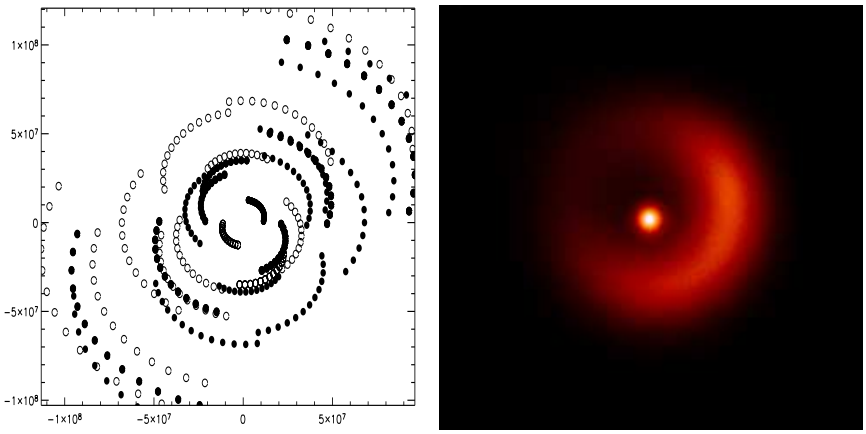



FIG. 5.3 – Objet vrai et couverture $u - v$ pour le fichier de données 1. Ces données simulent l'observation d'un objet correspondant au modèle de LkHa 101, de taille 12,1 $\times$ 12,1 mas, avec le Navy Prototype Optical Interferometer (NPOI). La majorité des mesures sont à faible rapport signal à bruit[Lawson *et al.* (2004)].

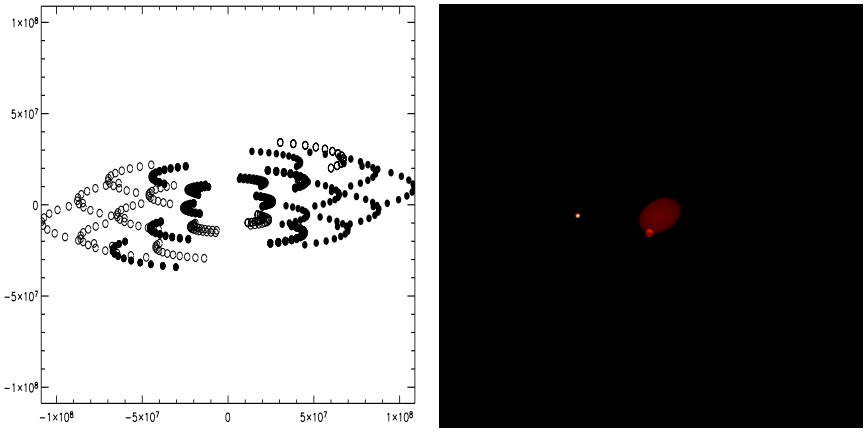
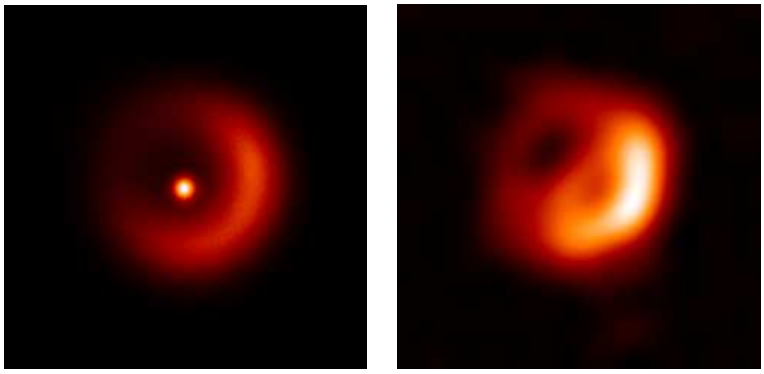
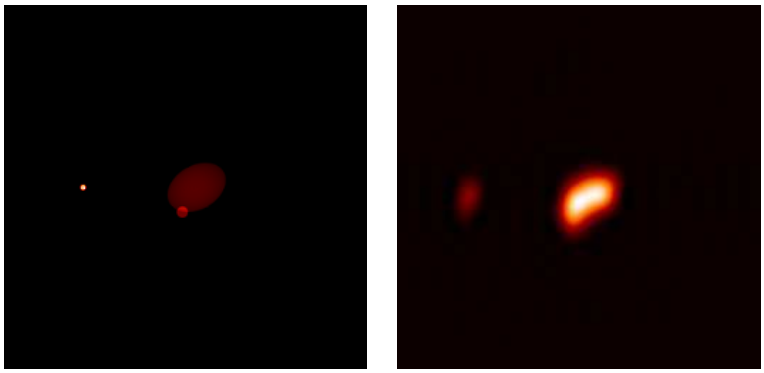



FIG. 5.4 – Objet vrai et couverture $u - v$ pour le fichier de données 1. Ces données simulent l'observation d'un objet correspondant à un modèle d'étoile double, de taille 30 $\times$ 30 mas, avec le Navy Prototype Optical Interferometer (NPOI). La majorité des mesures sont à fort rapport signal à bruit[Lawson *et al.* (2004)].



Fichier test 1 : Objet vrai Objet reconstruit

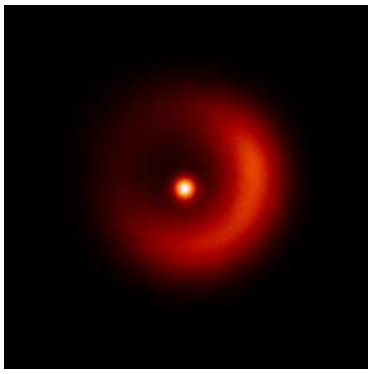


Fichier test 2 : Objet vrai Objet reconstruit

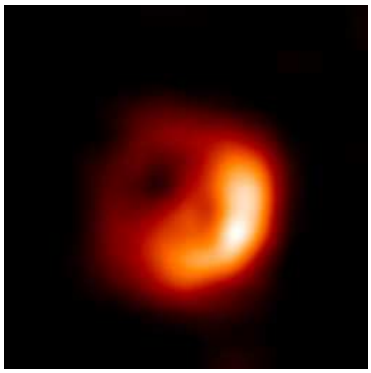
FIG. 5.5 – Objet original et reconstructions soumises à l'Imaging Beauty Contest, obtenues avec WISARD1, régularisation : L2 et positivité.

### 5.1.3 Amélioration des performances de l'algorithme

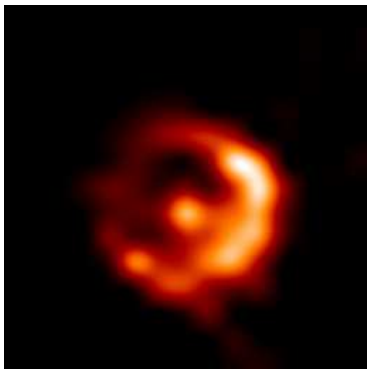
Depuis l'Imaging Beauty Contest, l'algorithme utilisé a profondément évolué (section 4.5 page 110). Nous avons proposé un nouveau type de régularisation préservant les hautes fréquences (WISARD2), adapté l'algorithme aux configurations à plus de 3 télescopes (WISARD2) et corrigé l'erreur sur la statistique des modules de visibilité (WISARD3). La figure 5.6 présente les reconstructions obtenues avec les versions successives de l'algorithme. Il existe bien sûr le risque que la connaissance du vrai objet ait pu orienter nos travaux dans cette seconde série de traitements.



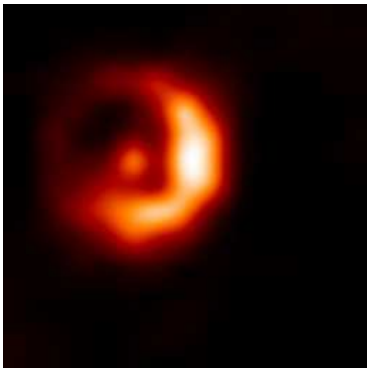
(a)Objet vrai



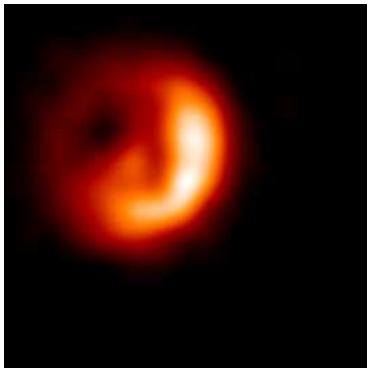
(b)Wisard1, *a priori* L2 et positivité



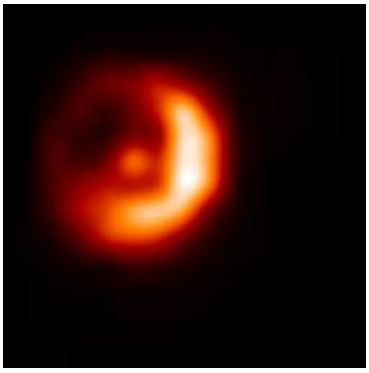
(c)Wisard1, *a priori* L1L2 et positivité



(d)Wisard2, L1L2w et positivité



(e)Wisard3, L2 et positivité



(f)Wisard3, L1L2w et positivité

FIG. 5.6 – Amélioration des performances de WISARD. Les reconstructions (b) et (c) s'appuient sur une modélisation à 3 télescopes, les reconstructions (d), (e) et (f) sur une modélisation à 6 télescopes (voir section 4.5)

### 5.1.4 Conclusions sur les simulations

On peut tirer les conclusions suivantes :

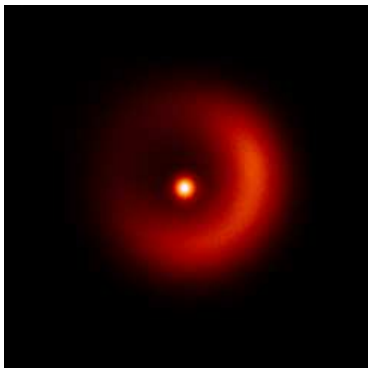
**La régularisation L1L2$_w$** (cf. section 1.4.2.3 page 40) permet de reconstruire avec succès des objets mixtes, c'est-à-dire constitués d'une composante lisse et de détails très piqués. On peut néanmoins formuler des critiques vis-à-vis de la meilleure reconstruction obtenue, correspondant à la figure 5.6(i) : le point central n'est pas correctement résolu, et les intensités maximales de ce pic et de l'enveloppe sont déséquilibrées par rapport à l'image originale. Pour résoudre correctement le pic, il est nécessaire de faire de la super-résolution, c'est-à-dire d'estimer la TF de l'objet à des fréquences au-delà de celles couvertes par l'interféromètre, ce qui ne peut être fait que par l'introduction d'*a priori*. C'est ce que nous faisons, de façon certes imparfaite, en utilisant une régularisation L1L2 blanche. Il est vraisemblablement possible d'améliorer encore la résolution en en utilisant un bi-modèle point source + objet étendu, par exemple avec la méthode proposée par Giovannelli & Coulais (2005). Plus généralement, la photométrie de la reconstruction est imparfaite. Cependant, l'ambition de notre méthode se limite à l'identification des structures globales de l'objet d'intérêt. Pour une reconstruction photométrique, une fois ces structures globales identifiées, il est aisé de n'optimiser que les quelques paramètres fixant les flux et positions de l'enveloppe et du pic central.

**Notre méthode est robuste aux erreurs de modélisation,** comme le montrent les figures 5.6 (b), (c) et (d) qui s'appuient sur un modèle statistique des amplitudes de visibilité erronée.

**Comparaison avec un télescope monolithique** A titre de comparaison, nous avons simulé les images que fourniraient un télescope de 10m de diamètre limité par la diffraction (c'est, au mètre près, le diamètre du Keck à Hawaï, c'est-à-dire le plus grand télescope monolithique en fonctionnement), un télescope de 66m de diamètre limité par la diffraction (c'est la plus grande base utilisée pour simuler les données interférométriques que nous avons traitées) et un télescope de 100m de diamètre (limite supérieure du diamètre du projet de télescope géant OWL). Ces images sont présentées figure 5.7. Il apparaît que le degré de résolution obtenu avec notre méthode de reconstruction est de l'ordre de celui d'un télescope de 100m de diamètre limité par la diffraction.

**Les reconstructions obtenues avec WISARD1 (figures 5.6(e,f)),** qui est dédié aux interféromètres à 3 télescopes, livre quelques enseignements :

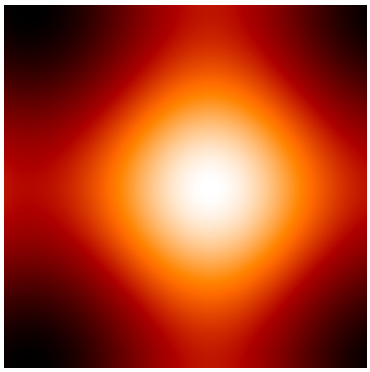
- il est possible de reconstruire une image correcte avec un interféromètre à 3 télescopes, pour une configuration $u - v$ riche et redondante ;
- néanmoins, la convergence est beaucoup plus difficile à 3 télescopes. En particulier, il est pas nécessaire de faire une série de reconstructions par fréquences croissantes (section 4.4.2 page 108), ce qui n'est pas le cas à 6 télescopes. Nous interprétons ceci comme provenant d'un plus grand nombre de minima locaux, le problème étant « plus myope » que dans un cas à 3 télescopes ;
- de ce fait, la convergence est beaucoup plus lente (20 fois plus lente environ lorsque l'on passe de 6 télescopes à 3 télescopes) ;
- à la lumière de ces différences entre les cas à 6 et 3 télescopes, il nous semble qu'il faudrait un Imaging Beauty Contest sur un cas à trois télescopes.



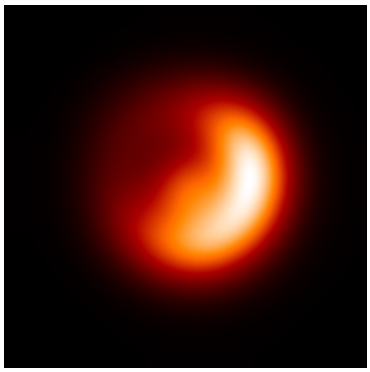
(a)Objet vrai



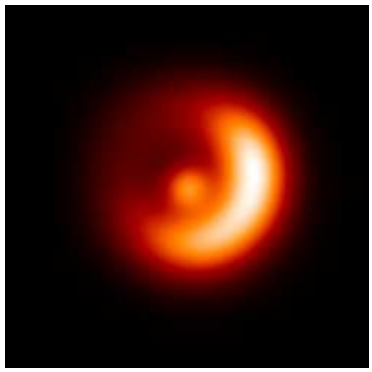
(b)Wisard3, L1L2w et positivité



(c)image d'un tél. de $\varnothing = 10m$



(d)image d'un tél. de $\varnothing = 66m$



(e)image d'un tél. de $\varnothing = 100m$

FIG. 5.7 – Illustration des performances de WISARD. Les images (c), (d) et (e) simulent les images que formeraient des télescopes monolithiques limités par la diffraction, de diamètres 10m, 66m et 100m.

## 5.2 Préparation des missions sur l'interféromètre IOTA

Dans cette section, nous présentons le fonctionnement de l'interféromètre IOTA (section 5.2.2). Le grand nombre de possibilités de positionnement des trois télescopes nous a conduit à proposer une méthode de choix de configuration pupillaire (section 5.2.3). Enfin, dans le but d'adapter ces méthodes aux sources observées, nous présentons brièvement les étoiles de type Mira, dont l'étude constitue l'objectif principal de ces missions (section 5.2.4).

### 5.2.1 Déroulement de la première mission

La première mission (section 5.2) a eu lieu en octobre 2004. L'équipe était constituée de Guy Perrin, qui dirigeait la mission, de Sylvestre Lacour, en cours de thèse avec Guy Perrin au Lesia, de Julien Woillez, en post-doctorat sur le Keck Interferometer à Hawaii, et de moi-même. L'objectif principal était l'étude d'étoiles de type Mira ou super-géante rouge (section 5.2.4). Nous n'avons pu réaliser de mesures que pendant 3 heures, sur les 11 nuits de mission, pour cause de météo défavorable. Si les données recueillies étaient en nombre trop faible pour permettre une reconstruction avec WISARD, cette mission m'a néanmoins permis de me familiariser avec le fonctionnement de l'instrument (section 5.2.2) et de proposer une méthode de choix de positionnement des télescopes de l'interféromètre, décrite section 5.2.3.

### 5.2.2 L'interféromètre IOTA

L'infrared Optical Telescope Array est un interféromètre reconfigurable à trois télescopes, situé au Fred Lawrence Whipple Observatory, en Arizona [Traub *et al.* (2003)]. Il est constitué de deux bras perpendiculaires, un bras long de 35 mètres, sur lequel se déplacent les télescopes A (ou télescope Nord) et C, et un bras court de 15 mètres, sur lequel se déplace le télescope B (ou télescope Sud) (Figure 5.8). Les stations possibles des télescopes sont situées tous les multiples de 5 et 7 mètres sur chaque bras (Figure 5.9).



FIG. 5.8 – Vue générale de l'interféromètre IOTA

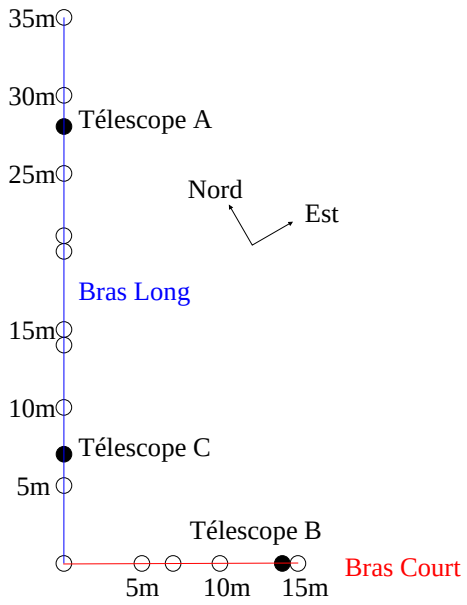



FIG. 5.9 – Vue schématique de IOTA

#### 5.2.2.1 Des sidérostats à la lame séparatrice

Chaque miroir collecteur de 45 cm de diamètre est équipé d'un sidérostat. Le faisceau lumineux issu du miroir primaire est réfléchi successivement sur un miroir secondaire, un miroir de tip/tilt et un miroir d'injection (ou « feed miror »). Le faisceau est alors acheminé vers le laboratoire dans des tubes sous vide.
L'interféromètre IOTA n'a d'abord compté que les deux télescopes A (Nord) et B (Sud). Le schéma de principe de l'interféromètre était alors analogue à celui de la figure 1.1.5 page 19. Les lignes à retard permettent de « ramener » le paquet de franges sur le détecteur en compensant la différence de marche due au fait que la source n'est pas à égale distance des deux ouvertures. Un miroir piezo électrique de balayage permet de faire défiler rapidement les franges sur le détecteur de façon à moduler l'interférogramme. La ligne à retard est constituée d'une ligne à retard longue (ou LD1), assurant le centrage approximatif du paquet de frange sur le détecteur, et d'une ligne à retard courte (SD1), assurant le réglage fin (l'étoile se déplaçant dans le ciel, le retard à introduire est réactualisé en permanence), et la modulation de l'interférogramme. Selon la position de la source dans le ciel, c'est soit le télescope A qu'il faut retarder - « North Delayed » -, soit le télescope B - « South Delayed ».
Le télescope C a été installé par la suite [Traub *et al.* (2000)], avec une deuxième ligne à retard (SD2, LD2). Le schéma de principe de l'interféromètre est présenté figure 5.10 (cas « North Delayed ») et 5.11 (cas « South Delayed »).

L'ensemble des miroirs de chaque voie (15 au total) doit être parfaitement aligné de façon à eviter un effet de vignettage[1] (ou silhouettage), entre autres pour que les flux venant des trois voies soient le plus équilibrés possibles. En particulier, l'alignement des miroirs des longues lignes à retard doit être vérifié chaque fois que celles-ci sont déplacées.

#### 5.2.2.2 La table visible et le suiveur d'étoile, ou « star tracker »

Une lame séparatrice permet d'envoyer la partie visible du faisceau sur un capteur CCD $10 \times 10$ pixels. 4 pixels par voie servent à calculer la position de l'étoile, et à la corriger en temps réel, en asservissant les miroirs de tip/tilt pour chacun des télescopes. Ce système simplifié

[1] On doit pouvoir voir la totalité du miroir primaire depuis le laboratoire

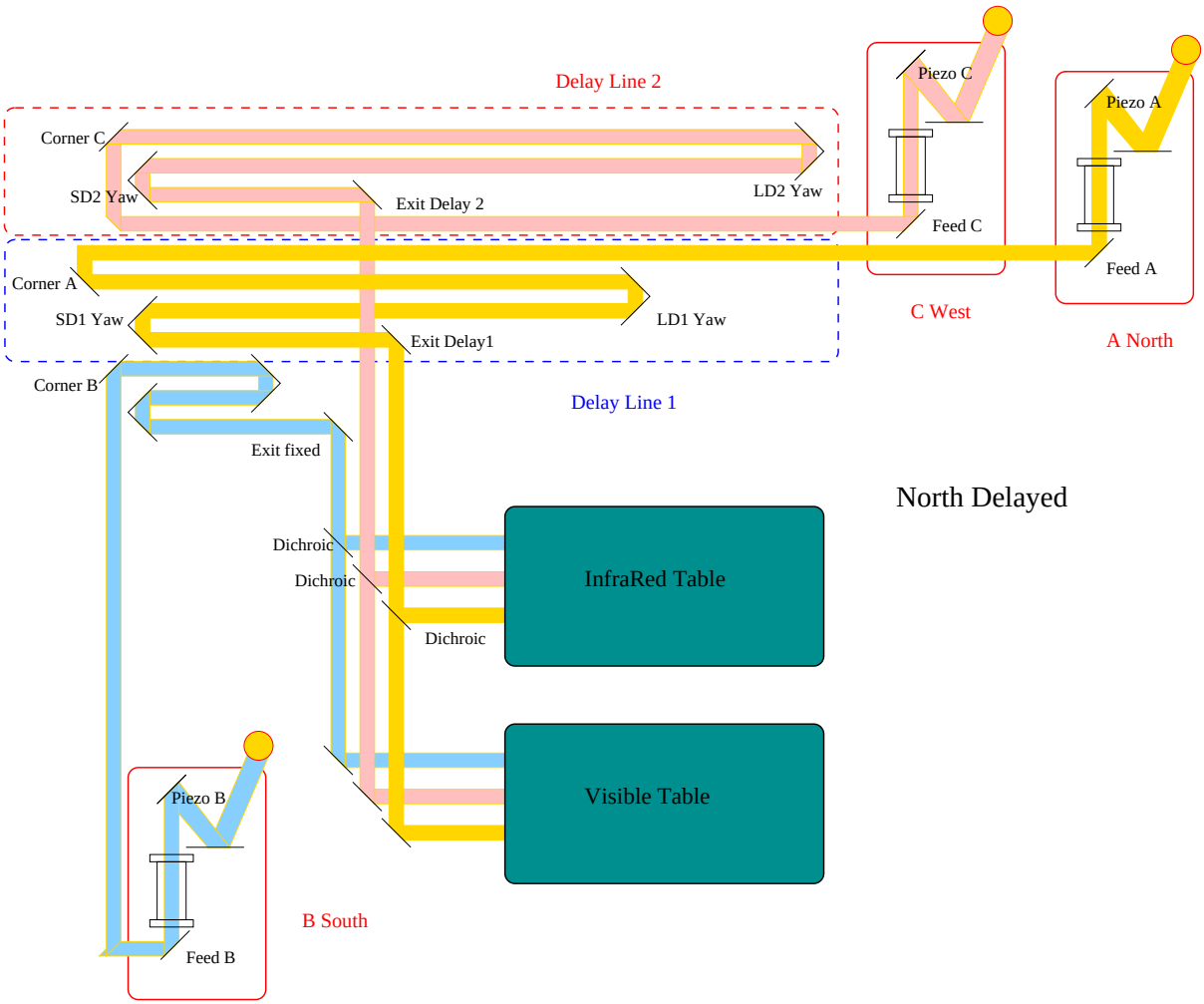



FIG. 5.10 – Schéma de principe de l'interféromètre IOTA, en configuration « North delayed »

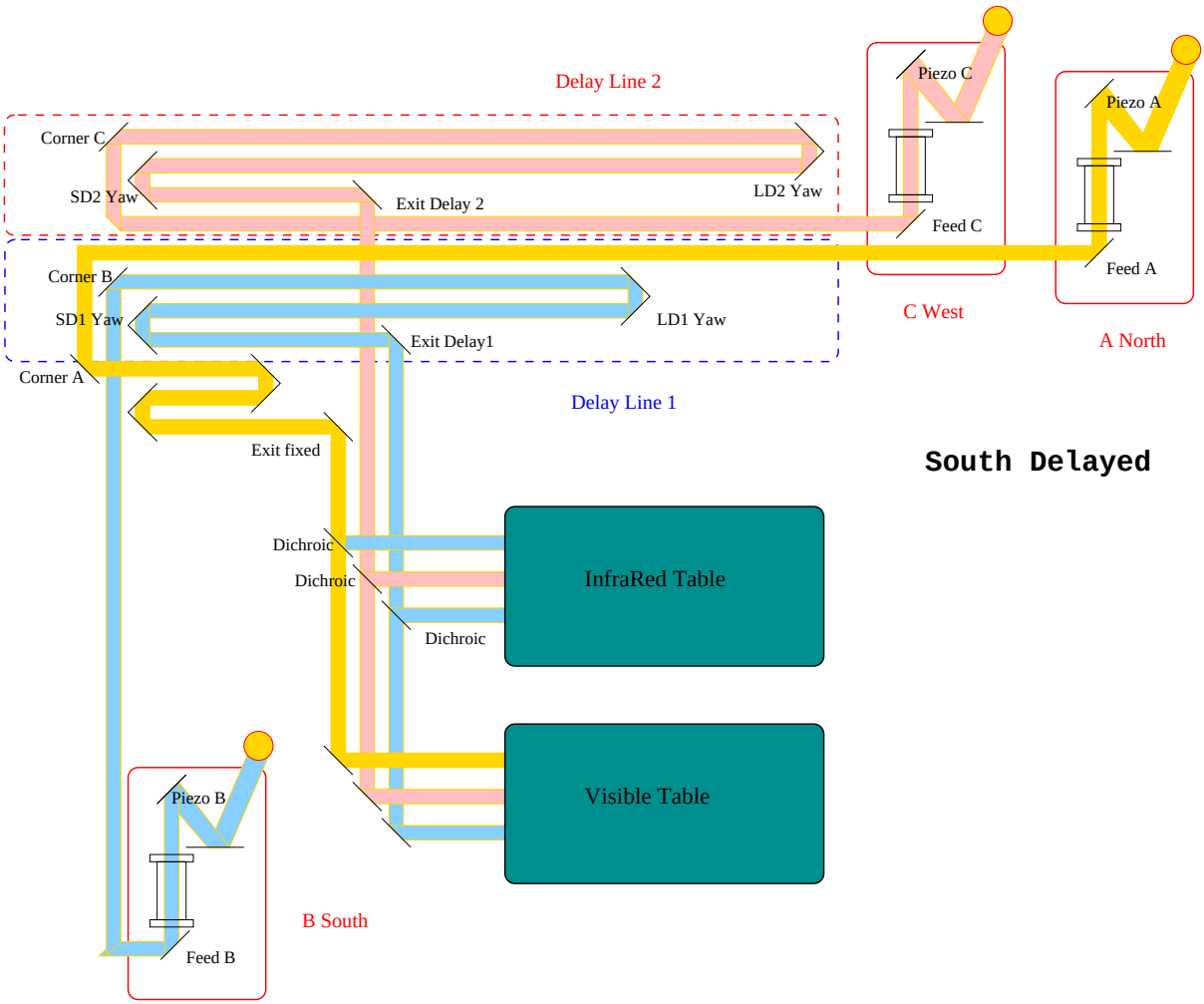



FIG. 5.11 – Schéma de principe de l'interféromètre IOTA, en configuration « South delayed »

d'optique adaptative compense à la fois les effets de tip/tilt dus à la turbulence atmosphérique et les erreurs dans le modèle de pointage.

#### 5.2.2.3 La table Infra rouge et le recombinateur IONIC

La partie infra-rouge du faisceau est envoyée sur des miroirs paraboliques. Ces derniers injectent le flux au foyer de fibres monomode. Ces fibres acheminent les trois faisceaux vers le recombinateur IONIC. Les trois interférogrammes produits sont imagés sur la caméra PICNIC, refroidie à l'azote liquide [Berger *et al.* (2003, 2001); Traub *et al.* (2003); Pedretti *et al.* (2004)]. Il est indispensable après chaque acquisition de franges, de mesurer les composantes de la matrice « kappa », de façon à etalonner la fonction de transfert du recombinateur fibré. De plus, lors d'observation d'objets très brillants (de magnitude inférieure à 1), il est nécessaire d'intercaler des filtres de densité entre le recombinateur et la caméra, de façon à ne pas saturer les détecteurs.
Les visibilités normalisées obtenues sont néanmoins sujettes à des erreurs instrumentales. Les visibilités obtenues sont etalonnées à l'aide d'objets stellaires, idéalement parfaitement connus, ou « calibrateurs ». Durant une nuit d'observation, l'opérateur doit donc alterner l'observation des sources d'intérêts avec l'observation de calibrateurs, le plus proche possible dans le ciel des sources, de façon à ne pas trop modifier le pointage et le réglage des lignes à retard. La bonne connaissance du calibrateur utilisé conditionne directement la fiabilité des mesures. Nous avons utilisé pour nos observations le catalogue de calibrateurs proposé par Mérand *et al.* (2005).

### 5.2.3 Une méthode de choix de configuration pupillaire

Le procédé de repositionnement des télescopes d'IOTA exige de remplir d'air les tuyaux sous vide, de déplacer les miroirs d'injection, de déplacer manuellement la structure des télescopes, puis de refaire le vide. Pour ces raisons, on se limite à une configuration par nuit d'observation, la reconfiguration de l'interféromètre se faisant de jour.
Il existe néanmoins 283 façons[2] de placer les télescopes pour chaque nuit. Une méthode de choix de configuration peut donc s'avérer utile…

#### 5.2.3.1 Une méthode d'optimisation de la FTO

Nous avons cherché à construire un critère de qualité de la configuration pupillaire. Pour cela, nous avons adapté au cas de l'interféromère IOTA la méthode de Mugnier *et al.* (1996) et Mugnier *et al.* (2004a) d'optimisation de configuration pupillaire d'un instrument imageur multi-pupilles. Cette méthode consiste à équilibrer la disposition des sous-pupilles de l'imageur, de façon à minimiser l'erreur quadratique moyenne (EQM) faite sur l'estimation de l'objet. Les auteurs montrent que cette EQM s'écrit :

$$EQM = \sum_{\boldsymbol{\nu}} \frac{\boldsymbol{S}_b(\boldsymbol{\nu})}{|FTO|^2(\boldsymbol{\nu}) + \frac{\boldsymbol{S}_b(\boldsymbol{\nu})}{\boldsymbol{S}_x(\boldsymbol{\nu})}}$$

Pour appliquer cette méthode au cas d'IOTA, il suffit de voir l'interféromètre comme un imageur, dont la pupille est constituée des trois miroirs collecteurs des télescopes A, B et C. Si A est en position A5 (à 5 mètres du coin sur le bras long), B en B5 (à 5 mètres du coin sur le bras court) et C en C0 (sur le coin), la pupille de cet instrument imageur virtuel correspond à celle de la figure 5.12. La FTO d'un tel instrument est l'autocorrelation de cette pupille. Comme on

[2] 12 positions pour le tél. A × 12 positions pour le tél. C × 6 positions pour le tél. B, dont on retire les configurations impossibles : A plus proche du coin que C, deux télescopes à moins de 5 mètres l'un de l'autre.

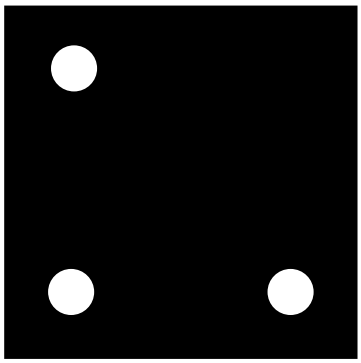


FIG. 5.12 – Pupille virtuelle de l'interféromètre IOTA en configuration A5-B5-C0.

considère plusieurs instants d'observation pour chaque nuit, et plusieurs nuits, il faut superposer les FTO obtenues pour chaque configuration. La FTO ainsi obtenue est la convolution de la couverture $u - v$ (consituée de Diracs) avec la FTO d'une seule pupille, qui a une forme de chapeau chinois.
Si on suppose que les DSP de l'objet $\boldsymbol{S}_x$ et du bruit $\boldsymbol{S}_b$ sont uniformes, l'EQM s'écrit :

$$EQM = \sum_{\boldsymbol{\nu}} \frac{\boldsymbol{S}_b}{|FTO|^2(\boldsymbol{\nu}) + \frac{\boldsymbol{S}_b}{\boldsymbol{S}_x}}$$

Ce critère pénalise les recouvrements entre les chapeaux chinois correspondant à chaque fréquence spatiale, c'est-à-dire qu'il pénalise les couvertures fréquentielles où les fréquences sont trop proches l'une de l'autre. On a donc bien un critère d'homogénéïté de la couverture fréquentielle. Ces recouvrements sont d'autant plus importants que les chapeaux chinois sont larges, c'est-à-dire que les diamètres des collecteurs des télescopes sont grands. Ceux d'IOTA mesurent 45 centimètres. Pour pénaliser plus fortement les configurations où deux fréquences spatiales mesurées sont proches, nous avons considéré des miroirs de 5 mètres de diamètre (au lieu de 45 centimètres), ce qui correspond à la distance minimale entre deux télescopes.

Il est possible de limiter la zone de fréquence d'intérêt en choisissant une DSP de l'objet prenant des valeurs très faibles au-delà d'une fréquence de coupure. Pour valider cette méthode, nous avons considéré une fréquence de coupure correspondant à une base de 20 mètres, avec une DSP de l'objet uniforme en dessous de cette fréquence.

#### 5.2.3.2 Résultats de simulation

Nous avons simulé un programme d'observation de 5 nuits, avec une configuration « classique » (que nous aurions intuitivement choisie pour une source dont seules les visibilités pour des bases inférieures à 20 mètres nous intéressent) et avec la configuration optimisée par notre méthode.
La figure 5.13 montre les FTO obtenues pour ces deux configurations. La zone plus sombre dans la FTO optimisée correspond aux fréquences au-delà de la fréquence de coupure considérée. On voit qu'à l'intérieur de la zone fréquentielle d'intérêt (cercle blanc en pointillés), la FTO est plus « riche » pour la configuration optimisée.

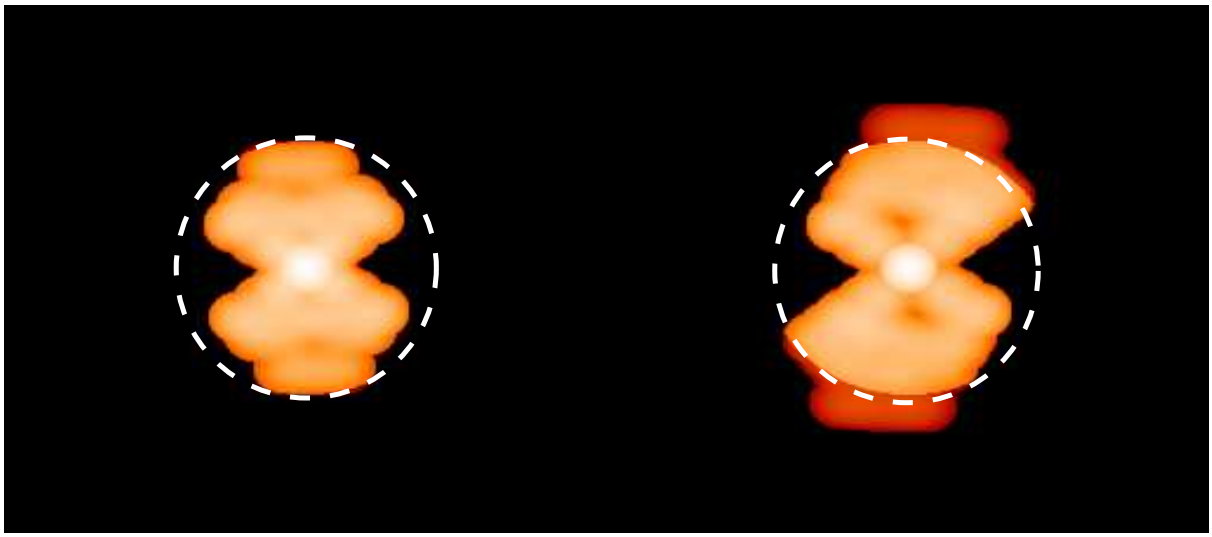


FIG. 5.13 – Comparaison des FTO obtenues pour les configurations pupillaires classique (à gauche) et optimisée (à droite).

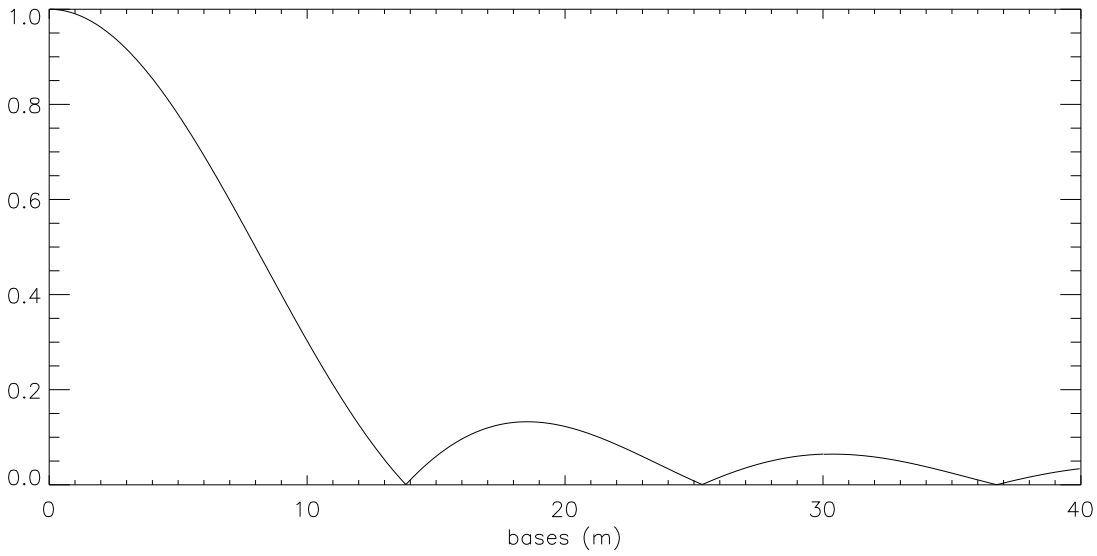



FIG. 5.14 – Visibilités en module, pour un disque uniforme de diamètre 30 mas, en bande H ($\lambda = 1,65$ microns).

#### 5.2.3.3 Un choix de configuration permettant la mesure de très faibles visibilités

En pratique, les conditions expérimentales n'étant pas idéales, beaucoup d'autres contraintes guident le choix de configuration de l'interféromètre, en particulier liées au RSB et au fait que l'on mesure des clôtures. Supposons une source circulaire d'un diamètre apparent de 30 millisecondes d'angle. La figure 5.14 montre le profil de visibilités obtenu en bande H (1,65 microns). Pour résoudre l'objet, il est nécessaire de couvrir des bases au delà du premier point d'annulation. Au maximum du deuxième lobe, les visibilités sont à 13%, et à moins de 7% au maximum du troisième lobe.
La difficulté pour les faibles visibilités consiste à repérer les franges. L'emplacement du paquet de frange dépend de la position des lignes à retard et de la direction de pointage. Les erreurs de modèle de pointage conduisent à balayer l'interférogramme sur plusieurs millimètres, par tranche de 20 microns environ. L'opérateur doit repérer les franges quand elles apparaissent. Un système de detection des franges existe, mais lorsque le contraste est très faible, l'œil humain est souvent plus efficace. Néanmoins, repérer des franges pour des visibilités inférieures à 5% devient très difficile.
Pour pouvoir néanmoins mesurer des franges de faible visibilité il est possible de choisir une configuration pupillaire telle que deux des trois bases correspondent à des visibilités hautes, et une seule à une visibilité basse. Une fois les franges repérées et centrées sur les deux premières bases (par exemple AB et BC), les franges sur la troisième base (AC) sont elles aussi sur l'ecran, même si on ne les distingue pas. En effet, si les franges sont centrées sur AB, c'est que la différence de marche $\mathrm{ddm}(AB)$ entre les faisceaux issus de A et de B est nulle. Si elles sont aussi centrées sur BC, alors on a :

$$\mathrm{ddm}(AC) = \mathrm{ddm}(AB) + \mathrm{ddm}(BC) = 0$$

Les franges sont donc aussi centrées sur AC.
Nous avons donc choisi les configurations pupillaires selon ce principe, et donc en fonction du profil de visibilité attendu sur les sources observées. La modélisation de la structure de $\chi$ du cygne, réalisée à partir des donnés recueillies par Perrin *et al.* (2004) en 2001 et 2002 (figure 5.15), nous a guidé dans le choix des premières configurations. Les données recueillies chaque nuit nous ont ensuite permis d'affiner au fur et à mesure notre modèle de profil de visibilité, et notamment de déterminer quelle configuration utiliser en fin de mission pour pouvoir former des franges avec une base de 38 mètres (nous nous attendions à une visibilité en module de l'ordre de $10\%$). Pour cette configuration, les visibilités en amplitudes prévues pour les deux autres bases, étaient environ $15\%$ et 0. En nous appuyant sur les deux bases pour lesquelles le contraste était suffisant, nous avons pu detecter des franges d'amplitude inférieure à $5\%$.

### 5.2.4 Étoiles de type Mira

Cette section rappelle les propriétés particulières des étoiles de type MIRA, et s'inspire largement de l'introduction de la thèse de doctorat de Verhoelst (2005).

#### 5.2.4.1 Types d'étoile considérés

Les étoiles sont en général classées en plusieurs groupes, correspondant à des régions du diagramme Hertzsprung-Russel, ou diagramme HR. Ce diagramme présente la répartition de l'ensemble d'étoiles considéré dans le plan Température-Luminosité. Les zones d'accumulation de ce diagramme, ou séquences, correspondent aux différents stades d'evolution des étoiles. Ce diagramme est particulièrement intéressant lorsqu'il concerne un ensemble d'étoiles

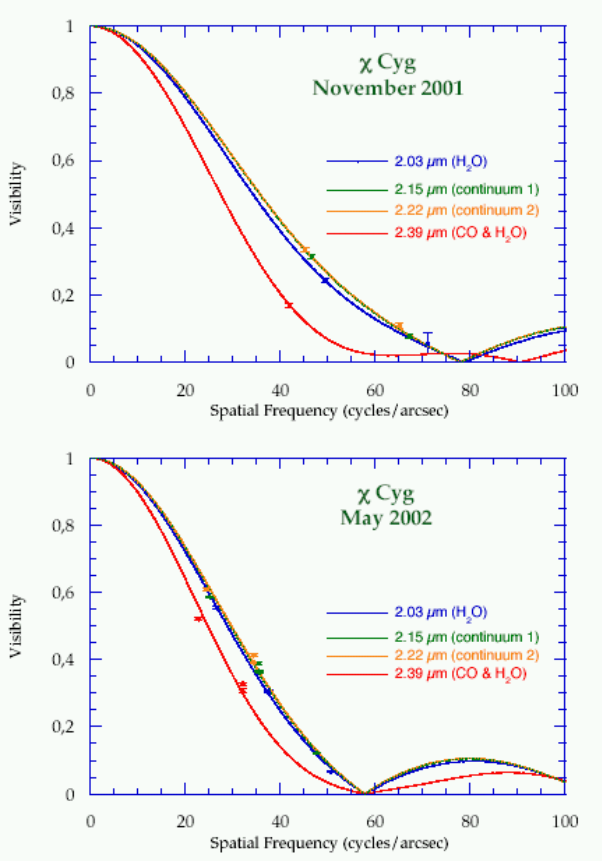



FIG. 5.15 – Modélisation de la structure de $\chi$ du cygne [Perrin *et al.* (2004)]

de même origine (tel qu'un amas globulaire), donc de même âge, ne se distinguant que par leur masse. La figure 5.16 présente le diagramme HR de l'amas globulaire M3.

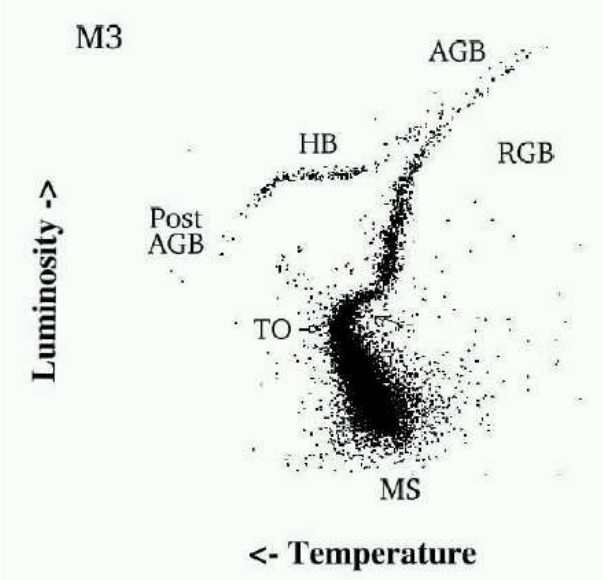



FIG. 5.16 – Diagramme HR de l'amas globulaire M3 [Renzini (1988)]

**Séquence Principale** Pendant 9 dixièmes de leur vie, les étoiles sont dans la Séquence Principale (Main Sequence ou MS). Cette zone correspond à une phase de combustion stable de l'hydrogène du noyau stellaire, transfomé en hélium. C'est la phase dans laquelle se trouve le Soleil. Cette phase dure moins longtemps pour les étoiles de grande masse, soumises à une forte gravité qui doit être équilibrée par une forte activité nucléaire. Un fois l'hydrogène du noyau épuisé, la gravité comprime l'étoile, provoquant un échauffement.

**Naines rouges et géantes rouges** Les étoiles de moins de 0,3 masse solaire, ou naines rouges, continuent à se refroidir et se comprimer jusqu'à devenir des naines blanches.
Pour les étoiles plus massives, le noyau d'hélium échauffe suffisament une fine couche d'hydrogène de l'enveloppe pour enclencher sa combustion. La luminosité augmente jusqu'à 2500 fois celle du Soleil, et l'étoile est en forte expansion. Elle « remonte » la branche des géantes rouges (Red Giant Branch ou RGB). Selon leur taille, on parle de géante rouge, comme Aldébaran, ou de supergéante rouge, comme Betelgeuse ($\alpha$ Orionis). Cette phase d'expansion et de combustion en couche s'accompagne de mouvements convectifs importants.

**Flash d'hélium et branche horizontale (Horizontal Branch ou HB)** Lorsque la combustion de la couche d'hydrogène porte le noyau d'hélium à environ 120 millions de degrés, la combustion de l'hélium en carbone et en oxygène (réaction « triple alpha »), devient possible. Pour les étoiles plus légères que 2 masses solaires environ, cette combustion s'acompagne d'un « Flash d'hélium ».
L'étoile s'échauffe à luminosité constante, se déplaçant vers la gauche, le long de la branche horizontale, à mesure que la combustion de l'hélium s'amplifie. Une fois passé un stade optimal, cette réaction se ralentit à mesure que les réserves d'hélium du noyau s'épuisent. L'étoile parcourt la branche horizontale dans l'autre sens, en se refroidissant et en se comprimant.

**Branche asymptotique des géantes (Asymptotic Giant Branch ou AGB)** Une fois le noyau d'hélium totalement converti en carbone et en oxygène, le noyau se contracte et se réchauffe, provoquant une combustion en couche similaire à celle intervenant à la fin de la combustion de l'hydrogène. Cependant, cette fois, la couche en combustion est une couche d'hélium. De nouveau, l'étoile entre en forte **expansion** en se rechauffant, dans une zone du diagramme HR proche de la branche des géantes rouges (RGB), appelée par conséquent branche asymptotique.

**Étoiles Mira** Cette phase d'expansion est accompagnée de forts phénomènes de convection, dûs à la combustion de la couche d'hélium proche du noyau, mais aussi d'une couche supérieure d'hydrogène. Cette structuration crée de grandes variations de luminosité de l'étoile, variations régulières et périodiques pour les étoiles de type Mira (dont l'étoile Mira est elle-même un exemple). Cette étape est aussi accompagnée d'une forte éjection de matière, créant une enveloppe de poussière autour de l'étoile. Dans le cas d'étoiles Mira, cette enveloppe circumstellaire peut masquer complètement le noyau de l'étoile. Les étoiles AGB contribuent à 75% de la perte de masse dans la galaxie, et contribuent à l'enrichissement du milieu interstellaire. Les molécules de poussière du milieu interstallaire sont la matière première notamment des planètes de type terrestre. Ces étoiles sont fondamentales dans l'évolution des galaxies, et constituent des cibles d'observation privilégiées.

**Étape post-AGB** L'ejection de matière se poursuit jusqu'à ce qu'il ne reste que le noyau de l'étoile. Pour les étoiles de moins de 8 masses solaires, il ne reste plus assez de matière superficielle pour relancer une réaction nucléaire. Le noyau se refroidit (le long de la branche post-AGB), jusqu'à devenir une naine blanche.

### 5.2.4.2 Objectifs astrophysiques de la mission

Perrin *et al.* (2004) proposent une modélisation laminaire des étoiles Mira, constitué d'une photosphère, d'une couche de vide, et d'une fine couche moléculaire (Figure 5.17). Cette enveloppe rayonne dans le proche IR, et peut donc être observée sur l'interféromètre IOTA.
Le but de cette mission a donc été l'imagerie en bande K d'étoiles de type MIRA, dont $\chi$ Cygni fait partie, et en particulier la mise en évidence d'éventuelles dissymétries dans la structure de la couche moléculaire.

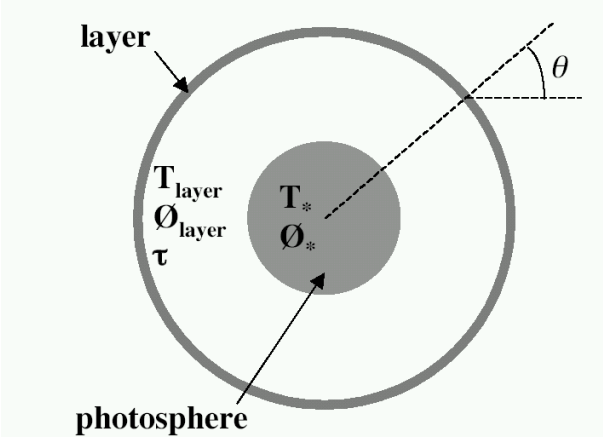



FIG. 5.17 – Modèle laminaire d'étoile MIRA [Perrin *et al.* (2004)]

## 5.3 Reconstruction de $\chi$ du Cygne

Une deuxième mission a été organisée en Mai 2005. Nous n'étions que deux sur le site (Sylvestre Lacour et moi-même), et avons pu apprécier très directement la difficulté d'obtenir des données interférométriques en nombre suffisant pour qu'une reconstruction d'image soit possible. Cette section présente le traitement des données expérimentales concernant l'étoile $\chi$ du Cygne.

### 5.3.1 Réduction et première reconstruction des données

#### 5.3.1.1 Couverture du plan $(u, v)$

Nous avons observé l'étoile $\chi$ Cygni en bande large H ($1{,}65$ microns $\pm 175nm$) durant quatre nuits. La couverture fréquentielle obtenue est présentée figure 5.18. Les mesures sur la

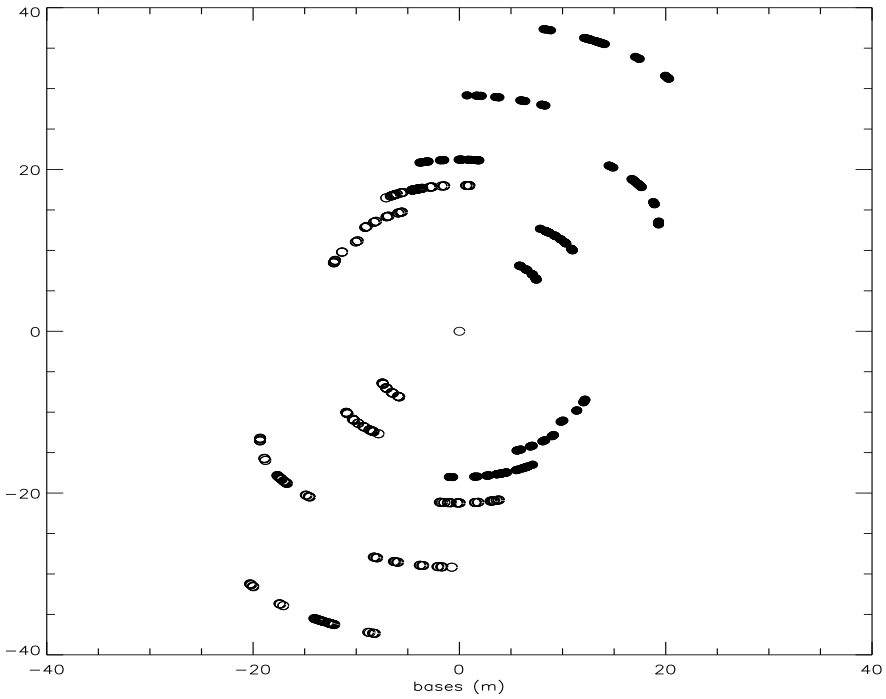



FIG. 5.18 – $\chi$ Cygni. Couverture $u - v$, IOTA mai 2005.

sources ont été encadrées de mesures sur les calibrateurs HD180450 et HD176670.

#### 5.3.1.2 Dissymétrie des sources observées

La figure 5.19 montre le profil de visibilités que nous avons obtenu. Les séries verticales de points sont dues en partie au bruit, mais traduisent aussi une dissymétrie de la source. Ces points correspondent à des fréquences de même module, mais de directions différentes. Cette dissymétrie est confirmée par l'étude des clôtures obtenues. La TF d'un objet à symétrie centrale est réelle, sa phase est donc $\pi$ ou $0$ en tout point. La clôture de phase étant la somme de trois phases modulo $2\pi$, elle est aussi de $\pi$ ou $0$ pour un objet à symétrie centrale. Nous avons obtenu des clôtures très différentes de $\pi$ ou 0 (Cf. figure 5.20), qui prouvent la dissymetrie de l'étoile $\chi$ Cygni.

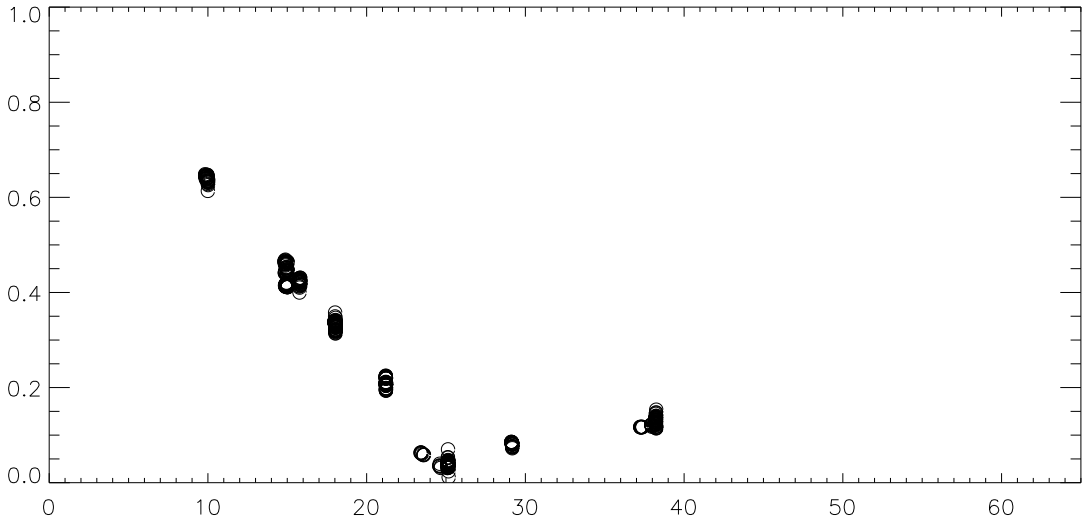



FIG. 5.19 – $\chi$ Cygni. Modules de visibilités, bande H, IOTA mai 2005.

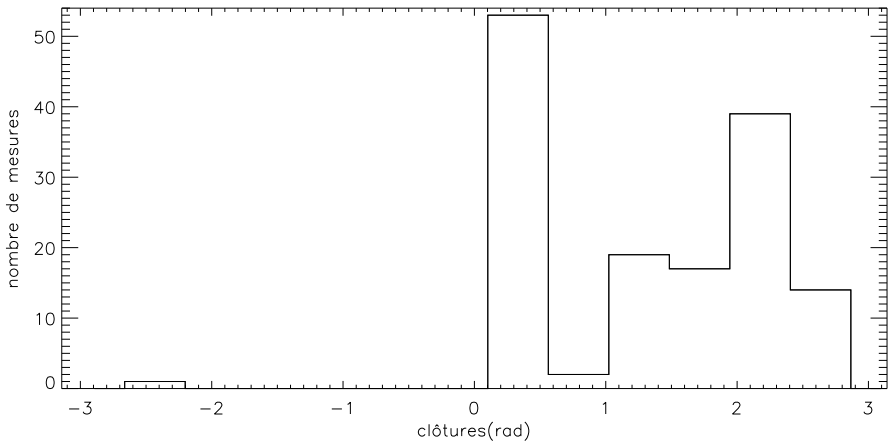



FIG. 5.20 – Histogramme de répartition des clôtures mesurées.

#### 5.3.1.3 Reconstruction paramétrique

Puisque nous disposions d'un modèle à deux couches d'étoile de type Mira (cf. section 5.2.4.2), nous avons procédé à un ajustement des paramètres de ce modèle sur les données mesurées (nous n'avons exploité que les données de visibilités carrées pour cette reconstruction). Nous obtenons une photosphère de diamètre $16mas$ et une couche moléculaire de diamètre $43mas$. La figure 5.21 montre l'adéquation entre les visibilités mesurées et celles correspondant au modèle.

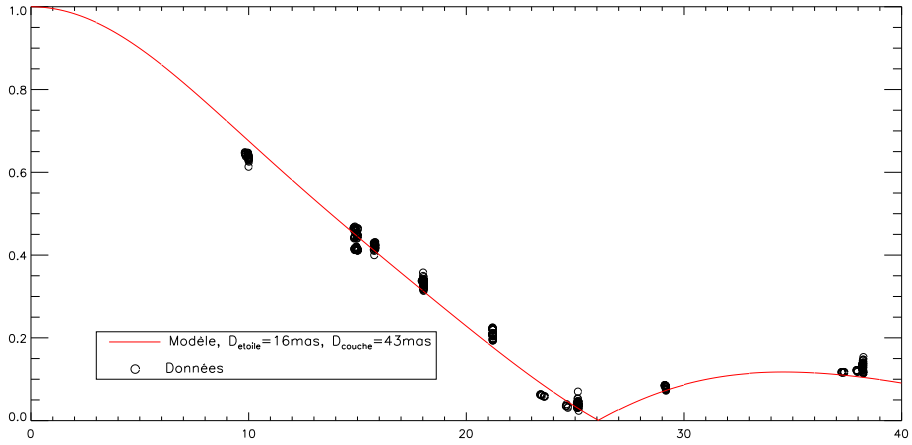



FIG. 5.21 – Reconstruction paramétrique de $\chi$ Cygni.

### 5.3.2 Sources d'erreurs et adaptation de l'algorithme

#### 5.3.2.1 Effets de bande large

L'algorithme WISARD ne tient pas compte de la bande spectrale dans laquelle sont effectuées les mesures. Or, le fait que plusieurs longueurs d'onde se superposent dans les interférogrammes conduit à des différences de mesure de visibilités carrées. En effet, la mesure est faite sur la TF de l'interférogramme. Ce dernier est le produit d'un sinus cardinal et d'un cosinus (Cf. section 1.1.5.2 page 20). La TF de l'interférogramme est donc un pic convolué par une fenêtre. Le calcul de la visibilité carrée est déduit de l'intégrale du carré de la TF de l'interférogramme. La mesure obtenue est donc une somme de visibilités carrées pour chaque longueur d'onde transmise par le filtre (voir [Perrin (2003)] pour plus d'informations sur les méthodes de calibration des données interférométriques).
La figure 5.22 page 137 simule les visibilités carrées que l'on obtiendrait pour un disque uniforme de diamètre 16 millisecondes d'angle et de spectre plat, en tenant compte de la largeur du filtre. Les erreurs sur les visibilités carrées dues à la non prise en compte de la largeur du filtre sont au maximum de 0,012. Ces erreurs peuvent être supérieures à quelques écarts-type, et donc affecter la qualité des reconstructions.

Nous avons donc réestimé les écarts-type sur les visibilités carrées, en leur ajoutant systématiquement 0,012. Une prise en compte plus précise de ces effets dans l'algorithme WISARD reste à faire.

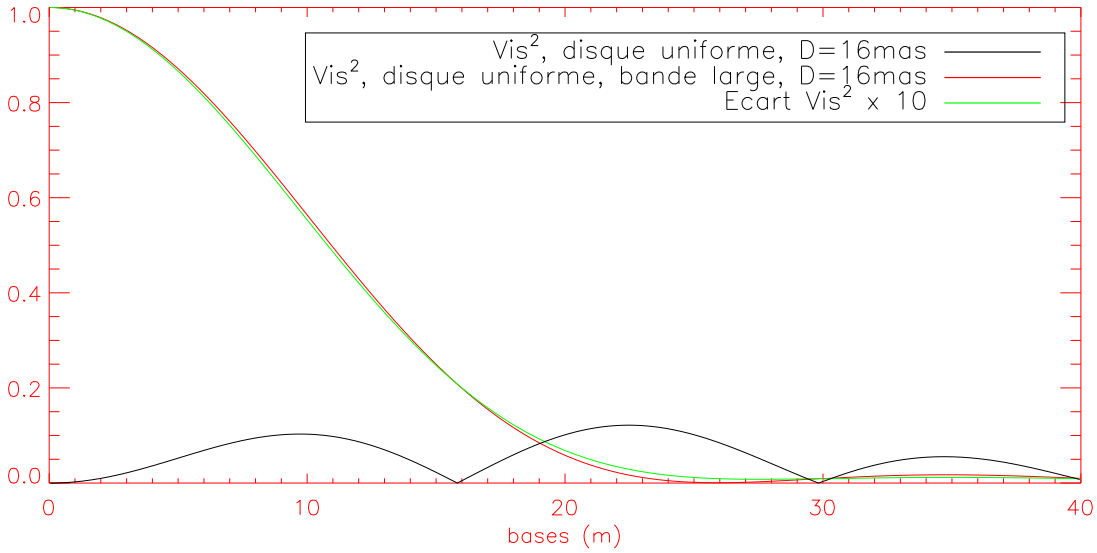



FIG. 5.22 – Erreur sur les visibilités carrées d'un disque uniforme de diamètre 16 mas, avec un filtre en bande H de largeur 350 nm.

#### 5.3.2.2 Effets parasites dûs à l'utilisation de densités

Il a été nécessaire d'utiliser des filtres à densité pour ne pas saturer les détecteurs. Étant donnée la différence de magnitude importante entre les calibrateurs utilisés et les sources, ces filtres n'ont été utilisés que pour les sources, et pas pour les calibrateurs. Nous avons néanmoins pu observer l'effet des densités sur les mesures, en observant la même source avec et sans densité. Les différences se traduisent par un écart en visibilité carrée de quelques pourcents. Pour l'instant, nous n'avons pas proposé de méthode de prise en compte de ces effets.

#### 5.3.2.3 Contraindre la reconstruction

Les premières reconstructions que nous avons obtenues présentaient des variations importantes en fonction de l'initialisation choisie, ou des paramètres de régularisation. Nous avons interprété cette variabilité comme liée à de nombreux minima locaux du critère, ou à des erreurs de modélisation du bruit. Nous avons alors pris le parti d'utiliser un *a priori* plus informatif que ceux utilisés en simulations. En particulier, le profil des mesures de $\chi$ du Cygne est proche de celui d'un disque uniforme de 16 millisecondes d'angle de diamètre, comme le montre la figure 5.23. Cette information *a priori* peut être prise en compte en choisissant comme objet de rappel de la régularisation un tel disque uniforme. L'objet reconstruit présentera alors des divergences mineures (en terme de reconstruction) par rapport à cette première approximation.

### 5.3.3 Reconstruction de $\chi$ du Cygne avec WISARD

Nous avons donc effectué une reconstruction de $\chi$ du Cygne, avec régularisation L2 avec pour objet *a priori* un disque uniforme de diamètre 16 mas, et en prenant en compte les effets dûs à la largeur du filtre utilisé. La figure 5.24 présente le profil des visibilités en module estimées et mesurées. On voit que l'écart entre mesure et estimation est plus grand pour des visibilités proches de 0, ce qui est dû à la prise en compte de la largeur du filtre. Les données reconstruites présentent une annulation, ce qui n'est pas le cas des visibilités mesurées.

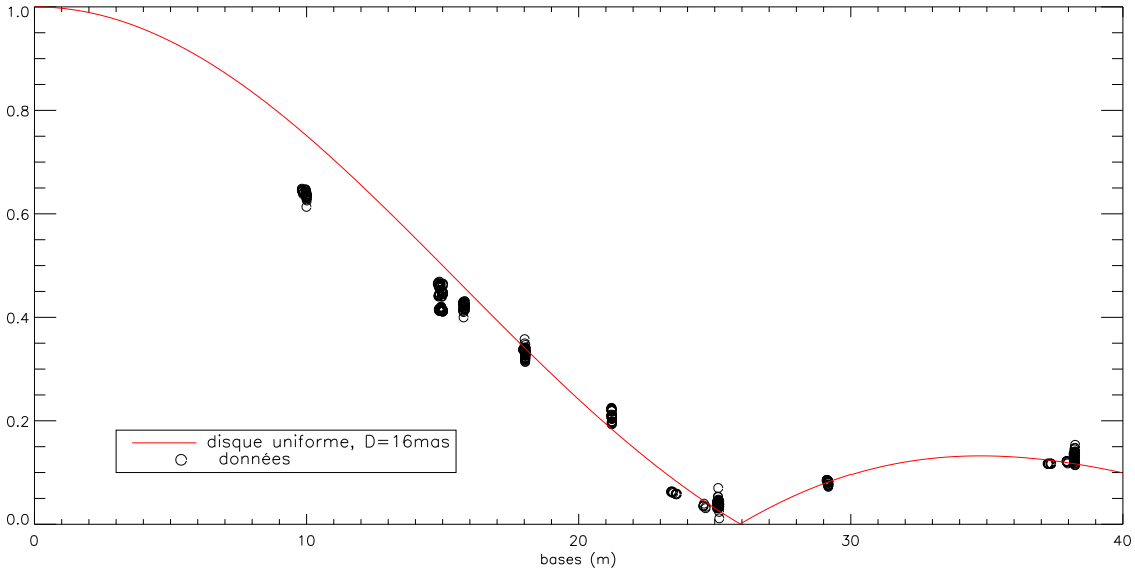



FIG. 5.23 – Comparaison des mesures de visibilités en module de $\chi$ du Cygne avec un modèle de disque uniforme de diamètre 16 mas, bande H.

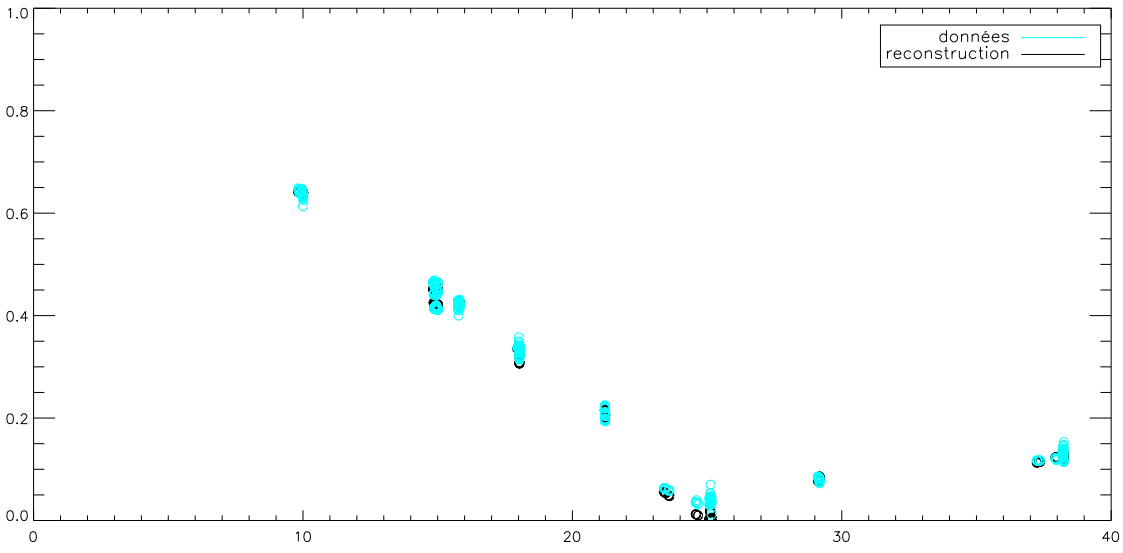



FIG. 5.24 – Visibilités mesurées et visibilités estimées pour l'étoile $\chi$ du Cygne.

L'image obtenue est présentée figure 5.25. Elle est constituée d'un disque uniforme correspondant à l'*a priori* utilisé, et d'une tache brillante concentrant près de 13% du flux total de l'étoile. La couche moléculaire rayonnant à peu près 10 fois moins que la photosphère à 1,65 microns, une première interprétation consiste à supposer une inhomogénéité dans la couche moléculaire, c'est-à-dire une zone où elle est plus fine et transmet une fraction plus importante du flux émis par la photosphère.

Ce résultat n'est qu'une première tentative, qui doit être notamment confronté à des modèles astrophysiques avant d'être validé. De plus, les mesures recueillies ne permettent de reconstruire une image que de quelques pixels. Les images sur-échantillonnées sont généralement plus agréables à l'oeil. Nous présentons néanmoins dans la figure 5.25 l'image pixélisée, c'est-à-dire avant sur-échantillonnage, qui traduit plus fidèlement le degré de résolution de la reconstruction. En particulier, il serait impossible d'estimer précisément la position ou le diamètre de cette tache brillante.

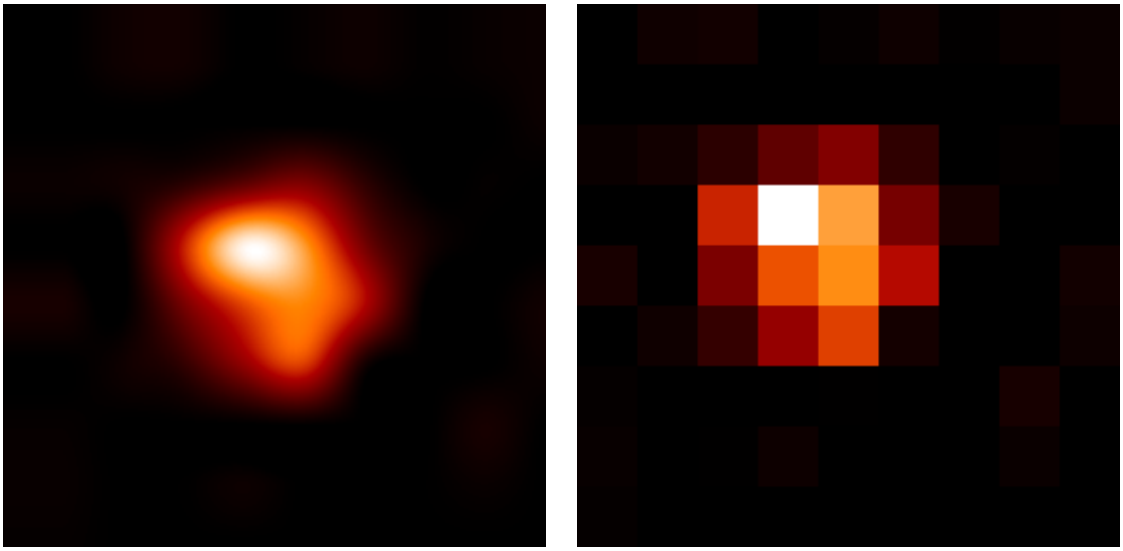


image sur-échantillonnée image pixélisée

FIG. 5.25 – $\chi$ du Cygne. Champ reconstruit :40mas x 40 mas. Reconstructions originale et sur-échantillonnée. La tache brillante concentre 13% du flux. Regularisation : rappel quadratique vers un disque uniforme de diamètre 16 mas

## 5.4 Conclusion

La première reconstruction d'images à partir de données expérimentales est généralement un passage difficile mais instructif. Dans le cas de WISARD, nous avons pu constater que la qualité de notre reconstruction s'est améliorée à la suite de l'utilisation d'une modélisation paramétrique de l'objet d'intérêt (en l'occurence un modèle simpliste de disque uniforme). Le cadre bayésien de notre méthode a permis une implantation satisfaisante de cet *a priori*,sous forme d'objet de rappel pour une régularisation quadratique. L'aller-retour entre d'une part l'estimation de paramètres de modèles astrophysiques de plus en plus complexes, et la reconstruction d'image d'autre part, devrait permettre d'affiner encore la reconstruction.

# Chapitre 6

# Conclusion

## 6.1 Objectifs et résultats

L'étude présentée dans ce manuscrit a consisté à concevoir une méthode de reconstruction d'image :

- qui utilise un modèle de bruit sur les données correspondant à celui effectivement constaté sur des interféromètres optiques ;
- qui prenne en entrée des visibilités carrées ou en module, et des clôtures de phase ;
- qui soit efficace sur les données expérimentales issues d'interféromètres optiques en fonctionnement comprenant au moins trois télescopes.

A l'issue de ce travail, voici les résultats que nous avons obtenus :

**RIME : un algorithme efficace dans le cadre de l'interférométrie optique à turbulence faible ou corrigée.** Les interféromètres optiques spatiaux ou à référence de phase fourniront des phases et modules de visibilités complexes bruités pour quelques fréquences spatiales. Nous avons proposé l'algorithme RIME, qui s'appuie sur une approximation quadratique convexe du modèle de bruit réel, et que nous avons validé sur simulations et sur données expérimentales. Cet algorithme a été développé dans un cadre bayésien et permet donc l'utilisation d'une grande variété de régularisations (chapitre 2).

**Une modélisation myope des effets turbulents.** En présence de turbulence, les observables en interférométrie optique sont les clôtures de phase et les visibilités carrées pour quelques fréquences spatiales. Après une étude détaillée de l'indétermination en phase, où nous avons explicité certaines dégénérescences du problème, nous avons proposé un modèle de pseudo-données de type visibilité complexe, où l'incomplétude des informations de phase apparaît explicitement sous forme d'aberrations myopes du système (chapitre 3).

**WISARD : une méthode originale de reconstruction d'images en optique par auto-calibration** Après avoir formé le critère d'attache aux données déduit de notre modèle myope et montré son équivalence avec la log-vraisemblance issue du modèle de bruit usuel, nous avons construit une méthode d'estimation conjointe des aberrations et de l'objet par auto-calibration (chapitre 4), pour laquelle :

- l'étape d'optimisation en objet à aberrations fixées bénéficie des apports du chapitre 2, c'est-à-dire que la dépendance en l'objet du critère d'attache aux données est quadratique ;
- l'étape d'optimisation en aberrations à objet fixé est globale dans le cas le plus difficile, c'est-à-dire pour des interféromètres à 3 télescopes.

**Validation sur données simulées et expérimentales** L'algorithme WISARD ainsi obtenu a été validé sur des simulations, au cours d'un concours international de reconstruction d'image en aveugle (cf. section 5.1 page 115). Il était alors au début de son développement et présentait des erreurs de modélisation, mais a pourtant été classé deuxième à ce concours.

Nous avons effectué deux missions sur l'interféromètre optique à 3 télescopes IOTA (chapitre 5), et la deuxième nous a permis de recueillir des données expérimentales sur les étoiles $\chi$ du Cygne et $\alpha$ du Bouvier. La version actuelle de WISARD a permis de reconstruire une image de l'étoile $\chi$ du Cygne, à une résolution d'environ 5 millisecondes d'angle, à partir de ces dernières données. Cette reconstruction permet d'identifier une tache brillante sur la surface de l'étoile. Une interprétation astrophysique de ce résultat est en cours.

## 6.2 Discussion

Un des enjeux majeurs dans la construction d'une méthode de traitement de données consiste à l'inscrire dans l'évolution des instruments auxquels elle s'applique. Il est en effet difficile de prévoir l'état des développements instrumentaux dans 10 ans, et l'opportunité à cette date des techniques que nous avons développées. Néanmoins, plusieurs arguments nous permettent d'être positifs quand à l'utilisation des algorithmes présentés dans ce manuscrit.

Nos travaux permettent en effet de répondre aux différentes échéances prévisibles dans l'évolution des interféromètres optiques :

1. Dès à présent, nous sommes capables de traiter avec succès des données expérimentales issues d'interféromètres à 3 télescopes ;
2. à moyen terme, on disposera de données issues d'interféromètres à plus de 3 ou 4 télescopes : nous avons montré lors de l'Imaging Beauty Contest que les performances atteintes par WISARD, avec des données issues d'un interféromètre à 6 télescopes dont la base maximale est de 66 mètres, sont comparables à celles d'un télescope monolithique de 100 mètres de diamètre ;
3. à plus long terme, les systèmes au sol à référence de phase (ou les interféromètres spatiaux) pourront fournir des données en visibilités complexes, qui constituent les observables d'entrée de l'algorithme RIME.

Ces extrapolations supposent que l'interférométrie optique ne soit pas abandonnée au profit de l'utilisation exclusive de grands télescopes. À cet egard, le développement de projets interférométriques à très longues bases, notamment d'interféromètres polaires, laisse à penser que l'interférométrie aura longtemps quelques hectomètres d'avance sur les plus grands télescopes monolithiques, tels que le futur Thirty Meter Telescope américain, ou le projet de télescope européen géant OWL.

Il est apparu, notamment dans le cadre du concours de traitement d'image organisé par Peter Lawson [Lawson *et al.* (2004)], que la reconstruction d'images à partir de clôtures de phase et de visibilités carrées obtenues sur un interféromètre à 6 télescopes est possible, dans des conditions assez larges sur le choix de la méthode et de l'algorithmie. Toutefois, un choix judicieux d'*a priori* ($\mathrm{L1L2}_w$, décrit section 1.4.2.4) permet d'exploiter au maximum les données, y compris haute fréquence, comme le montre la comparaison avec un grand télescope monolithique. Cet exemple montre que, contrairement à ce qui est fait dans les algorithmes CLEAN ou WIPE, où l'on a tendance à limiter le support fréquentiel de la reconstruction (par exemple, dans CLEAN, en convoluant la carte de Diracs reconstruite par le « Clean Beam »),

il ne faut pas se « censurer » sur la résolution de la reconstruction. Pourvu que l'*a priori* utilisé soit judicieux, une extrapolation spectrale partielle est possible.

Bien entendu, le problème devient beaucoup plus ardu lorsque l'interféromètre compte seulement 3 ou 4 télescopes, ce qui est majoritairement le cas aujourd'hui. Nous avons proposé des contributions originales sur ce sujet, et traité des données expérimentales (ce qui n'est pas le cas, à notre connaissance, pour toutes les méthodes de traitement de données interférométriques optiques déjà publiées). Les enseignements tirés de cette expérience nous conduisent à choisir les axes d'amélioration de WISARD suivants :

- D'un point de vue méthodologique, la fiabilité de WISARD pourrait être validée sur plus d'exemples test, notamment dans le cas 3 télescopes – par exemple au cours du prochain Imaging Beauty Contest, organisé en 2006 ;
- Parallèlement, le caractère multimodal de l'estimation de l'objet d'intérêt nous a poussés à contraindre le problème avec des *a priori* suffisamment bien choisis pour atteindre une bonne solution. A ce titre, l'utilisation d'un bi-modèle, c'est-à-dire la reconstruction simultanée d'une carte de points sources et d'une carte d'objet étendu, nous semble une solution adaptée. Des méthodes d'optimisation globale sont certainement à envisager. Nous les avons écartées pour des raisons de temps de convergence, mais cette lenteur relative devrait être pondérée par l'amélioration éventuelle que de telles méthodes apporteraient.

Pour choisir telle ou telle direction dans la création de méthodes nouvelles, ou l'amélioration de l'existant, il faut définir ce qu'en seraient les objectifs. Il me semble qu'un « bon » algorithme de traitement de données interférométriques doit avant tout être stable et robuste, c'est-à-dire qu'il doit pouvoir fournir une reconstruction fiable de l'objet sans être trop sensible à un bruit important sur les données ou à une erreur de modélisation. La précision photométrique de la reconstruction est secondaire, car une reconstruction correcte permet de modéliser approximativement la source. Ce modèle peut alors être injecté dans une nouvelle reconstruction, soit de façon stricte – on n'estime que les paramètres du modèle – soit sous forme d'un objet de rappel dans un terme de régularisation – cette deuxième solution a l'avantage d'autoriser des divergences par rapport au modèle. Un tel aller-retour entre reconstruction et modélisation astrophysique de la source me semble la meilleure façon d'obtenir une reconstruction tenant compte des *a priori* complémentaires proposés d'une part par le traiteur de signal, d'autre part par l'astrophysicien.

## 6.3 Perspectives

Dans notre obstination à atteindre l'objectif principal de ces travaux, c'est-à-dire traiter des données que nous aurions recueillies avec une méthode que nous aurions développée, certains points d'intérêt ont été laissés de côté.
Parmi ceux-ci, les ambiguïtés entières, évoquées section 3.4.2, mériteraient une étude approfondie. L'implantation de la contrainte de positivité par « slack variables » et l'utilisation d'un bi-modèle, toutes deux inspirées des travaux de Giovannelli & Coulais (2005), me semblent pouvoir améliorer notablement les performances de l'algorithme.
La comparaison de la méthode WISARD avec ses concurrentes, bien que partiellement réalisée lors de l'Imaging Beauty Contest, est encore incomplète. Il est pour l'instant difficile de préciser les domaines d'excellence de chaque méthode. L'étude croisée des mécanismes d'auto-calibration dans les algorithmes MIRA et WISARD est en cours. Ces deux méthodes seront délivrées prochainement à la Communauté Européenne, dans le cadre de l'Eropean In-

terferometry Initiative (EII).
La perspective la plus proche consiste à traiter les données que nous avons recueillies lors de la seconde mission sur IOTA, correspondant à l'étoile $\alpha$ du Bouvier. Par ailleurs, une collaboration est esquissée avec l'observatoire de Grenoble, pour traiter d'autres jeux de données expérimentales.
Une perspective plus lointaine, mais qui me semble de tout premier plan, est l'inclusion d'images obtenues sur un instrument monolithique corrigé par optique adaptative dans la reconstruction de données interférométriques. Ce couplage entre les deux techniques d'imagerie optique haute résolution les plus performantes est assurément prometteur.

# Bibliographie


Anterrieu, E., 1998. Advanced Imaging Methods : WIPE. In Guilloteau (1998b), chapter 14, pp. 147–158.

Berger, J. P., Haguenauer, P., Kern, P., Perraut, K., Malbet, F., Schanen, I., Severi, M., Millan-Gabet, R., & Traub, W., 2001. Integrated optics for astronomical interferometry. IV. First measurements of stars **376**, L31–L34.

Berger, J.-P., *et al.*, 2003. An integrated-optics 3-way beam combiner for IOTA. In *Interferometry for Optical Astronomy II. Edited by Wesley A. Traub . Proceedings of the SPIE, Volume 4838, pp. 1099-1106 (2003).*, pp. 1099–1106.

Blanc, A., 2002. *Identification de réponse impulsionnelle et restauration d'images : apports de la diversité de phase*. Ph.D. thesis, Université Paris XI Orsay.

Blanc, A., Mugnier, L. M., & Idier, J., 2003. Marginal estimation of aberrations and image restoration by use of phase diversity. *J. Opt. Soc. Am. A* **20** 6, 1035–1045.

Bouman, C. & Sauer, K., 1993. A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation. *IEEE Trans. Image Processing* **2** 3, 296–310.

Buscher, D. F., 1994. Direct Maximum-Entropy Image Reconstruction from the Bispectrum. In *Very High Angular Resolution Imaging : proceedings of the 158th Symposium of the International Astronomical Union held at the Women's College, University of Sydney, Australia, 11-15 January 1993*.

Cassaing, F., 1997. *Analyse d'un instrument à synthèse d'ouverture optique : méthodes de cophasage et imagerie à haute résolution angulaire*. Ph.D. thesis, Université Paris XI Orsay.

Conan, J.-M., Mugnier, L. M., Fusco, T., Michau, V., & Rousset, G., 1998. Myopic Deconvolution of Adaptive Optics Images using Object and Point Spread Function Power Spectra. *Appl. Opt.* **37** 21, 4614–4622.

Cornwell, T. J. & Wilkinson, P. N., 1981. A new method for making maps with unstable radio interferometers. *Mon. Not. R. Astr. Soc.* **196**, 1067–1086.

Delage, L., Reynaud, F., & Lannes, A., 2000. Laboratory Imaging Stellar Interferometer with Fiber Links. *Appl. Opt.* **39**, 6406–6420.

Delage, L., Reynaud, F., & Thiébaut, E., 1999. Imaging laboratory test on a fiber linked telescope array. *Opt. Commun.* **160**, 27–32.

Demoment, G., 1989. Image Reconstruction and Restoration : Overview of Common Estimation Structures and Problems. *IEEE Trans. Acoust. Speech Signal Process.* **37** 12, 2024–2036.

Demoment, G. & Goussard, Y., 2001. Problèmes inverses en imagerie optique à travers la turbulence. In *Approche bayésienne pour les problèmes inverses* (J. Idier, ed.), chapter 3, pp. 59–78, Hermès, Paris.

Dyck, H. M., 1999. Interferometry with two telescopes. In *Principles of Long Baseline Stellar Interferometry* (P. R. Lawson, ed.), chapter 12, pp. 185–202, Jet Propulsion Laboratory.

Faucherre, M., Merkle, F., & Vakili, F., 1989. Beam combination in aperture synthesis from space : field of view limitations and (u,v) plane coverage optimisation. In *New Technologies for Astronomy* (J.-P. Swings, ed.), volume 1130 of *Proc. Soc. Photo-Opt. Instrum. Eng.*, pp. 138–145.

Fienup, J. R., 1982. Phase retrieval algorithms : a comparison. *Appl. Opt.* **21** 15, 2758–2769.

Garcia, P. J. V., Cabrit, S., Ferreira, J., & Binette, L., 2001. Atomic T Tauri disk winds heated by ambipolar diffusion. II. Observational tests. *Astron. Astrophys.* **377**, 609–616.

Gerchberg, R. W. & Saxton, W. O., 1972. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35**, 237–246.

Gilbert, J. C., 2002. *Optimisation Différentiable : Théorie et Algorithmes*. INRIA.

Giovannelli, J.-F. & Coulais, A., 2005. Positive deconvolution for superimposed extended source and point sources. *Astron. Astrophys.* **439**, 401–412.

Giovannelli, J.-F. & Idier, J., 2001. Bayesian interpretation of periodograms. *IEEE Trans. Signal Processing* **49** 7, 1988–1996.

Golay, M. J. E., 1971. Point Arrays Having Compact, Nonredundant Autocorrelations. *J. Opt. Soc. Am.* **61**, 272–273.

Goodman, J. W., 1968. *Introduction to Fourier optics*. McGraw-Hill.

Green, P. J., 1990. Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm. *IEEE Trans. Med. Imag.* **9**, 84–93.

Guilloteau, S., 1998a. The Imaging Principles. In Guilloteau (1998b), chapter 13, pp. 137–146.

Guilloteau, S. (ed.), 1998b. *Proceedings of the IRAM Millimeter Interferometry Summer School*, Institut de Radio Astronomie Millimétrique, Saint Martin d'Hères.

Haniff, C. A., Mackay, C. D., Titterington, D. J., Sivia, D., Baldwin, J. E., & Warner, P. J., 1987. The first images from optical aperture synthesis. *Nature (London)* **328**, 694–696.

Hogbom, J., 1974. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. Astrophys. Suppl. Ser.* **15**, 417–426.

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

Idier, J., Mugnier, L., & Blanc, A., à paraître 2005. Statistical behavior of joint least square estimation in the phase diversity context. *IEEE Trans. Image Processing* .

Jennison, R. C., 1958. A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distribution of small angular extent. *Mon. Not. R. Astr. Soc.* **118**, 276–284.

Jennison, R. C., 1961. The Michelson stellar interferometer : a phase sensitive variation of the optical instrument. *Proc. Phys. Soc.* **78**, 596–599.

Lannes, A., 1990. Remarkable algebraic structures of phase-closure imaging and their algorithmic implications in aperture synthesis. *J. Opt. Soc. Am. A* **7**, 500–512.

Lannes, A., 1998a. Weak-phase imaging in optical interferometry. *J. Opt. Soc. Am. A* **15** 4, 811–824.

Lannes, A., 1998b. Weak-phase imaging in optical interferometry. *J. Opt. Soc. Am. A* **15** 4, 811–824.

Lannes, A., 2001. Integer ambiguity resolution in phase closure imaging. *Optical Society of America Journal A* **18**, 1046–1055.

Lannes, A., Anterrieu, E., & Bouyoucef, K., 1994. Fourier Interpolation and Reconstruction via Shannon-Type Techniques Part I : Regularization Principle. *J. Mod. Opt.* **41**, 1537–1574.

Lannes, A., Anterrieu, E., & Bouyoucef, K., 1996. Fourier Interpolation and Reconstruction via Shannon-Type Techniques Part II : Technical developpements and applications. *J. Mod. Opt.* **43**, 105–138.

Lannes, A., Anterrieu, E., & Maréchal, P., 1997. Clean and Wipe. *Astron. Astrophys. Suppl. Ser.* **123**, 183–198.

Lannes, A., Roques, S., & Casanove, M. J., 1987. Resolution and robustness in image processing : a new regularization principle. *Optical Society of America Journal A* **4**, 189–199.

Lawson, P. R., *et al.*, 2004. An interferometric imaging beauty contest. In *New frontiers in stellar interferometry* (W. A. Traub, ed.), volume 5491, pp. 886–899, Proc. Soc. Photo-Opt. Instrum. Eng.

Le Besnerais, G., 1993a. *Méthode du maximum d'entropie sur la moyenne, critères de reconstruction d'image et synthèse d'ouverture en radio-astronomie*. Thèse de doctorat, Université de Paris-Sud, Orsay.

Le Besnerais, G., 1993b. Méthodes de calibration et d'autocalibration en radio-astronomie par voie interférométrique. *Traitement du Signal* **10** 5-6, 491–497.

Le Besnerais, G., Navaza, J., & Demoment, G., 1991. Synthèse d'ouverture en radio-astronomie par maximum d'entropie sur la moyenne. In *Actes du 13$^{ème}$ colloque GRETSI*, pp. 217–220, Juan-les-Pins, France.

Lucy, L. B., 1974. An iterative technique for the rectification of observed distributions. *Astron. J.* **79** 6, 745–754.

Meimon, S., Mugnier, L. M., & Le Besnerais, G., accepté pour publication. A convex approximation of the likelihood in optical interferometry. *J. Opt. Soc. Am. A* .

Meimon, S. C., Mugnier, L. M., & Le Besnerais, G., 2004. A novel method of reconstruction for weak-phase optical interferometry. In *New frontiers in stellar interferometry* (W. A. Traub, ed.), volume 5491, pp. 909–919, Proc. Soc. Photo-Opt. Instrum. Eng., Date conférence : June 2004, Glasgow, UK.

Monnier, J. D., 1999. An Introduction to Closure Phases. In *Principles of Long Baseline Stellar Interferometry* (P. R. Lawson, ed.), chapter 13, pp. 203–239, Jet Propulsion Laboratory.

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

Mugnier, L. & Le Besnerais, G., 2001. Problèmes inverses en imagerie optique à travers la turbulence. In *Approche bayésienne pour les problèmes inverses* (J. Idier, ed.), chapter 10, pp. 241–270, Hermès, Paris.

Mugnier, L. M., Fusco, T., & Conan, J.-M., 2004b. 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.

Mugnier, L. M., Rousset, G., & Cassaing, F., 1996. Aperture Configuration Optimality Criterion for Phased Arrays of Optical Telescopes. *J. Opt. Soc. Am. A* **13** 12, 2367–2374.

Mérand, A., Bordé, P., & Coudé du Foresto, V., 2005. A catalog of bright calibrator stars for 200-m baseline near-infrared stellar interferometry. *Astron. Astrophys.* **433**, 1155–1162.

Narayan, R. & Nityananda, R., 1986. Maximum Entropy Image Restoration in Astronomy. *Ann. Rev. Astron. Astrophys.* **24** 2, 127–170.

Navaza, J., 1985. On the maximum-entropy estimate of the electron density function. *Acta Cryst.* **A41**, 232–244.

Nityananda, R. & Narayan, R., 1982. Maximum Entropy Image Reconstruction-A practical Non-Information-Theoretic Approach. *J. Astrophys. Astr.* **3**, 419–450.

Oppenheim, A. V., Hayes, M. H., & Lim, J. S., 1980. Iterative procedures for signal reconstruction from phase. In *1980 International Optical Computing Conference, Washington, DC, April 8-11, 1980, Proceedings. Book 1. (A81-37755 17-35) Bellingham, WA, Society of Photo-Optical Instrumentation Engineers, 1980, p. 121-129.*, pp. 121–129.

Oppenheim, A. V. & Lim, J. S., 1981. The importance of phase in signals. *IEEE Proceedings* **69**, 529–541.

Pauls, T. *et al.*, 2004. A data exchange standard for optical (visible/IR) interferometry. In *New frontiers in stellar interferometry*, volume 5491, Proc. Soc. Photo-Opt. Instrum. Eng.

Pedretti, E., Millan-Gabet, R., Monnier, J. D., Traub, W. A., Carleton, N. P., Berger, J.-P., Lacasse, M. G., Schloerb, F. P., & Brewer, M. K., 2004. The PICNIC Interferometry Camera at IOTA. *Pub. Astron. Soc. Pacific* **116**, 377–389.

Perrin, G., 2003. The calibration of interferometric visibilities obtained with single-mode optical interferometers. Computation of error bars and correlations. *Astron. Astrophys.* **400**, 1173–1181.

Perrin, G., *et al.*, 2004. Unveiling Mira stars behind the molecules. Confirmation of the molecular layer model with narrow band near-infrared interferometry. *Astron. Astrophys.* **426**, 279–296.

Renzini, 1988. Tests of Evolutionary Sequences Using Color-Magnitude Diagrams of Globular Clust ers. *Annual Review of Astronomy and Astrophysics* **26**, 199–244.

Roddier, F., 1981. The effects of atmospherical turbulence in optical astronomy. In *Progress in Optics* (E. Wolf, ed.), volume XIX, pp. 281–376, North Holland, Amsterdam.

Rogstad, D. H., 1968. A technique for measuring visibility phase with an optical interferometer in the presence of atmospheric seeing. *Appl. Opt.* **7**, 585–588.

Schwab, 1983. Optimal Gridding of Visibility Data in Radio-interferometry. In *Indirect Imaging, Proc. of an International Symposium* (J. Roberts, ed.), pp. 333–346, Cambridge Univ. Press, Sidney.

Skilling, J. & Bryan, R., 1984. Maximum entropy image reconstruction : general algorithm. *Mon. Not. R. Astr. Soc.* **211**, 111–124.

Thiébaut, E., Garcia, P. J. V., & Foy, R., 2003. Imaging with Amber/VLTI : the case of microjets. *Astrophys. Space. Sci.* **286**, 171–176.

Thiébaut, E., 2002. Optimization issues in blind deconvolution algorithms. In *Astronomical Data Analysis*, volume 4847, Proc. Soc. Photo-Opt. Instrum. Eng.

Thompson, Moran, & Swenson, 1986. *Interferometry and synthesis in Radio-astronomy*. Wiley Interscience, New-York.

Titterington, D. M., 1985. General structure of regularization procedures in image reconstruction. *Astron. Astrophys.* **144**, 381–387.

Traub, W. A., *et al.*, 2000. Third telescope project at the IOTA interferometer. In *Proc. SPIE Vol. 4006, p. 715-722, Interferometry in Optical Astronomy, Pierre J. Lena ; Andreas Quirrenbach ; Eds.*, pp. 715–722.

Traub, W. A., *et al.*, 2003. New Beam-Combination Techniques at IOTA. In *Interferometry for Optical Astronomy II. Edited by Wesley A. Traub. Proceedings of the SPIE, Volume 4838, pp. 45-52 (2003).*, pp. 45–52.

Verhoelst, T., 2005. *Evolved stars : a combined view from interferometry and spectroscopy*. Ph.D. thesis, Katholieke Universiteit Leuven, Faculty of Science.

Wakker, B. P. & Schwarz, U. J., 1988. The Multi-Resolution CLEAN and its application to the short-spacing problem in interferometry. *Astron. Astrophys.* **200**, 312–322.

# Annexe A

# Approximation optimale par une loi gaussienne.

## A.1 Démonstration dans le cas général

Nous démontrons ici que, pour toute distribution $f(\boldsymbol{y})$, la distribution gaussienne définie par :

$$g(\boldsymbol{y}) = \frac{1}{\sqrt{2\pi \det \boldsymbol{R_y}}} \mathrm{e}^{-\frac{1}{2}P(\boldsymbol{y})}$$

avec $\boldsymbol{R_y}$ une matrice symétrique définie positive et

$$P(\boldsymbol{y}) = (\boldsymbol{y} - \bar{\boldsymbol{y}})^t \, \boldsymbol{R}_{\boldsymbol{y}}^{-1} \, (\boldsymbol{y} - \bar{\boldsymbol{y}})$$

qui minimise la distance de Kullback-Leibler à $f$ $\delta(f, g)$, est telle que :

$$\begin{aligned} \bar{\boldsymbol{y}} &= \mathrm{E}_f \{\boldsymbol{y}\} \\ \boldsymbol{R_y} &= \mathrm{E}_f \left\{ (\boldsymbol{y} - \mathrm{E}_f \{\boldsymbol{y}\}) \, (\boldsymbol{y} - \mathrm{E}_f \{\boldsymbol{y}\})^t \right\} \\ &= \mathrm{Var}(\boldsymbol{y}) \end{aligned} \tag{A.1}$$

Cette propriété est très certainement un corrolaire de résultats généraux de la théorie des probabilités, mais nous en proposons ici une démonstration simple et rapide.

### A.1.1 Définition de la distance de Kullback-Leibler

La distance $\delta(f, g)$ est définie par

$$\begin{aligned} \delta(f, g) &= \int f(\boldsymbol{y}) \log \frac{f(\boldsymbol{y})}{g(\boldsymbol{y})} \mathrm{d}y \\ &= -\mathrm{E}_f \{\log g\} + cst \end{aligned}$$

Donc

$$\delta(f, g) = \frac{1}{2} \left( \mathrm{E}_f \{P(\boldsymbol{y})\} + \log \det \boldsymbol{R_y} \right) + cst$$

### A.1.2 Termes du premier ordre

$$\begin{aligned}\frac{\partial}{\partial \bar{\boldsymbol{y}}}\delta(f,g)=0 \Rightarrow & \frac{\partial \mathrm{E}_f\left\{P(\boldsymbol{y})\right\}}{\partial \bar{\boldsymbol{y}}}=0\\ \Rightarrow & \mathrm{E}_f\left\{\frac{\partial P(\boldsymbol{y})}{\partial \bar{\boldsymbol{y}}}\right\}=0\\ \Rightarrow & \mathrm{E}_f\left\{2\boldsymbol{R}_{\boldsymbol{y}}^{-1}\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\right\}=0\\ \Rightarrow & \mathrm{E}_f\left\{\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\right\}=0\\ \Rightarrow & \mathrm{E}_f\left\{\boldsymbol{y}\right\}=\bar{\boldsymbol{y}}\end{aligned}$$

### A.1.3 Termes du second ordre

$$\begin{aligned}&\frac{\partial\delta(f,g)}{\partial \boldsymbol{R}_{\boldsymbol{y}}}=0\\ &\Rightarrow \frac{\partial}{\partial \boldsymbol{R}_{\boldsymbol{y}}}\left[\mathrm{E}_f\left\{P\right\}+\log\det \boldsymbol{R}_{\boldsymbol{y}}\right]=0\\ &\Rightarrow \mathrm{E}_f\left\{\frac{\partial P}{\partial \boldsymbol{R}_{\boldsymbol{y}}}\right\}+\frac{\partial \log\det \boldsymbol{R}_{\boldsymbol{y}}}{\partial \boldsymbol{R}_{\boldsymbol{y}}}=0\\ &\Rightarrow \mathrm{E}_f\left\{-\boldsymbol{R}_{\boldsymbol{y}}^{-t}\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)^t\boldsymbol{R}_{\boldsymbol{y}}^{-t}\right\}+\boldsymbol{R}_{\boldsymbol{y}}^{-t}=0\end{aligned}$$

Puisque $\boldsymbol{R}_{\boldsymbol{y}}$ est symmetrique, on a : $\boldsymbol{R}_{\boldsymbol{y}}^{-t}=\boldsymbol{R}_{\boldsymbol{y}}^{-1}$, d'où

$$\begin{aligned}&\frac{\partial\delta(f,g)}{\partial \boldsymbol{R}_{\boldsymbol{y}}}=0\\ &\Rightarrow \boldsymbol{R}_{\boldsymbol{y}}^{-t}\mathrm{E}_f\left\{\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)^t\right\}\boldsymbol{R}_{\boldsymbol{y}}^{-t}=\boldsymbol{R}_{\boldsymbol{y}}^{-t}\\ &\Rightarrow \mathrm{E}_f\left\{\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)^t\right\}=\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{R}_{\boldsymbol{y}}^{-t}\boldsymbol{R}_{\boldsymbol{y}}\\ &\Rightarrow \boldsymbol{R}_{\boldsymbol{y}}=\mathrm{E}_f\left\{\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)\left(\boldsymbol{y}-\bar{\boldsymbol{y}}\right)^t\right\}\\ &\Rightarrow \boldsymbol{R}_{\boldsymbol{y}}=\mathrm{Var}(\boldsymbol{y})\end{aligned}$$

qui conclue la démonstration.

### A.1.4 Cas particulier plan

Dans le cas où l'espace vectoriel considéré est de dimension 2, soit $\boldsymbol{y}=(y_1,y_2)$, le résultat précédent prend la forme particulière suivante :

$$\begin{aligned}g(y_1,y_2)&=\frac{1}{\sqrt{2\pi\det \boldsymbol{R}_{\boldsymbol{y}}}}\mathrm{e}^{-\frac{1}{2}P(\boldsymbol{y})}\\ P(\bar{y_1},\bar{y_2})&=\begin{bmatrix} y_1-\bar{y_1}\\ y_2-\bar{y_2}\end{bmatrix}^t\boldsymbol{R}_{\boldsymbol{y}}^{-1}\begin{bmatrix} y_1-\bar{y_1}\\ y_2-\bar{y_2}\end{bmatrix}\end{aligned}$$

L'approximation optimale $g$ de $f$ a pour paramètres :

$$\begin{aligned} \bar{y_1} &= \mathrm{E}_f\left\{y_1\right\} \\ \bar{y_2} &= \mathrm{E}_f\left\{y_2\right\} \\ \boldsymbol{R_y} &= \mathrm{E}_f\left\{ \begin{bmatrix} y_1 - \bar{y_1} \\ y_2 - \bar{y_2} \end{bmatrix} \begin{bmatrix} y_1 - \bar{y_1} \\ y_2 - \bar{y_2} \end{bmatrix}^t \right\} \end{aligned} \tag{A.2}$$

## A.2 Application au contexte interférométrique optique

Dans ce manuscrit, on s'intérresse au cas suivant :

– L'espace vectoriel considéré est de dimension 2, soit $\boldsymbol{y} = (y_1, y_2)$
– La loi que l'on souhaite approcher par une loi gaussienne s'écrit

$$f(y_1, y_2) = f((a_0 + a)\cos\phi - a_0, (a_0 + a)\sin\phi)$$

où les variables $a$ et $\phi$ ont pour distribution :

$$F(a, \phi) \propto \exp -\frac{1}{2} \begin{bmatrix} a \\ \phi \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \sigma_a^{-2} & 0 \\ 0 & \sigma_\phi^{-2} \end{bmatrix} \begin{bmatrix} a \\ \phi \end{bmatrix}$$

Avec cette distribution, on a :

$$\begin{aligned} \mathrm{E}_f\left\{a\right\} &= 0 \\ \mathrm{E}_f\left\{\phi\right\} &= 0 \\ \mathrm{E}_f\left\{a^2\right\} &= \sigma_a^2 \\ \mathrm{E}_f\left\{\phi^2\right\} &= \sigma_\phi^2 \\ \mathrm{E}_f\left\{\cos k\phi\right\} &= \mathrm{e}^{-\frac{(k\sigma_\phi)^2}{2}} \\ \mathrm{E}_f\left\{\sin k\phi\right\} &= 0 \\ \mathrm{E}_f\left\{\cos^2\phi\right\} &= \mathrm{E}_f\left\{\frac{1+\cos 2\phi}{2}\right\} \\ &= \frac{1+\mathrm{e}^{-\sigma_\phi^2}}{2} \\ \mathrm{E}_f\left\{\sin^2\phi\right\} &= \mathrm{E}_f\left\{\frac{1-\cos 2\phi}{2}\right\} \\ &= \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2} \\ \mathrm{E}_f\left\{\sin\phi\cos\phi\right\} &= \mathrm{E}_f\left\{\frac{\sin 2\phi}{2}\right\} \\ &= 0 \end{aligned}$$

### A.2.1 Termes du premier ordre

La moyenne $\bar{\boldsymbol{y}}$ de l'approximation optimale s'écrit donc :

$$\begin{aligned}
\bar{y_1} &= \mathrm{E}_f\left\{y_1\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)\cos\phi - a_0\right\} \\
&= (a_0 + \mathrm{E}_f\left\{a\right\})\mathrm{E}_f\left\{\cos\phi\right\} - a_0 \\
&= a_0 \mathrm{e}^{-\frac{\sigma_\phi^2}{2}} - a_0 \\
\bar{y_2} &= \mathrm{E}_f\left\{y_2\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)\sin\phi\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)\right\}\mathrm{E}_f\left\{\sin\phi\right\} \\
&= 0
\end{aligned}$$

En particulier, la variable centrée $\boldsymbol{y} - \bar{\boldsymbol{y}}$ s'écrit

$$\left[\begin{array}{c} y_1 - \bar{y_1} \\ y_2 - \bar{y_2} \end{array}\right] = \left[\begin{array}{c} (a_0+a)\cos\phi - a_0\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} \\ (a_0+a)\sin\phi \end{array}\right]$$

### A.2.2 Termes du second ordre

La matrice de covariance $\boldsymbol{R}_y$ de l'approximation optimale s'écrit donc :

$$\begin{aligned}
\boldsymbol{R_y} &= \mathrm{E}_f\left\{\left[\boldsymbol{y} - \bar{\boldsymbol{y}}\right]\left[\boldsymbol{y} - \bar{\boldsymbol{y}}\right]^{\mathrm{T}}\right\} \\
&= \left[\begin{array}{cc} r_1 & r_{12} \\ r_{12} & r_2 \end{array}\right]
\end{aligned}$$

avec :

$$
\begin{aligned}
r_1 &= \mathrm{E}_f\left\{(y_1-\bar{y_1})^2\right\} \\
&= \mathrm{E}_f\left\{\left[(a_0+a)\cos\phi - a_0\mathrm{e}^{-\frac{\sigma_\phi^2}{2}}\right]^2\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)^2\cos^2\phi - 2a_0(a_0+a)\cos\phi\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} + a_0^2\mathrm{e}^{-\sigma_\phi^2}\right\} \\
&= \frac{1+\mathrm{e}^{-\sigma_\phi^2}}{2}\mathrm{E}_f\left\{(a_0+a)^2\right\} - 2a_0^2\mathrm{e}^{-\sigma_\phi^2} + a_0^2\mathrm{e}^{-\sigma_\phi^2} \\
&= \frac{(1+\mathrm{e}^{-\sigma_\phi^2})(a_0^2+\sigma_a^2)}{2} - a_0^2\mathrm{e}^{-\sigma_\phi^2} \\
&= \frac{\sigma_a^2(1+\mathrm{e}^{-\sigma_\phi^2})}{2} + \frac{a_0^2(1+\mathrm{e}^{-\sigma_\phi^2})}{2} - a_0^2\mathrm{e}^{-\sigma_\phi^2} \\
\Rightarrow r_1 &= \frac{1+\mathrm{e}^{-\sigma_\phi^2}}{2}\,\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,a_0^2
\end{aligned}
$$

$$
\begin{aligned}
r_2 &= \mathrm{E}_f\left\{(y_2-\bar{y_2})^2\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)^2\sin^2\phi\right\} \\
&= (a_0^2+\sigma_a^2)\frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2} \\
\Rightarrow r_2 &= \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,a_0^2
\end{aligned}
$$

$$
\begin{aligned}
r_{12} &= \mathrm{E}_f\left\{(y_1-\bar{y_1})(y_2-\bar{y_2})\right\} \\
&= \mathrm{E}_f\left\{(a_0+a)\sin\phi\left((a_0+a)\cos\phi - a_0\mathrm{e}^{-\frac{\sigma_\phi^2}{2}}\right)\right\} \\
&= \underbrace{\mathrm{E}_f\left\{\sin\phi\cos\phi\right\}}_{=0} - \mathrm{E}_f\left\{(a_0+a)^2\right\}\underbrace{\mathrm{E}_f\left\{\sin\phi\right\}}_{=0}a_0\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} \\
\Rightarrow r_{12} &= 0
\end{aligned}
$$

### A.2.3 Bilan

Les paramètres de l'approximation gaussienne optimale sont :

$$
\begin{cases}
\bar{\boldsymbol{y}} = \begin{bmatrix} a_0\left(\mathrm{e}^{-\frac{\sigma_\phi^2}{2}} - 1\right) \\ 0 \end{bmatrix} \\
\boldsymbol{R_y} = \begin{bmatrix} \frac{1+\mathrm{e}^{-\sigma_\phi^2}}{2}\,\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,a_0^2 & 0 \\ 0 & \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,\sigma_a^2 + \frac{1-\mathrm{e}^{-\sigma_\phi^2}}{2}\,a_0^2 \end{bmatrix}
\end{cases}
$$

# Annexe B

# Clôtures de phase independantes.

Cette annexe a pour but d'identifier une base des combinaisons linéaires des phases mesurées annulant les effets turbulents. Une telle combinaison linéaire, représentée par un vecteur $\boldsymbol{v}$, vérifie

$$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{\phi}^{\mathrm{data}} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\varphi}^{turb} + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{\phi}_b = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{\phi}_b$$

On a alors :

$$\begin{aligned}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\varphi}^{turb} = \boldsymbol{0},\ \forall\boldsymbol{\varphi}^{turb} &\Leftrightarrow \boldsymbol{v}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{0}^{\mathrm{T}}\\ &\Leftrightarrow \boldsymbol{B}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{0}\\ &\Leftrightarrow \boldsymbol{v} \in \ker \boldsymbol{B}^{\mathrm{T}}\end{aligned}$$

## B.1 Définitions

On considère un interféromètre optique à $n$ télescopes. Les télescopes élémentaires du réseau interférométrique sont numérotées de 1 à n. Soit un polygone P dont les $m$ sommets sont les télescopes $T_{P(1)}$, $T_{P(2)}$,..., $T_{P(m)}$. On lui associe une clôture de phase à $m$ télescopes définie par la somme des phases de visibilité complexe pour les fréquences spatiales correspondant aux arrêtes du polygone P :

$$\beta_P = \boldsymbol{\phi}_{P(1)P(2)} + \boldsymbol{\phi}_{P(2)P(3)} + \ldots + \boldsymbol{\phi}_{P(m)P(1)} \tag{B.1}$$

Par exemple, pour le polygone $P = 2457$, la clôture correspondante s'écrit :

$$\beta_{2457} = \boldsymbol{\phi}_{24} + \boldsymbol{\phi}_{45} + \boldsymbol{\phi}_{57} + \boldsymbol{\phi}_{72}$$

On définit dans l'espace $\boldsymbol{F}_{phases}$ des phases de visibilités la base suivante :

$$\left\{\vec{e}_{ij} = \boldsymbol{\phi}_{ij},\ 1 \leq i < j \leq m\right\}$$

Il y a autant de bases que de façons de choisir 2 points parmis n, soit

$$N_b = \mathrm{card}(\boldsymbol{\phi}) = \dim \boldsymbol{F}_{phases} = \binom{n}{2} = \frac{n(n-1)}{2}. \tag{B.2}$$

Avec cette définition, on réécrit donc la clôture $\beta_{2457}$ de la façon suivante :

$$\beta_{2457} = \vec{e}_{24} - \vec{e}_{27} + \vec{e}_{45} + \vec{e}_{57}$$

Les clôtures de phase sont des exemples de combinaisons linéaires des phases, représentées par un vecteur $\boldsymbol{v}$ de coefficients $0$, $1$ ou $-1$. De plus, elles sont incluses dans le noyau de $\boldsymbol{B}^{\mathrm{T}}$

(Eq. 1.18 page 30). Nous nous proposons de démontrer que l'ensemble $\mathcal{C}_1^{(3)}$ des clôtures de phase à 3 telescopes incluant le télescope 1 (la numérotation est arbitraire), de la forme

$$\beta_{1ij} = \boldsymbol{\phi}_{1i} + \boldsymbol{\phi}_{ij} + \boldsymbol{\phi}_{j1} = \boldsymbol{\phi}_{1i} - \boldsymbol{\phi}_{1j} + \boldsymbol{\phi}_{ij}$$

constitue une base du noyau de $\boldsymbol{B}^{\mathrm{T}}$, et en particulier une base de l'ensemble des clôtures à m télescopes, pour tout $m$.

## B.2 Combinaisons linéaires de phases indépendantes annulant la turbulence

Pour chacune des $\frac{n(n-1)}{2}$ bases, la phase mesurée sur les interférogramme s'ecrit :

$$\boldsymbol{\phi}^{\mathrm{data}} = \boldsymbol{\phi}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{\varphi}^{turb} + \boldsymbol{\phi}_b$$

avec

$$\boldsymbol{B} = \left[\begin{array}{c|c|c} -\mathbf{1}_{n-1} & \boldsymbol{I}_{n-1} & \\ \hline \mathbf{0}_{n-2} & -\mathbf{1}_{n-2} & \boldsymbol{Id}_{n-1} \\ \hline \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots} \\ \hline \multicolumn{2}{c|}{\mathbf{0}_{n-2}^{\mathrm{T}}} & -1 \;\; 1 \end{array}\right] \tag{B.3}$$

Pour $n = 4$, cet opérateur s'écrit : $\boldsymbol{B} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$.

$\boldsymbol{B}$ est une application linéaire d'un espace $\boldsymbol{E}_{tel}$ de dimension $n$ (l'espace des phases turbulentes sur chaque telescope, isomorphe à l'espace des sommets du polygone) vers l'espace des phases $\boldsymbol{F}_{phases}$ (isomorphe à l'espace des arrêtes du polygone), de dimension $\frac{n(n-1)}{2}$ (Eq. B.2) :

$$\dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \dim \boldsymbol{F}_{bases} - \mathrm{rg}(\boldsymbol{B}^{\mathrm{T}}) = \frac{n(n-1)}{2} - \mathrm{rg}(\boldsymbol{B}^{\mathrm{T}}). \tag{B.4}$$

Or on peut extraire de $\boldsymbol{B}^{\mathrm{T}}$ la matrice $\boldsymbol{I}_{n-1}$ donc $\mathrm{rg}(\boldsymbol{B}^{\mathrm{T}}) \geq n-1$. D'autre part, il est facile de voir que la somme des colonnes de $\boldsymbol{B}$ vaut 0, donc $\mathrm{rg}(\boldsymbol{B}^{\mathrm{T}}) \leq n-1$. On en déduit que le rang de $\boldsymbol{B}^{\mathrm{T}}$ est exactement $n-1$. L'équation B.4 entraine :

$$\boxed{\dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \frac{(n-1)(n-2)}{2}} \tag{B.5}$$

## B.3 Clôtures indépendantes à 3 télescopes

Considérons l'ensemble des clôtures à 3 telescopes contenant le telescope 1 $\mathcal{C}_1^{(3)}$. Il y en a autant que de façons de choisir 2 points parmis $n-1$ (le point 1 etant exclus), soit

$$\mathrm{card}\,\mathcal{C}_1^{(3)} = \binom{n-1}{2} = \frac{(n-1)(n-2)}{2} \tag{B.6}$$

Ces clôtures sont de la forme :

$$\beta_{1ij} = \boldsymbol{\phi}_{1i} + \boldsymbol{\phi}_{ij} - \boldsymbol{\phi}_{1j},\ 2 \leq i < j \leq n$$

Le vecteur de $\boldsymbol{F}_{phases}$ associé à cette combinaison linéaire est :

$$v_{1ij} = \vec{e}_{1i} + \vec{e}_{ij} - \vec{e}_{1j}.$$

Ces vecteurs forment une famille libre. En effet, on a :

$$\begin{aligned}\sum_{2\leq i<j\leq n} w_{ij}v_{1ij} = 0 \Rightarrow & \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{1i} + w_{ij}\vec{e}_{ij} - w_{ij}\vec{e}_{1j} = 0 \\ \Rightarrow & \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{1i} - w_{ij}\vec{e}_{1j} + \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{ij} = 0 \\ \Rightarrow & \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{1j} - w_{ij}\vec{e}_{1i} = \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{ij}\end{aligned}$$

Le terme de gauche est une combinaison linéaire des vecteurs $\vec{e}_{1k}$, le terme de droite une combinaison linéaire de vecteurs $\vec{e}_{ij}$ **ne contenant pas de vecteur** $\vec{e}_{1k}$. La base des vecteurs $\vec{e}_{ij}$ étant libre, les deux termes sont donc nuls :

$$\begin{aligned}\sum_{2\leq i<j\leq n} w_{ij}v_{1ij} = 0 \Rightarrow & \sum_{2\leq i<j\leq n} w_{ij}\vec{e}_{ij} = 0 \\ & \Rightarrow w_{ij} = 0,\ 2 \leq i < j \leq n\end{aligned}$$

Par conséquent, l'ensemble des clôtures à 3 telescopes contenant le telescope 1 $\mathcal{C}_1^{(3)}$ est une famille libre.

On peut obtenir ce résultat de façon intuitive : dans $\mathcal{C}_1^{(3)}$, seule la clôture $\beta_{1ij}$ contient la phase $\phi_{ij}$, d'où l'indépendance des clôtures à trois telescopes contenant le telescope 1.

## B.4 Ces clôtures forment une base du noyau de $\ker(\boldsymbol{B}^{\mathrm{T}})$

Nous avons montré que :

1. $\dim(\ker(\boldsymbol{B}^{\mathrm{T}})) = \frac{(n-1)(n-2)}{2}$ (Eq. B.5) ;
2. $\mathcal{C}_1^{(3)}$ est une famille libre de $\ker(\boldsymbol{B}^{\mathrm{T}})$ ;
3. $\mathcal{C}_1^{(3)}$ est de cardinal $\frac{(n-1)(n-2)}{2}$ (Eq. B.6).

| $\mathcal{C}_1^{(3)}$ est donc une famille libre maximale de $\ker(\boldsymbol{B}^{\mathrm{T}})$, et donc une base de $\ker(\boldsymbol{B}^{\mathrm{T}})$. |
|---|

En particulier, cette base génère l'ensemble des clôtures, dont par exemple :

– les clôtures à 3 telescopes ne contenant pas le telescope 1 :

$$\begin{aligned}v_{ijk} &= \vec{e}_{ij} + \vec{e}_{jk} - \vec{e}_{ik} \\ &= (\vec{e}_{1i} + \vec{e}_{ij} - \vec{e}_{1j}) + (\vec{e}_{1j} + \vec{e}_{jk} - \vec{e}_{1k}) + (\vec{e}_{1k} - \vec{e}_{ik} - \vec{e}_{1i}) \\ &= v_{1ij} + v_{1jk} - v_{1ik}\end{aligned}$$

– les clôtures à 4 telescopes :

$$\begin{aligned}v_{ijkl} &= \vec{e}_{ij} + \vec{e}_{jk} + \vec{e}_{kl} - \vec{e}_{il} \\ &= (\vec{e}_{1i} + \vec{e}_{ij} - \vec{e}_{1j}) + (\vec{e}_{1j} + \vec{e}_{jk} - \vec{e}_{1k}) + (\vec{e}_{1k} + \vec{e}_{kl} - \vec{e}_{1l}) + (\vec{e}_{1l} - \vec{e}_{il} - \vec{e}_{1i}) \\ &= v_{1ij} + v_{1jk} + v_{1kl} - v_{1il}\end{aligned}$$

## B.5 Les opérateurs $C$ et $\bar{B}$

On définit $\bar{\boldsymbol{B}}$ comme la matrice formées de toutes les colonnes de $\boldsymbol{B}$ sauf la première. On définit l'opérateur de clôture $\boldsymbol{C}$ comme la matrice dont chaque ligne correspond à une clôture à 3 télescopes incluant le télescope 1, soit pour 4 télescopes :

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

Avec ce qui précède, les lignes de $\boldsymbol{C}$ forment une famille libre. Le rang de $\boldsymbol{C}$ est donc exactement le nombre de ses lignes, c'est-à-dire (cf eq. B.6) :

$$\mathrm{rg}(\boldsymbol{C}) = \mathrm{card}\, \mathcal{C}_1^{(3)} = \frac{(n-1)(n-2)}{2} \tag{B.7}$$

De plus, il est facile de voir que cet opérateur est de la forme :

$$\boldsymbol{C} = \left[ \; * \; \middle| \; \mathbf{Id} \; \right] \tag{B.8}$$

L'equation B.3 montre que $\bar{\boldsymbol{B}}$ est de rang $n-1$. De plus, comme $\boldsymbol{CB} = 0$, l'image de $\boldsymbol{B}$ est incluse dans le noyau de $\boldsymbol{C}$. Comme l'image de $\bar{\boldsymbol{B}}$ est incluse dans celle de $\boldsymbol{B}$, alors elle est aussi dans le noyau de $\boldsymbol{C}$ :

$$\boldsymbol{C}\bar{\boldsymbol{B}} = 0.$$

L'opérateur $\boldsymbol{C}$ a pour espace de départ l'espace des phases $\boldsymbol{F}_{phases}$, de dimension $\frac{n(n-1)}{2}$ (Eq. B.2), donc :

$$\dim(\ker \boldsymbol{C}) + \mathrm{rg}(\boldsymbol{C}) = \frac{n(n-1)}{2}$$

Avec l'équation B.7, on obtient donc :

$$\dim(\ker \boldsymbol{C}) = \frac{n(n-1)}{2} - \mathrm{rg}(\boldsymbol{C}) = n-1$$

Comme $\bar{\boldsymbol{B}}$ est de rang $n-1$, et que $\boldsymbol{C}\bar{\boldsymbol{B}} = 0$, on en déduit que :

$$\boxed{\ker \boldsymbol{C} = \mathrm{Im}\ \bar{\boldsymbol{B}}} \tag{B.9}$$

# Annexe C

# Optimisation quadratique sous contrainte de positivité

Cette annexe reprend la description de la méthode d'optimisation quadratique sous contrainte de positivité par activation de contraintes de Gilbert (2002).

Si le modèle de formation des données est additif gaussien, et que le terme de régularisation est lui aussi quadratique, la minimisation du critère composite $\mathcal{J} = \mathcal{J}_{\text{data}} + \mathcal{J}_{\text{prior}}$ sous contrainte de positivité est un problème d'optimisation quadratique convexe, c'est à dire :

$$(P) \begin{cases} \min f(x) \\ Ax \leq b, \end{cases} \tag{C.1}$$

avec :

$$\begin{cases} f(x) &= \frac{1}{2}x^T Q x + q^T x, \\ A &= -Id, \\ b &= 0 \end{cases}$$

## C.1 Théorème de Karush, Kuhn et Tucker

### C.1.1 Formulation générale d'un problème d'optimisation sous contraintes

Problème général :

$$(P) \begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ c_i(x) = 0,\ i \in \mathbb{E} \\ c_i(x) \leq 0,\ i \in \mathbb{I} \end{cases} \tag{C.2}$$

Problème de minimisation avec positivité :

$$(P_p) \begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ x_i \geq 0,\ \forall i \end{cases} \tag{C.3}$$

### C.1.2 Caracterisation d'optimalité

Le but de cette section est de determiner des conditions necessaires d'optimalités suffisament pratiques pour inspirer un algorithme de minimisation.
Dans certains cas, il existe plusieurs solutions au problème d'optimisation. On pose $\Omega$ l'ensemble des $x$ acceptables :

$\Omega = \{x \in \mathbb{R}^n : c_i(x) = 0,\ i \in \mathcal{E} \text{ et } c_i(x) \leq 0,\ i \in \mathcal{I}\}$
$x_*$ est une solution locale si on a un voisinage $\mathcal{V} \subset \Omega$ de $x_*$ tel que :
$f(x) \geq f(x_*),\ \forall x \in \mathcal{V}$
S'il n'existe qu'une seule solution locale, c'est la solution globale.

Dans le cas de contraintes de type inégalité, on définit le caractère actif ou inactif d'une contrainte :

- $c_i$ est une contrainte active en $x$ : $c_i(x) = 0$
- $c_i$ est une contrainte inactive en $x$ : $c_i(x) > 0$
- $\mathcal{A}(x) = \{i \in \mathcal{I} \ :\ c_i(x) = 0\}$. $\mathcal{A}(x)$ est l'ensemble des contraintes actives en $x$.

L'utilisation de lagrangiens permet d'exprimer simplement les conditions d'optimalité d'un point. On définit donc les notions suivantes :

$\lambda$ est un multiplicateur de Lagrange pour les contraintes $c_i$ en $x$ si :

$$\nabla f(x) = \sum_i \lambda_i(x) \nabla c_i(x)$$

De plus, on pose $\lambda^*$ le multiplicateur de Lagrange en $x_*$, et on définit :

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$$

$\mathcal{L}$ est le **lagrangien associé au problème**.

Le resultat suivant est à la base de la methode de minimisation que nous utilisons :

**Théorème 1 (CN1 de Karush-Kuhn-Tucker)** *Soit $x_*$ un minimum local de $(P)$. Supposons que $f$ et $c$ sont dérivables en $x_*$. Alors il existe un $\lambda^*$ tel que l'on ait*

$$(KKT) \begin{cases} (a)\ \nabla f(x_*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(x) = 0 \Leftrightarrow \nabla_x \mathcal{L}(x, \lambda) = 0 \\ (b)\ c_i(x_*) = 0,\ \forall i \in \mathcal{E} \\ (c)\ c_i(x_*) \geq 0,\ \forall i \in \mathcal{I} \\ (d)\ \lambda_i^* \geq 0,\ \forall i \in \mathcal{I} \\ (e)\ c_i(x_*)\lambda_i^* = 0,\ \forall i \in \mathcal{I} \end{cases} \tag{C.4}$$

Pour une classe importante de problèmes, $\lambda^*$ est déterminé de façon unique. En effet, si l'ensemble $\{\nabla c_i(x^*) \ :\ i \in \mathcal{A}(x^*)\}$ est linéairement indépendant, alors $\lambda^*$ vérifiant (C.4) est unique. C'est en particulier le cas lorsque les contraintes sont de type positivité.

## C.2 Application à l'optimisation quadratique

Dans le cadre de l'optimisation quadratique, la fonction $f$ à minimiser s'écrit :

$$f(x) = \frac{1}{2} x^t H x + x^t d$$

### C.2.1 Cas des contraintes égalités

Les contraintes sont de la forme $Ax = b$. $A$ est appelé **Jacobien des contraintes**.
On a donc : $\mathcal{L}(x, \lambda) = \frac{1}{2}x^t Hx + x^t d - \lambda^t(Ax - b)$. Les equations d'optimalité sont :

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) = 0 &\Leftrightarrow x^{*t} H + d^t - \lambda^{*t} A = 0 \\ &\Leftrightarrow Hx^* + d - A^t \lambda^* = 0 \quad \text{(C.5)} \\ Ax^* &= b \quad \text{(C.6)} \end{aligned}$$

Ceci peut s'exprimer par l'equation matricielle suivante :

$$\begin{pmatrix} H & -A^t \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} -d \\ b \end{pmatrix}$$

On appelle matrice de Karush-Kuhn-Tucker (matrice KKT) la matrice

$$KKT = \begin{pmatrix} H & -A^t \\ A & 0 \end{pmatrix}$$

### C.2.2 Cas des contraintes inégalités

Dans le cas des contraintes inégalités, le principe est de se fixer un ensemble de contraintes actives, et de chercher l'optimal correspondant. Si un tel point optimal existe, on teste s'il valide les autres contraintes. On dispose alors d'un candidat à l'optimalité. Une approche systématique consisterait à reproduire ce calcul pour toutes les combinaisons de contraintes possibles.
On va plutot chercher à proceder de façon plus progressive, en faisant evoluer notre ensemble de contraintes actives, ou *active set*, vers celui correspondant à la solution du problème complet :

**Algorithme 1 (activation des contraintes)** *L'algorithme suivant converge (hors cyclage*) pour $H \geq 0$, avec une region admissible convexe :*

1. *Choisir un point de départ $x_0$. On prend $\mathcal{C}_0 = \mathcal{A}(x_0)$*
2. **Pour** $k = 0, 1, 2, \ldots$, *determiner $p_k$ tel que $\hat{x} = x_k + p_k$ soit optimal pour les contraintes $\mathcal{C}_k$.*

   **Si** $p_k = 0$ *($x_k$ optimal pour les contraintes $\mathcal{C}_k$)*
   - *Calculer $\hat{\lambda}_i$ en $\hat{x} = x_k + p_k$ pour les contraintes $\mathcal{C}_k$.*
   - **Si** $\hat{\lambda}_i \geq 0,\ \forall i \in \mathcal{C}_k \cup \mathcal{I}$, **STOP**, *avec $x^* = x_k$*

   **Sinon** – $j =$ *l'indice du $\hat{\lambda}_i$ « le plus négatif », pour $i \in \mathcal{C}_k \cap \mathcal{I}$*
   - $x_{k+1} = x_k$
   - $\mathcal{C}_{k+1} = \mathcal{C}_k \setminus \{j\}$ $\Longrightarrow$ **On retire une contrainte**

   **FinSi**

   **Sinon** *($p_k \neq 0$)*
   - *Calculer $\alpha_k = \max\{\alpha \ :\ x_k + \alpha p_k$ vérifie $\mathcal{C}_k\}$*
   - $x_{k+1} = x_k + \alpha_k p_k$
   - **Si** $x_{k+1}$ *a activé au moins une contrainte $c$ de plus que $x_k$,*
     *on prend $\mathcal{C}_{k+1} = \mathcal{C}_k + c$* $\Longrightarrow$ **On rajoute une contrainte**

     **Sinon** $\mathcal{C}_{k+1} = \mathcal{C}_k$

     **FinSi**

   **FinSi**

   **FinPour**

**Le cyclage intervient quand on enleve et on rajoute toujours la même contrainte.*

## C.3 Application à la contrainte de positivité

Dans le cas où la contrainte est une clause de positivité sur l'ensemble des pixels, les $c_i$ sont des fonctions identité. On exprime les conditions nécéssaires d'optimalité pour le problème de minimisation avec certaines contraines actives (des pixels nuls) pour ensuite deduire l'algorithme de recherche de l'optimal du probleme de depart.

### C.3.1 Pixels « gelés »

On considère un ensemble $\mathcal{C}$ de pixels que l'on impose à 0, et on cherche à minimiser $f$. Quitte à réordonner les pixels, on considère que ce sont les $k$ premiers que l'on doit annuler. La matrice $A$ correspondante est donc, **à une permutation des lignes près** :

$$A = \left[\ \mathbb{I}_k \ \middle|\ 0\ \right]$$

La matrice KKT s'écrit :

$$KKT = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{H}} & -\mathbb{I}_k \\ & & 0 \\ \hline \mathbb{I}_k & 0 & 0 \end{array}\right)$$

à une permutation des lignes de $A$ près. Les lignes de $A$ etant linéairement indépendantes, l'équation

$$\left(\begin{array}{cc|c} \multicolumn{2}{c|}{\text{H}} & \begin{array}{c}-\mathbb{I}_k\\ \hline 0\end{array} \\ \hline \mathbb{I}_k & 0 & 0 \end{array}\right)\left(\begin{array}{c} x^* \\ \hline \lambda^* \end{array}\right) = \left(\begin{array}{c} \text{-d} \\ \hline 0 \end{array}\right)$$

a une unique solution $(\hat{x}, \hat{\lambda})$.

## C.3.2 Algorithme

**Algorithme 2 (optimisation quadratique sous positivité)** *$\mathcal{P}_k$ représente l'ensemble des pixels gelés (i.e. nuls) à l'étape $k$.*

1. *Choisir un point de départ $x_0$ où tous les pixels sont positifs. On prend $\mathcal{P}_0 = \{i \ : \ (x_0)_i = 0\}$*
2. **Pour** $k = 0, 1, 2, \ldots$, *determiner $p_k$ tel que $\hat{x} = x_k + p_k$ soit optimal, avec $x_i = 0 \ \forall i \in \mathcal{P}_k$.*

   **Si** $p_k = 0$ *($x_k$ optimal avec les pixels $\mathcal{P}_k$ nuls)*
   - *Calculer $\hat{\lambda}_i$ en $\hat{x} = x_k + p_k$ pour les pixels $\mathcal{P}_k$ nuls.*
   - **Si** $\hat{\lambda}_i \geq 0, \ \forall i \in \mathcal{P}_k \cup \mathcal{I}$, **STOP**, *avec $x^* = x_k$*

     **Sinon** – *$j$ = l'indice du $\hat{\lambda}_i$ « le plus négatif », pour $i \in \mathcal{P}_k$*
     - $x_{k+1} = x_k$
     - $\mathcal{P}_{k+1} = \mathcal{P}_k \setminus \{j\}$ $\Longrightarrow$ **On dégèle le j[ème] pixel**

     **FinSi**

   **Sinon** *($p_k \neq 0$)*
   - *Calculer $\alpha_k = \max\{\alpha \ : \ x_k + \alpha p_k$ avec les pixels $\mathcal{P}_k$ nuls$\}$*
   - $x_{k+1} = x_k + \alpha_k p_k$
   - **Si** $x_{k+1}$ *a au moins un pixel nul þ de plus que $x_k$,*
     $\mathcal{P}_{k+1} = \mathcal{P}_k + \{$*þ est nul*$\}$ $\Longrightarrow$ **On gèle le pixel þ**.

     **Sinon** $\mathcal{P}_{k+1} = \mathcal{P}_k$

     **FinSi**

   **FinSi**

**FinPour**

## C.3.3 Cas non quadratique

Si l'on choisit un modèle *a priori* non gaussien, le critère composite $\mathcal{J} = \mathcal{J}_{\text{data}} + \mathcal{J}_{\text{prior}}$ n'est pas quadratique. Dans ce cas, on peut approximer la fonction à minimiser par une fonction quadratique. Un pas de descente permet de déterminer un $\hat{x}$ « presque » stationnaire.

# Annexe D

# Imagerie et configuration pupillaire

$$\begin{array}{ccc} \Psi_{\boldsymbol{o}}(\boldsymbol{\xi}) & \xrightarrow[\mathcal{H}]{\text{propagation+imagerie}} & \Psi_{\boldsymbol{im}}(\boldsymbol{\xi}) \\ \text{detection}\downarrow \langle ||^2 \rangle_t & & \text{detection}\downarrow \langle ||^2 \rangle_t \\ \boldsymbol{o} & \xrightarrow[\mathcal{G}]{\text{propagation+imagerie}} & \boldsymbol{im} \end{array}$$

## D.1 L'operateur $\mathcal{H} : \Psi_{\boldsymbol{o}} \longrightarrow \Psi_{\boldsymbol{im}}$ est LIT

$$\begin{aligned} \mathcal{H}(\Psi_o * \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_0)) &= \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_o * \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_0)\right)\right) \\ &= \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_o\right) \mathrm{TF}\left(\delta(\boldsymbol{\xi} - \boldsymbol{\xi}_0)\right)\right) \\ &= \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_o\right)\right) * \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \\ &= \mathcal{H}(\Psi_o) * \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \\ \mathcal{H}(\Psi_{o,1} + \alpha \times \Psi_{o,2}) &= \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_{o,1} + \alpha \times \Psi_{o,2}\right)\right) \\ &= \mathrm{TF}\left(p\ \left(\mathrm{TF}\left(\Psi_{o,1}\right) + \alpha \times \mathrm{TF}\left(\Psi_{o,2}\right)\right)\right) \\ &= \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_{o,1}\right)\right) + \alpha \times \mathrm{TF}\left(p\ \mathrm{TF}\left(\Psi_{o,2}\right)\right) \\ &= \mathcal{H}(\Psi_{o,1}) + \alpha \times \mathcal{H}(\Psi_{o,2}) \end{aligned}$$

Puisque $\mathcal{H}$ est LIT, il revient à convoluer par la reponse impulsionnelle (RI) de $\mathcal{H}$ :

$$\boxed{\mathcal{H}(\Psi_o) = \Psi_o * \mathcal{H}(\delta(0,0)) = \Psi_o * RI(\mathcal{H})}$$

## D.2 L'operateur $\mathcal{G} : \boldsymbol{o} \longrightarrow \boldsymbol{im}$ est LIT

Supposons deux champs objet $\Psi_{o,1}, \Psi_{o,2}$ décorrelés :

$$\langle \Psi_{o,1} \times \Psi_{o,2} \rangle_t = 0$$

On a alors :

$$\begin{aligned}
\left\langle |\Psi_o(\boldsymbol{\xi}) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0)|^2 \right\rangle_t &= \left\langle |\Psi_o(\boldsymbol{\xi}-\boldsymbol{\xi}_0)|^2 \right\rangle_t \\
&= \left\langle |\Psi_o(\boldsymbol{\xi})|^2 \right\rangle_t * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0) \\
&= \boldsymbol{o}(\boldsymbol{\xi}) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0) \\
\left\langle |\mathcal{H}(\Psi_o(\boldsymbol{\xi}) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0))|^2 \right\rangle_t &= \left\langle |\mathcal{H}(\Psi_o(\boldsymbol{\xi})) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0)|^2 \right\rangle_t \\
&= \left\langle |\mathcal{H}(\Psi_o(\boldsymbol{\xi}))|^2 \right\rangle_t * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0) \\
&= \boldsymbol{im} * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0)
\end{aligned}$$

d'où

$$\boxed{\mathcal{G}(\boldsymbol{o}(\boldsymbol{\xi}) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0))\mathcal{G}(\boldsymbol{o}(\boldsymbol{\xi})) * \delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0)}$$

$$\begin{aligned}
\left\langle |\Psi_{o,1} + \alpha \times \Psi_{o,2}|^2 \right\rangle_t &= \left\langle |\Psi_{o,1}|^2 \right\rangle_t + \left\langle |\alpha \times \Psi_{o,2}|^2 \right\rangle_t + 2\Re \underbrace{\left\langle \Psi_{o,1} \alpha \times \Psi_{o,2} \right\rangle_t}_{=0} \\
&= \left\langle |\Psi_{o,1}|^2 \right\rangle_t + \alpha^2 \left\langle |\Psi_{o,2}|^2 \right\rangle_t \\
&= \boldsymbol{o}_1 + \alpha^2 \boldsymbol{o}_2 \\
\left\langle |\mathcal{H}(\Psi_{o,1} + \alpha\Psi_{o,2})|^2 \right\rangle_t &= \left\langle |\mathcal{H}(\Psi_{o,1}) + \alpha\mathcal{H}(\Psi_{o,2})|^2 \right\rangle_t \\
&= \left\langle |\mathcal{H}(\Psi_{o,1})|^2 \right\rangle_t + \left\langle |\alpha \times \mathcal{H}(\Psi_{o,2})|^2 \right\rangle_t + 2\Re \left\langle \alpha\mathcal{H}(\Psi_{o,1})\mathcal{H}(\Psi_{o,2}) \right\rangle_t \\
&= \left\langle |\mathcal{H}(\Psi_{o,1})|^2 \right\rangle_t + \alpha^2 \left\langle |\mathcal{H}(\Psi_{o,2})|^2 \right\rangle_t + 2\alpha\Re \left\langle (\Psi_{o,1} * RI(\mathcal{H}))(\Psi_{o,2} * RI(\mathcal{H})) \right\rangle_t \\
&= \boldsymbol{im}_1 + \alpha^2 \boldsymbol{im}_2 + 2\alpha\Re \left\langle (\Psi_{o,1} \times \Psi_{o,2}) * RI(\mathcal{H}) \right\rangle_t \\
&= \boldsymbol{im}_1 + \alpha^2 \boldsymbol{im}_2 + 2\alpha\Re \left[ \underbrace{\left\langle (\Psi_{o,1} \times \Psi_{o,2}) \right\rangle_t}_{=0} * RI(\mathcal{H}) \right] \\
&= \boldsymbol{im}_1 + \alpha^2 \boldsymbol{im}_2
\end{aligned}$$

d'où

$$\boxed{\mathcal{G}(\boldsymbol{o} + \alpha^2 \boldsymbol{o}_2) = \mathcal{G}(\boldsymbol{o}_1) + \alpha^2 \mathcal{G}(\boldsymbol{o}_2)}$$

Puisque $\mathcal{G}$ est LIT, on a :

$$\boxed{\mathcal{G}(\boldsymbol{o}) = \boldsymbol{o} * RI(\mathcal{G})}$$

Avec :

$$\begin{aligned}
\left\langle |\delta(\boldsymbol{\xi})|^2 \right\rangle_t &= \delta(\boldsymbol{\xi}) \\
\left\langle |\mathcal{H}(\delta(\boldsymbol{\xi}))|^2 \right\rangle_t &= \left\langle |\mathrm{TF}\,(p\ \mathrm{TF}\,(\delta(\boldsymbol{\xi})))|^2 \right\rangle_t \\
&= \left\langle |\mathrm{TF}\,(p\,)|^2 \right\rangle_t \\
&= |\mathrm{TF}\,(p\,)|^2 \\
&= \mathrm{TF}\,(p\,)\,\mathrm{TF}\,(p*\,) \\
&= \mathrm{TF}\,(p * p*\,) \\
&= \mathrm{TF}\,(p \otimes p) \\
\mathcal{G}(\delta(\boldsymbol{\xi})) &= \mathrm{TF}\,(p \otimes p)
\end{aligned}$$

Donc

$$\boxed{RI(\mathcal{G}) = \mathrm{TF}\,(p \otimes p)}$$

# Annexe E

# Gradients de $\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha})$

## E.1 Gradient de $\mathcal{J}_{WISARD}(\boldsymbol{x}|\boldsymbol{\alpha})$ par rapport à $\boldsymbol{x}$

$\mathcal{J}_{WISARD}$ s'exprime sous la forme d'une somme de termes

$$\mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_t \mathcal{J}_{ell}(\boldsymbol{x}, \boldsymbol{\alpha}(t), t)$$

et chacun de ces termes s'écrit

$$\mathcal{J}_{ell}(\boldsymbol{x}, \boldsymbol{\alpha}(t), t) = \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}}(t) \cdot \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}\right)$$

On a donc :

$$\begin{aligned}
\frac{\partial \mathcal{J}_{WISARD}(\boldsymbol{x}, \boldsymbol{\alpha})}{\partial \boldsymbol{x}} &= \frac{\partial \sum_t \mathcal{J}_{ell}(\boldsymbol{x}, \boldsymbol{\alpha}(t), t)}{\partial \boldsymbol{x}} \\
&= \sum_t \frac{\partial \mathcal{J}_{ell}(\boldsymbol{x}, \boldsymbol{\alpha}(t), t)}{\partial \boldsymbol{x}} \\
&= \sum_t \frac{\partial \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}}(t) \cdot \left(\bar{\bar{\boldsymbol{y}}}^{data}(t) - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t), t)\boldsymbol{x}\right)}{\partial \boldsymbol{x}} \\
&= \sum_t \bar{\bar{\boldsymbol{H}}}^{\mathrm{T}}(\boldsymbol{\alpha}(t), t) \cdot \boldsymbol{W}_{\bar{\bar{y}}}(t) \cdot \left(\bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha}(t), t)\boldsymbol{x} - \bar{\bar{\boldsymbol{y}}}^{data}(t)\right)
\end{aligned}$$

## E.2 Gradient de $\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})$ par rapport à $\boldsymbol{\alpha}$

On a :

$$\begin{aligned}
\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x}) &= \sum_t \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t) \\
\Rightarrow \frac{\partial \mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})}{\partial \boldsymbol{\alpha}} &= \sum_t \frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \boldsymbol{\alpha}} \\
&= \sum_t \frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \boldsymbol{\alpha}(t)}
\end{aligned}$$

Il suffit donc de calculer le gradient

$$\frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \boldsymbol{\alpha}(t)}$$

Pour alléger les notations, on pourra omettre de préciser l'instant de mesure $t$.
De plus, on pose :

$$\begin{aligned}\boldsymbol{y}(\boldsymbol{x}, \boldsymbol{\alpha}) &= \boldsymbol{H}(\boldsymbol{\alpha})\boldsymbol{x} = \text{Diag}\left\{\exp \mathrm{i}\bar{\boldsymbol{B}}\boldsymbol{\alpha}\right\}\boldsymbol{H}\boldsymbol{x} \\ \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{\alpha}) &= \boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha}\end{aligned}$$

On a alors

$$\begin{aligned}\frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}|\boldsymbol{x})}{\partial \boldsymbol{\alpha}} &= \frac{\partial \left(\bar{\bar{\boldsymbol{y}}}^{data} - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha})\boldsymbol{x}\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}} \cdot \left(\bar{\bar{\boldsymbol{y}}}^{data} - \bar{\bar{\boldsymbol{H}}}(\boldsymbol{\alpha})\boldsymbol{x}\right)}{\partial \boldsymbol{\alpha}} \\ &= \frac{\partial \left(\bar{\bar{\boldsymbol{y}}}^{data} - \bar{\bar{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{\alpha})\right)^{\mathrm{T}} \cdot \boldsymbol{W}_{\bar{\bar{y}}} \cdot \left(\bar{\bar{\boldsymbol{y}}}^{data} - \bar{\bar{\boldsymbol{y}}}(\boldsymbol{x}, \boldsymbol{\alpha})\right)}{\partial \boldsymbol{\alpha}}\end{aligned}$$

### E.2.1 Réécriture de $\mathcal{J}_{ell}$

Si cette notation matricielle est compacte et pratique pour calculer le gradient du critère par rapport à $\boldsymbol{x}$, elle ne permet pas facilement de calculer le gradient par rapport à $\boldsymbol{\alpha}$.

On dispose d'une autre façon d'écrire le critère :

$$\mathcal{J}_{ell} = \sum_i \mathcal{J}_{ell}^i$$

Et pour $\mathcal{J}_{ell}^i$, on a (équation 2.19 page 55) :

$$\mathcal{J}_{ell}^i = \|y_{\text{res},i}\|_{ell,i}^2$$

avec

$$\begin{aligned}y_{\text{res},i} &= y_i(\boldsymbol{x}) - y_i^{\text{data}} \\ \|z\|_{ell,i} &= \frac{(z.\boldsymbol{e}_{\text{rad},i})^2}{\sigma_{\text{rad},i}^2} + \frac{(z.\boldsymbol{e}_{\text{tan},i})^2}{\sigma_{\text{tan},i}^2} \\ &= \frac{\Re^2(z\mathrm{e}^{-\mathrm{i}\,\phi_i^{\text{data}}})}{\sigma_{\text{rad},i}^2} + \frac{\Im^2(z\mathrm{e}^{-\mathrm{i}\,\phi_i^{\text{data}}})}{\sigma_{\text{tan},i}^2}\end{aligned}$$

Pour le modèle myope, les résidus en phase et amplitude s'écrivent (Eq. 4.4 page 99) :

$$\begin{cases}\phi_{\text{res},i} \triangleq -\phi_{b,i} = \phi_i(\boldsymbol{x}, \boldsymbol{\alpha}) - \phi_i^{\text{data}} \\ a_{\text{res},i} \triangleq -a_{b,i} = a_i(\boldsymbol{x}) - a_i^{\text{data}}\end{cases}$$

Le résidu en visibilités complexes s'écrit donc de la façon suivante :

$$\begin{aligned}y_{\text{res},i} &= y_i(\boldsymbol{x}, \boldsymbol{\alpha}) - y_i^{\text{data}} \\ &= a_i(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\,\phi_i(\boldsymbol{x},\boldsymbol{\alpha})} - a_i^{\text{data}}\mathrm{e}^{\mathrm{i}\,\phi_i^{\text{data}}} \\ &= \left(a_i(\boldsymbol{x})\underbrace{\mathrm{e}^{\mathrm{i}\left[\phi_i(\boldsymbol{x},\boldsymbol{\alpha}) - \phi_i^{\text{data}}\right]}}_{\exp \mathrm{i}\phi_{\text{res},i}} - a_i^{\text{data}}\right)\mathrm{e}^{\mathrm{i}\,\phi_i^{\text{data}}} \\ &= \left(a_i(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\,\phi_{\text{res},i}} - a_i^{\text{data}}\right)\mathrm{e}^{\mathrm{i}\,\phi_i^{\text{data}}}\end{aligned}$$

D'où

$$\begin{aligned}
\mathcal{J}_{ell}^{i} &= \|y_{\mathrm{res},i}\|_{ell,i}^{2} \\
&= \left\| \left(a_i(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\,\phi_{\mathrm{res},i}} - a_i^{\mathrm{data}}\right)\mathrm{e}^{\mathrm{i}\,\phi_i^{\mathrm{data}}} \right\|_{ell,i}^{2} \\
&= \frac{\Re^2(a_i(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\,\phi_{\mathrm{res},i}} - a_i^{\mathrm{data}})}{\sigma_{\mathrm{rad},i}^2} + \frac{\Im^2(a_i(\boldsymbol{x})\mathrm{e}^{\mathrm{i}\,\phi_{\mathrm{res},i}} - a_i^{\mathrm{data}})}{\sigma_{\mathrm{tan},i}^2} \\
&= \frac{(a_i(\boldsymbol{x})\cos\phi_{\mathrm{res},i} - a_i^{\mathrm{data}})^2}{\sigma_{\mathrm{rad},i}^2} + \frac{(a_i(\boldsymbol{x})\sin\phi_{\mathrm{res},i})^2}{\sigma_{\mathrm{tan},i}^2} \\
&= \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{rad},i}^2}\cos^2\phi_{\mathrm{res},i} + \frac{a_{\mathrm{data},i}^2}{\sigma_{\mathrm{rad},i}^2} - 2\frac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma_{\mathrm{rad},i}^2}\cos\phi_{\mathrm{res},i} + \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{tan},i}^2}\sin^2\phi_{\mathrm{res},i}
\end{aligned} \tag{E.1}$$

### E.2.2 Calcul du gradient de $\mathcal{J}_{ell}^{i}$ par rapport à $\phi_{\mathrm{res},i}$

On obtient

$$\begin{aligned}
\frac{\partial \mathcal{J}_{ell}^{i}}{\partial \phi_{\mathrm{res},i}} &= -\frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{rad},i}^2}\sin 2\phi_{\mathrm{res},i} + 2\frac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma_{\mathrm{rad},i}^2}\sin\phi_{\mathrm{res},i} + \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{tan},i}^2}\sin 2\phi_{\mathrm{res},i} \\
&= \left[\frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{tan},i}^2} - \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{rad},i}^2}\right]\sin 2\phi_{\mathrm{res},i} + 2\frac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma_{\mathrm{rad},i}^2}\sin\phi_{\mathrm{res},i}
\end{aligned}$$

On pose

$$\begin{aligned}
w_{1,i} &= \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{tan},i}^2} - \frac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{rad},i}^2} \\
w_{2,i} &= 2\frac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma_{\mathrm{rad},i}^2}
\end{aligned}$$

On a donc :

$$\frac{\partial \mathcal{J}_{ell}^{i}}{\partial \phi_{\mathrm{res},i}} = w_{1,i}\sin 2\phi_{\mathrm{res},i} + w_{2,i}\sin\phi_{\mathrm{res},i} \tag{E.2}$$

### E.2.3 Calcul du gradient de $\mathcal{J}_{ell}$ par rapport à $\boldsymbol{\phi}_{\mathrm{res}}$

En posant

$$\begin{cases}
\boldsymbol{W}_1 \triangleq \mathrm{Diag}\left\{w_{1,i}\right\} = \mathrm{Diag}\left\{\dfrac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{tan},i}^2} - \dfrac{a_i^2(\boldsymbol{x})}{\sigma_{\mathrm{rad},i}^2}\right\} \\
\boldsymbol{W}_2 \triangleq \mathrm{Diag}\left\{w_{2,i}\right\} = \mathrm{Diag}\left\{2\dfrac{a_i(\boldsymbol{x})a_i^{\mathrm{data}}}{\sigma_{\mathrm{rad},i}^2}\right\}
\end{cases} \tag{E.3}$$

on obtient :

$$\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\phi}_{\mathrm{res}}} = \boldsymbol{W}_1\sin 2\boldsymbol{\phi}_{\mathrm{res}} + \boldsymbol{W}_2\sin\boldsymbol{\phi}_{\mathrm{res}} \tag{E.4}$$

La dérivée de $\boldsymbol{\phi}_{\mathrm{res}}$ par rapport à $\boldsymbol{\alpha}$ s'écrit :

$$\begin{aligned}
\frac{\partial \boldsymbol{\phi}_{\mathrm{res}}}{\partial \boldsymbol{\alpha}} &= \frac{\partial\left[\boldsymbol{\phi}(\boldsymbol{x}) + \bar{\boldsymbol{B}}\boldsymbol{\alpha} - \boldsymbol{\phi}^{\mathrm{data}}\right]}{\partial \boldsymbol{\alpha}} \\
&= \bar{\boldsymbol{B}}^{\mathrm{T}}
\end{aligned}$$

Le gradient de $\mathcal{J}_{ell}$ par rapport à $\boldsymbol{\alpha}$ de décompose en

$$\frac{\partial \boldsymbol{\phi}_{\text{res}}}{\partial \boldsymbol{\alpha}} \cdot \frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\phi}_{\text{res}}}$$

On obtient donc

$$\boxed{\frac{\partial \mathcal{J}_{ell}}{\partial \boldsymbol{\alpha}} = \bar{\boldsymbol{B}}^{\text{T}} \boldsymbol{W}_1 \sin 2\boldsymbol{\phi}_{\text{res}} + \bar{\boldsymbol{B}}^{\text{T}} \boldsymbol{W}_2 \sin \boldsymbol{\phi}_{\text{res}}} \tag{E.5}$$

### E.2.4 Conclusion : gradient de $\mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})$ par rapport à $\boldsymbol{\alpha}$

$$\begin{aligned}\frac{\partial \mathcal{J}_{WISARD}(\boldsymbol{\alpha}|\boldsymbol{x})}{\partial \boldsymbol{\alpha}} &= \sum_t \frac{\partial \mathcal{J}_{ell}(\boldsymbol{\alpha}(t)|\boldsymbol{x}, t)}{\partial \boldsymbol{\alpha}} \\ &= \sum_t \bar{\boldsymbol{B}}^{\text{T}} \boldsymbol{W}_1(t) \sin 2\boldsymbol{\phi}_{\text{res}}(t) + \bar{\boldsymbol{B}}^{\text{T}} \boldsymbol{W}_2(t) \sin \boldsymbol{\phi}_{\text{res}}(t)\end{aligned}$$

On note que la matrice $\bar{\boldsymbol{B}}$ ne dépend pas de l'instant de mesure, puisqu'elle ne dépend que du nombre de télescopes du réseau interférométrique.

# Annexe F

# Publications.

## F.1 JOSAA 2005

PROOF COPY [57938] 014511JOA



# Convex approximation to the likelihood criterion for aperture synthesis imaging

**Serge Meimon and Laurent M. Mugnier**

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

**Guy Le Besnerais**

*Office National d'Études et de Recherches Aérospatiales, Département Traitement de l'Information et Modélisation BP 72, F-92322 Châtillon cedex, France*

Received December 1, 2004; revised manuscript received April 1, 2005; accepted April 8, 2005

Aperture synthesis allows one to measure visibilities at very high resolutions by coupling telescopes of reasonable diameters. We consider the case where visibility amplitudes and phase are measured separately. It leads to an estimation problem where the noise model yields a nonconvex data-likelihood criterion. We show how to optimally approximate the noise model while keeping the criterion convex. This approximation has been validated both on simulations and on experimental data. © 2005 Optical Society of America

*OCIS codes:* 120.3180, 100.3020, 100.3190, 110.6770.

## 1. INTRODUCTION

Aperture synthesis allows one to reach very high angular resolution by coupling telescopes of reasonable diameters in an interferometric array. Because current interferometers do not provide images directly, the data have to be processed through an appropriate imaging software.

The basic observables of an interferometer are the complex visibilities extracted from each fringe pattern formed by the instrument. In the absence of noise, complex visibilities amplitudes and phases are corrupted by atmospheric path length fluctuations and by imperfect knowledge of the source position and of the interferometer geometry.

At radio wavelengths, it is usually possible to consider these errors as part of the noise and to use complex visibility amplitudes and phases directly. In contrast, at optical wavelengths, path length fluctuations due to atmospheric turbulence make visibility phases unexploitable. Thus, the observables of current interferometers at optical and infrared wavelengths are quantities independent of turbulent phases, such as squared visibilities and closure phases.

There are various ways of circumventing turbulence effects. A first one is obviously to locate the instrument where there is no turbulence, i.e., in space. In this case, complex visibilities are measurable. Second, if the u–v plane, i.e., the frequency coverage, is redundant enough, visibility phases can be estimated successfully from closure phases. This is the method used by Delage *et al.*[1] to form complex visibilities from experimental squared visibilities and closure phases. However, redundancy techniques reduce the frequency coverage. Another promising way of obtaining complex visibilities with an optical interferometer in the presence of turbulence is to use phase reference, as in the Very Large Telescope Interferometer (VLTI) instrument PRIMA (Phase-Referenced Imaging and Microarcsecond Astrometry).[2] This method will allow astronomers to measure complex visibilities without constraining the u–v coverage.

Finally, self-calibration algorithms,[3] first developed for radio interferometry, allow one to estimate both turbulent phases and the object by alternating turbulent-phase estimation steps with a known object and object reconstruction steps with known turbulent phases. The latter are strictly identical to Fourier synthesis problems without turbulence, i.e., to object reconstruction problems from noisy complex visibilities.

In this paper, we address object reconstruction from complex visibilities for both optical and radio wavelengths. The noise witnessed on complex visibilities yields a nonconvex data-likelihood criterion (Subsection 3.D), which makes reconstruction difficult.

After stating the interferometric data model that we consider (Section 3), we compute an optimal approximation of it that yields a quadratic data-likelihood criterion (Section 4). This approximation is then validated on simulations and used to process experimental data[1] (Section 5).

## 2. FOURIER SYNTHESIS

The basic observable of an interferometer is complex visibility, which can be measured from the fringe pattern obtained by combining the beams of two correctly phased telescopes. According to the Van Cittert–Zernike theorem,[4] complex visibilities are related to the sky brightness distribution $x(a,b)$ through a Fourier transform,

1084-7529/05/110001-0/$15.00 © 2005 Optical Society of America

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



$$V\left(\boldsymbol{\nu}=\begin{bmatrix}u\\v\end{bmatrix}\right)=\int\int x(a,b)\exp(-2\pi i(ua+vb))\mathrm{d}a\mathrm{d}b, \tag{1}$$

$a$ and $b$ being angular positions in the sky and $\boldsymbol{\nu}$ the 2D spatial frequency. For a couple of telescopes $(T_1,T_2)$, the spatial frequency $\boldsymbol{\nu}$ is given by $\vec{\nu}=(\vec{r_2}-\vec{r_1})/\lambda$, where $\vec{r_1}$ (resp. $\vec{r_2}$) denotes the position vector of $T_1$ (resp. $T_2$) projected onto a plane normal to the observation axis; $\vec{r_2}-\vec{r_1}$ is the corresponding baseline.

An interferometer is a device allowing one to measure the Fourier transform of an object at a set $\boldsymbol{\nu}$ of spatial frequencies. The aim of interferometry imaging is to retrieve the observed object from the set of measured Fourier samples. We adopt a Bayesian approach to solve this inverse problem, in which the first step is to design a data-formation model, both accurately fitting the actual physical process and yielding a tractable estimation problem.

## 3. RECONSTRUCTION MODEL

### A. Matrix Formulation

Let us suppose that the sky brightness $x(a,b)$ is discretized over a cardinal sine basis. It is thus represented by a vector of real coefficients $\mathbf{X}=[X_1,\ldots,X_j,\ldots,X_{N_p}]$, and Eq. (1) reads

$$V(\boldsymbol{\nu})=\sum_j^{N_p} h(j,\boldsymbol{\nu})X_j, \tag{2}$$

the $h(m,\boldsymbol{\nu})$ being complex coefficients.

We derive the following matrix formulation

$$\mathbf{V}(\mathbf{X})=\mathbf{H}\mathbf{X}, \tag{3}$$

with vector $\mathbf{V}$ and matrix $\mathbf{H}$ defined by

$$V_i=V(\boldsymbol{\nu}_i),$$

$$h_{i,j}=h(j,\boldsymbol{\nu}_i),$$

where $\boldsymbol{\nu}_i$ denotes the $i$th measurement spatial frequency.

### B. Noise Statistics

We consider that measured visibility moduli and phases follow Gaussian distributions. Although our method generalizes to any Gaussian distribution of the visibility moduli and phases, we will assume in this paper that the cross correlations are either not available or negligible. Then the measured visibilities $V_i^{\mathrm{meas}}$ are linked to the "true" ones $V_i$ by the following model:

$$|V_i^{\mathrm{meas}}|=|V_i(\mathbf{X})|+b_{\|,i},$$

$$\arg V_i^{\mathrm{meas}}=\arg V_i(\mathbf{X})+b_{\arg,i}, \tag{4}$$

with all the noises centered, decorrelated, and Gaussian. Let $\sigma_{\|,i}$ be the standard deviation of $b_{\|,i}$ and $\sigma_{\arg,i}$ the standard deviation of $b_{\arg,i}$. Model 4 applies to the output of an unstable radio interferometer.[3] In optical interferometry, it corresponds to the noise witnessed in various experimental settings where turbulence effects are either nonexistent or sufficiently corrected (see Section 1).

### C. Bayesian Estimation

Owing to the poor spectral coverage, the object reconstruction is an ill-posed inverse problem and must be regularized (see Refs. 5–7 for reviews on regularization) in the sense that some *a priori* information must be introduced in their resolution for the solution to be unique and robust to noise. In Bayesian estimation, the data likelihood $p(\mathrm{data}|\mathbf{X})$ is associated with a prior distribution $p(\mathbf{X})$. The maximum *a posteriori* estimation is obtained by maximizing the joint probability

$$p(\mathbf{X}|\mathrm{data})\propto p(\mathrm{data}|\mathbf{X})p(\mathbf{X})$$

or by minimizing the opposite of its logarithm:

$$-\log p(\mathbf{X}|\mathrm{data})=-\log p(\mathrm{data}|\mathbf{X})-\log p(\mathbf{X})+\mathrm{const}.$$

Hence, it reduces to the minimization of a compound criterion,

$$J=J_{\mathrm{data}}+\lambda J_{\mathrm{prior}}, \tag{5}$$

where $\lambda$ accounts for the confidence in the prior and is called the regularization parameter. With a Gaussian prior on $\mathbf{X}$, i.e., if we consider that the distribution of the object $\mathbf{X}$ is Gaussian, $J_{\mathrm{prior}}$ is quadratic.

Here, we focus on the data-likelihood term, which is yielded directly by the noise model:

$$J_{\mathrm{data}}(\mathbf{X})\propto-\log(p(\mathbf{V}^{\mathrm{meas}}|\mathbf{X})).$$

Taking into account data model (4), the data-likelihood term reduces to $J_1$:

$$J_1(\mathbf{X})=\sum_{i=0}^{N_b-1}\left(\frac{|V_i^{\mathrm{meas}}|-|V_i(\mathbf{X})|}{\sigma_{\|,i}}\right)^2+\sum_{i=0}^{N_b-1}\left(\frac{\arg V_i^{\mathrm{meas}}-\arg V_i(\mathbf{X})}{\sigma_{\arg,i}}\right)^2, \tag{6}$$

with $N_b$ the number of baselines for which the Fourier transform of the object is measured.

### D. A Nonconvex Criterion

The strict convexity of the criterion is a sufficient condition of uniqueness of its minimum and ensures good behavior of the classical minimization algorithms.[8] We show now that the functional $J_1$ of Eq. (6) is not convex. Because $\mathbf{V}$ is linked to $\mathbf{X}$ by a linear operator [see Eq. (3)], the convexity of $\tilde{J}_1$ defined by $\tilde{J}_1(\mathbf{V}(\mathbf{X}))=J_1(\mathbf{X})$ is equivalent to the convexity of $J_1$. Because $\tilde{J}_1$ is a sum of $N_b$ independent terms, we can deal with the case $N_b=1$ without loss of generality. Then $\mathbf{V}_{\mathrm{meas}}$ reduces to a complex number $z_0$ and $\mathbf{V}$ to a complex number $z$. $\tilde{J}_1$ reads

$$\tilde{J}_1(z)=\frac{(|z|-|z_0|)^2}{\sigma_{\|}^2}+\frac{(\arg z-\arg z_0)^2}{\sigma_{\arg}^2}.$$

The choice of $z_1=z_0\exp[(2i\pi)/3]$ and $z_2=z_0\exp[(-2i\pi)/3]$ yields $\tilde{J}_1(z_1)=(2\pi/3)^2/\sigma_{\arg}^2$, $\tilde{J}_1(z_2)=\tilde{J}_1(z_1)$, and $z_1+z_2=-z_0$. Hence, we get

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



$$\tilde{J}_1\left(\frac{z_1+z_2}{2}\right)=\frac{(|z_0|/2)^2}{\sigma_\parallel^2}+\frac{\pi^2}{\sigma_{\arg}^2},$$

$$\frac{1}{2}(\tilde{J}_1(z_1)+\tilde{J}_1(z_2))=0+\frac{4\pi^2/9}{\sigma_{\arg}^2}.$$

So $\tilde{J}_1[(z_1+z_2)/2]>\frac{1}{2}(\tilde{J}_1(z_1)+\tilde{J}_1(z_2))$, which contradicts the convexity of $\tilde{J}_1$ (actually, this example shows the nonconvexity of both the phase term and the modulus term of $\tilde{J}_1$).

## 4. AN EQUIVALENT ADDITIVE GAUSSIAN NOISE

In this section, we design an additive Gaussian approximation of the noise distribution, optimally "close" to the true one (in terms of a distance to be defined below), which yields a quadratic data-likelihood criterion. We first recall the true distribution; then we state the general shape of any complex Gaussian distribution, expressed in a convenient basis; and we conclude by selecting the parameters of the optimal distribution.

### A. Statement of the True Distribution

Once again, we have to study only the complex unidimensional problem, which is generalized without any difficulty. We consider the following model:

$$|z|=|z_0|+r,$$

$$\arg z=\arg z_0+\varphi. \tag{7}$$

Hence, $z=(|z_0|+r)\exp[i(\arg z_0+\varphi)]$ with $r$ and $\varphi$ following Gaussian centered distributions of variances $\mathrm{Var}(r)=\sigma_r^2$ and $\mathrm{Var}(\varphi)=\sigma_\varphi^2$. It is the model of Eqs. (4).

The probability distribution of $z$ is

$$p(z=(|z_0|+r)\exp[i(\arg z_0+\varphi)])=f(r,\varphi),$$

$$f(r,\varphi)\propto\exp\left[-\frac{1}{2}\left(\frac{r^2}{\sigma_r^2}+\frac{\varphi^2}{\sigma_\varphi^2}\right)\right]. \tag{8}$$

We want to approximate this distribution by an additive one. So, we have to recast model (7) into an additive one,

$$z=z_0+B, \tag{9}$$

and we choose to write $B$ as

$$B=(x+iy)\exp(i\arg z_0). \tag{10}$$

Identification of relations (8)–(10) yields

$$x=(|z_0|+r)\cos\varphi-|z_0|,$$

$$y=(|z_0|+r)\sin\varphi. \tag{11}$$

It is simple to see that $x$ and $y$ are the coordinates of $B$ in the Cartesian basis $(\mathbf{u}_x,\mathbf{u}_y)$, corresponding to the canonical $(\Re,\Im)$ one, rotated by angle $\arg z_0$ (see Fig. 1).

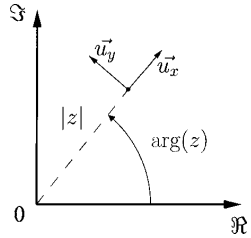



Fig. 1. Polar and Cartesian coordinate systems in $\mathbb{C}$.

### B. Statement of a Complex Gaussian Distribution

A complex noise is Gaussian if its vector representation in Cartesian coordinates is Gaussian. We choose the aforementioned Cartesian basis $(\mathbf{u}_x,\mathbf{u}_y)$. The change of basis is achieved by a rotation matrix $\mathcal{R}(\arg z_0)$, with

$$\mathcal{R}(\psi)=\begin{bmatrix}\cos\psi & \sin\psi\\ -\sin\psi & \cos\psi\end{bmatrix}. \tag{12}$$

Any additive Gaussian noise distribution can be written $p(z=z_0+(x+iy)\exp[i\arg z_0])=f_g(x,y)$, with

$$f_g(x,y)=\frac{1}{\sqrt{2\pi\det\Sigma}}\exp\left\{-\frac{1}{2}\begin{bmatrix}x-\bar{x}\\ y-\bar{y}\end{bmatrix}^t\Sigma^{-1}\begin{bmatrix}x-\bar{x}\\ y-\bar{y}\end{bmatrix}\right\} \tag{13}$$

where $\Sigma$ is a symmetric positive definite matrix.

We now compare Eq. (13) with the true distribution of $(x,y)$ stated in relations (8) and (11). APC: #1

### C. Kullback–Leibler Divergence Minimization

In order to choose the additive Gaussian distribution closest to the true one, we have to define a distance between two distributions. A convenient and well-known one is the Kullback–Leibler divergence, defined by

$$\delta(f_1,f_2)=\int f_1\log\left(\frac{f_1}{f_2}\right).$$

Note that technically this divergence is not a distance, because it is not symmetric. It is, however, often used as a discrepancy measure of $f_1$ with respect to (w.r.t.) $f_2$ because it is positive and equal to 0 only for $f_1=f_2$. $\delta(f_1,f_2)$ is the expectation of the "log distance" between two distributions $\log(f_1/f_2)$ w.r.t. the probability distribution $f_1$. To fit a Gaussian distribution $f_g$ on the true distribution $f$, it is therefore natural to minimize $\delta(f,f_g)$ rather than $\delta(f_g,f)$.

As proved in Appendix A, the minimization of $\delta(f,f_g)$ yields the following optimal parameters:

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



$$\bar{x} = \mathrm{E}_f\{x\} = |z_0|[\exp(-\sigma_\varphi^2/2) - 1],$$

$$\bar{y} = \mathrm{E}_f\{y\} = 0,$$

$$\Sigma = \mathrm{Diag}\{\sigma_1^2, \sigma_2^2\},$$

$$\sigma_1^2 = \mathrm{E}_f\{(x-\bar{x})^2\} = \frac{|z_0|^2 + \sigma_r^2}{2}[1 + \exp(-2\sigma_\varphi^2)] - |z_0|^2 \exp(-\sigma_\varphi^2),$$

$$\sigma_2^2 = \mathrm{E}_f\{(y-\bar{y})^2\} = \frac{|z_0|^2 + \sigma_r^2}{2}[1 - \exp(-2\sigma_\varphi^2)]. \tag{14}$$

APC: #2

The radial bias $\bar{x}$ can be estimated from $z$ and $\sigma_\varphi$ as $\bar{x} \approx |z|[\exp(-\sigma_\varphi^2/2) - 1]$. We shall denote $m$ the complex bias of coordinates $(\bar{x}, \bar{y})$.

### D. Two Gaussian Approximations

*1. Circular approximation*

This simple isotropic Gaussian approximation, inherited from radio imaging, is obtained by setting $\sigma_1$ and $\sigma_2$ in Eqs. (14) to the same value. Such an approximation is valid in radio imaging with stable interferometers and has also been used in optical interferometry.[9] However, it is not adapted to noise distributions in which the modulus standard deviation is different from the phase standard deviation, which is often the case in optical interferometry. We show here how to design an approximation specifically dedicated to process optical interferometry data.

*2. Optimal approximation*

Instead of a circular approximation, we propose a second-order expansion of the optimal Gaussian approximation stated in Eqs. (14); i.e., we consider that $\sigma_\varphi/2\pi$ and $\sigma_r/|z_0|$ are small w.r.t. 1:

$$\bar{x} = 0,$$

$$\bar{y} = \mathrm{E}_f\{y\} = 0,$$

$$\sigma_1^2 = \sigma_r^2,$$

$$\sigma_2^2 = |z_0|^2 \sigma_\varphi^2. \tag{15}$$

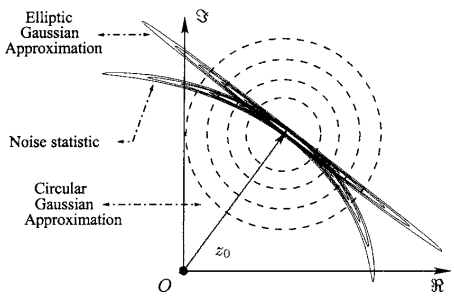



Fig. 2. Noise distribution contour lines, for $\sigma_r/|z_0| < \sigma_\varphi$.

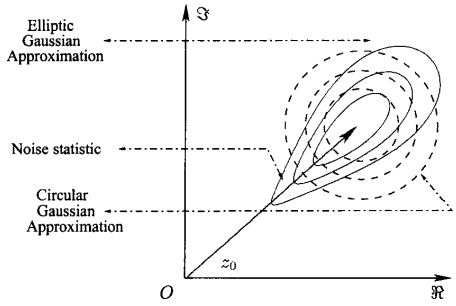



Fig. 3. Noise distribution contour lines, for $\sigma_r/|z_0| > \sigma_\varphi$.

*3. Why choose the optimal approximation?* APC: #3

The contours of the distribution of $z$ around $z_0$ are plotted in Figs. 2 and 3 for the true noise statistics for the optimal Gaussian approximation (more precisely, its second-order expansion) and for the circular one. In Fig. 3, the radial noise level, i.e., in the direction $\mathbf{u}_x$, given by $\sigma_r/|z_0|$ is greater than the one in the direction $\mathbf{u}_y$ given by $\sigma_\varphi$, whereas it is the opposite in Fig. 2.

For both configurations, these contour maps illustrate that our approximation fits the true distribution better.

### E. *N*-Dimensional Case

With our Gaussian approximation, the data likelihood for one measurement $V_0^{\mathrm{meas}}$ is

$$J_g(\mathbf{X}) = -2 \log f_g(V_0^{\mathrm{meas}} - V_0(\mathbf{X})).$$

With Eqs. (12) and (13), we get

$$J_g(\mathbf{X}) = \|V_0^{\mathrm{meas}} - V_0(\mathbf{X}) - m_0\|^2_{\Sigma_{0,\mathcal{R}}} \tag{16}$$

with

$$\Sigma_{0,\mathcal{R}} = \mathcal{R}(-\arg V_0^{\mathrm{meas}}) \Sigma \mathcal{R}(\arg V_0^{\mathrm{meas}}) \tag{17}$$

and

$$\|\cdot\|^2_{\Sigma_{0,\mathcal{R}}} l = \begin{bmatrix} \Re(\cdot) \\ \Im(\cdot) \end{bmatrix}^t \Sigma^{-1}_{0,\mathcal{R}} \begin{bmatrix} \Re(\cdot) \\ \Im(\cdot) \end{bmatrix}. \tag{18}$$

This expression can be easily generalized for $N$ measurements:

$$J_g(\mathbf{X}) = \sum_{i=0}^{N-1} \|V_i^{\mathrm{meas}} - V_i(\mathbf{X}) - m_i\|^2_{\Sigma_{i,\mathcal{R}}} \tag{19}$$

$$= \|\mathbf{V}^{\mathrm{meas}} - \mathbf{V}(\mathbf{X}) - \mathbf{m}\|^2_{\Sigma}, \tag{20}$$

$\Sigma$ being a bloc diagonal, with its blocks equal to the $\Sigma_{i,\mathcal{R}}$.

## 5. VALIDATION ON SIMULATIONS AND ON EXPERIMENTAL DATA

In this section, we compare circular approximation and our optimal Gaussian approximated distribution, denoted elliptic, in terms of reconstruction performance. To do so, we use either a circular or an elliptic noise model to build the data-likelihood term, which we associate with the same prior term (see Subsection 5.A.2) in a Bayesian reconstruction process.

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



Although our model clearly fits the noise distribution better, its performance is highly dependent on the noise outcome affecting the data. Hence, we will generate a hundred noise outcomes in order to assess the average gain induced by our approximation. We will then show that our method performs satisfactorily on real data.

### A. Simulations

*1. Simulated data*

The data that we process simulate VLTI measurements when observing an object corresponding to the model of the Ru Lupus Micro-jet developed by Garcia *et al.*[10] The frequency coverage [Fig. 5(b) below] has to be chosen rich enough to highlight the differences between reconstructions. Indeed, our method focuses on the data likelihood, which affects all the more the reconstruction quality as there are many data. The frequency coverage corresponds to six nights of observation of the same source with three telescopes of the VLTI, with 20 measurements each night. As already mentioned, we consider that the effects of turbulence are corrected enough to be included in the noise. The complex visibilities are corrupted by noise according to model (4), with $\sigma_{\|,i}=|V_i^{\mathrm{meas}}|\times 4.65\%$ and $\sigma_{\arg,i}=0.27$ rad for all $i$.

*2. Regularization and constraints*

We choose a Gaussian and shift-invariant prior distribution for $\mathbf{X}$,[11] so the distribution of its Fourier transform $\tilde{\mathbf{X}}$ is a Gaussian distribution with a diagonal covariance matrix, and the diagonal components are the values of the object power spectral density (PSD) $\mathrm{PSD}(\boldsymbol{\nu})$. Thus, the prior term reads

$$J_{\mathrm{prior}}(\mathbf{X})=\sum_{\boldsymbol{\nu}}\frac{|\tilde{\mathbf{X}}(\boldsymbol{\nu})-\tilde{\mathbf{X}}_m(\boldsymbol{\nu})|^2}{\mathrm{PSD}(\boldsymbol{\nu})}.$$

The mean object $\tilde{\mathbf{X}}_m$ is assumed to be constant, with its flux equal to the measured flux, i.e., the null frequency measured visibility.

The PSD model chosen is the function

$$\mathrm{PSD}(\boldsymbol{\nu})=\frac{K}{(|\boldsymbol{\nu}|/\rho_0)^p+1}.$$

The parameters $K$, $p$, and $\rho_0$ are estimated by a maximum likelihood on the data.

As noted Ref. 11, $K$ plays the role of the regularization parameter $\lambda$ [see Eq. (5)] and can be estimated directly from the data. Thus, the method is completely unsupervised, i.e., *no parameter has to be set by the user*.

Reconstruction uses a BFGS method (Broyden–Fletcher–Goldfarb–Shanno) software OP-VMLM, designed by Eric Thiébaut[13] and is performed under a positivity constraint. To compare the circular approximation with our method, we compute for each noise outcome an improvement of the root mean square error (IRMSE) in decibels (dB). A positive IRMSE means a better reconstruction with our method. Figure 4 shows the IRMSE repartition histogram for the 100 noise outcomes.

The improvement is 4 dB on average, and 95% of the reconstructions have an IRMSE of more than 2 dB. We can conclude that our elliptic approximation performs much better than the circular one in terms of reconstructed image quality.

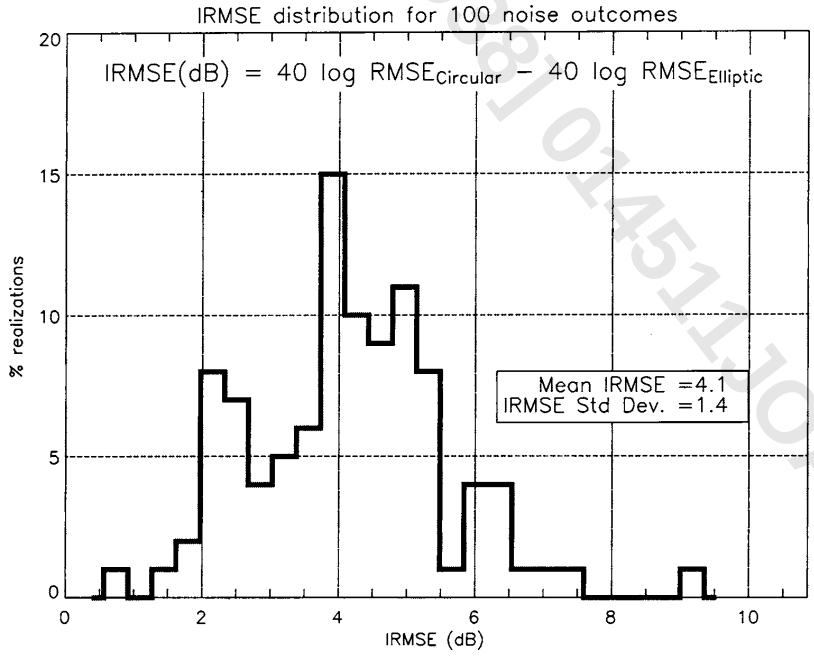



Fig. 4. IRMSE repartition histogram

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



As mentioned before, reconstructions are performed with $\lambda = 1$. To measure the influence of $\lambda$ on the IRMSE, we have processed the same data with $\lambda = 0.1$ and $\lambda = 10$. Table 1 provides IRMSE means and standard deviations over the 100 reconstructions for different $\lambda$.

For a variation of a decade around the nominal $\lambda$ value, we still witness a clear reconstruction improvement with our method.

*3. Reconstructions*

To further illustrate the interest of using our method, we show in Fig. 5 typical reconstructions for both methods:

**Table 1. Influence of Regularization Parameter on IRMSE**

| $\lambda$ | Mean (IRMSE) | Std. Dev. (IRMSE) |
|---|---|---|
| 0.1 | 2.6 | 1.8 |
| 1 | 4.1 | 1.4 |
| 10 | 7.8 | 0.9 |

we have selected among 100 noise outcomes the one yielding an IRMSE close to the mean value. Our method obviously helps reducing the noise, yielding an IRMSE worth 4 dB on average.

In next section, we process real experimental data in order to demonstrate the efficiency of our method with realistic noise distribution and frequency coverage.

### B. Validation on Experimental Data

*1. Experimental setup*

Experimental data were graciously supplied by Laurent Delage and François Reynaud and correspond to the experiment described in Ref. 1. The object is made of four stars of various magnitudes and is observed through a fiber link interferometer featuring 61 frequency measurements. The data model used corresponds to system (4), because only the standard deviations of measurements are provided.

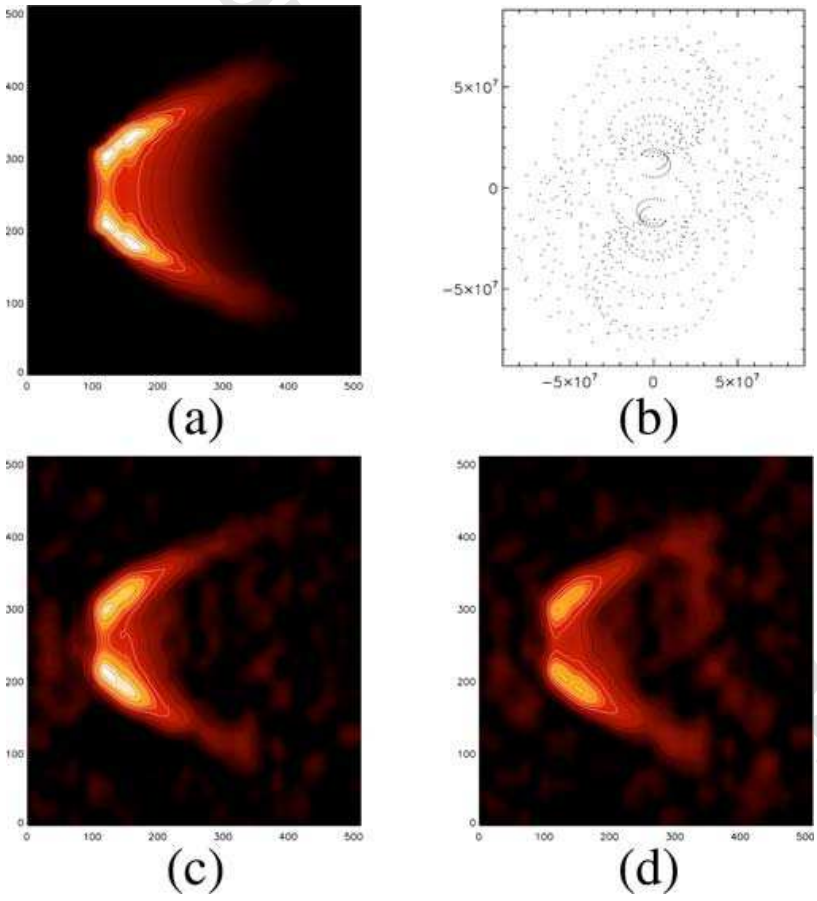



Fig. 5. Simulation results. (a) true object, (b) u–v coverage—360 frequencies, (c) reconstruction with elliptic approximation, (d) reconstruction with circular approximation. 256×256 pixels. Pixel size 0.2 marcsec.

APC: #13

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



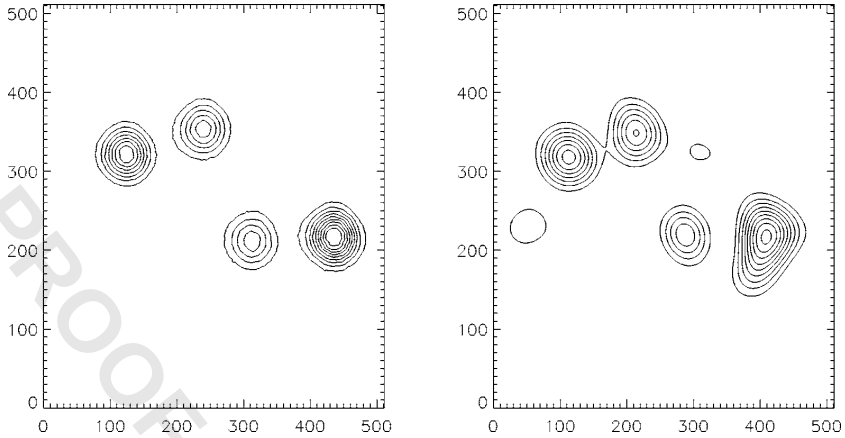



Fig. 6. True object (left) and restored one (right). Contour levels: 10%, 20%,...,100% of the maximum.

**Table 2. Relative Positions and Flux of the Three Faintest Stars w.r.t. the Brightest One**

| | Position Error (w.r.t. main star diameter $D$) | | Intensity ($-2.5\log_{10}$ flux) | |
|---|---|---|---|---|
| Star # | $\Delta x/D$ | $\Delta y/D$ | True | Reconstructed |
| 1 | 0 | 0 | 1.33 | 1.07±19% |
| 2 | 0.13 | 0.05 | 1.89 | 1.72±9% |
| 3 | 0.03 | 0.06 | 2.20 | 1.95±11% |
| 4 | 0.01 | 0.08 | 2.44 | 2.41±2% |

*2. Regularization*

Reconstructions are done under the positivity constraint. We also use the quadratic regularization term described in Subsection 5.A.2.

*3. Reconstruction*

APC: #4 Figure 6 shows the contour maps of the true object and the restored one. The four structuring elements are correctly reconstructed, although quadratic regularization slightly oversmoothed them. Table 2 shows that our reconstruction is correct in terms of relative positions of the peaks. We here validate that our method is efficient and usable on experimental data.

## 6. CONCLUDING COMMENTS

We have designed an accurate data-likelihood criterion that closely mimics the noise model while keeping the criterion convex. Our method performed satisfactorily both on simulated data and on experimental material. However, more sophisticated regularization should be investigated. Additionally, this paper did not address how to deal with closure phases instead of visibility phases. This can be done by using "self-calibration" methods, which alternate transfer function estimation steps with object reconstruction steps.[9,13] We are currently developing an original self-calibration procedure that uses the likelihood approximation techniques developed in this paper.[14]

## APPENDIX A: KULLBACK–LEIBLER DISTANCE OPTIMIZATION

We show here that for any given distribution $f(\mathbf{X})$, the Gaussian distribution defined by

$$g(\mathbf{X}) = \frac{1}{\sqrt{2\pi \det \Sigma}} \exp\left[-\frac{1}{2}P(\mathbf{X})\right],$$

with $\Sigma$ a symmetric positive definite matrix and

$$P(\mathbf{X}) = (\mathbf{X}-\bar{\mathbf{X}})^t\Sigma^{-1}(\mathbf{X}-\bar{\mathbf{X}}),$$

which reaches the minimum of the Kullback–Leibler distance $\delta(f,g)$, is such that

$$\begin{aligned}\bar{\mathbf{X}} &= \mathrm{E}_f\{\mathbf{X}\},\\ \Sigma &= \mathrm{E}_f\{(\mathbf{X}-\mathrm{E}_f\{\mathbf{X}\})(\mathbf{X}-\mathrm{E}_f\{\mathbf{X}\})^t\}\\ &= \mathrm{Var}(\mathbf{X}).\end{aligned} \tag{A1}$$

This property may result from general results of the probability theory, but we provide here a compact and self-contained proof.

**1. Definition of the Kullback–Leibler Distance**

The distance $\delta(f,g)$ is defined by

$$\delta(f,g) = \int f(\mathbf{X})\log\frac{f(\mathbf{X})}{g(\mathbf{X})}d\mathbf{x} = -\mathrm{E}_f\{\log g\} + \text{const.}$$

So APC: #6

$$\delta(f,g) = \tfrac{1}{2}(\mathrm{E}_f\{P(\mathbf{X})\} + \log\det\Sigma) + \text{const.}$$

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



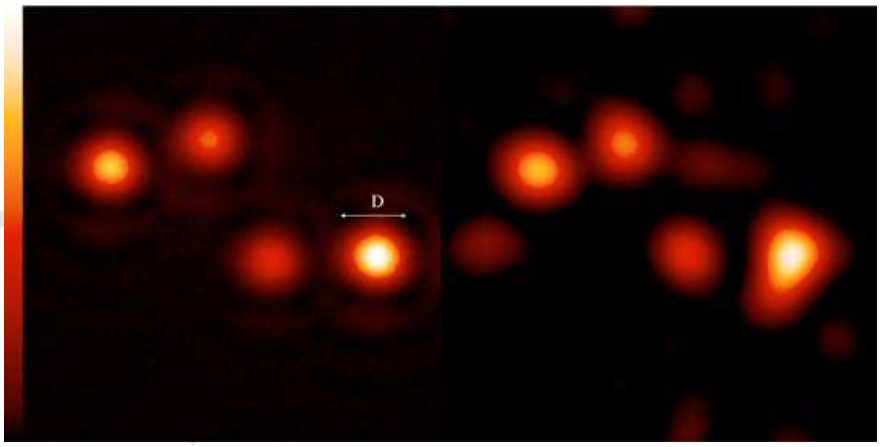



Fig. 7. True object (left) and restored one (right). D is the diameter of the main star used in Table 2.

APC: #7

### 2. First-Order Terms

$$\frac{\partial \delta(f,g)}{\partial \bar{X}} = 0$$

$$\Rightarrow \frac{\partial \mathrm{E}_f\{P(\mathbf{X})\}}{\partial \bar{X}} = 0$$

$$\Rightarrow \mathrm{E}_f\left\{\frac{\partial P(\mathbf{X})}{\partial \bar{X}}\right\} = 0$$

$$\Rightarrow \mathrm{E}_f\{2\Sigma^{-1}(\mathbf{X}-\bar{\mathbf{X}})\} = 0$$

$$\Rightarrow \mathrm{E}_f\{(\mathbf{X}-\bar{\mathbf{X}})\} = 0$$

$$\Rightarrow \mathrm{E}_f\{\mathbf{X}\} = \bar{\mathbf{X}}.$$

APC: #8

### 3. Second-Order Terms

$$\frac{\partial \delta(f,g)}{\partial \Sigma} = 0$$

$$\Rightarrow \frac{\partial}{\partial \Sigma}[\mathrm{E}_f\{P\} + \log\det\Sigma] = 0$$

$$\Rightarrow \mathrm{E}_f\left\{\frac{\partial P}{\partial \Sigma}\right\} + \frac{\partial \log\det\Sigma}{\partial \Sigma} = 0$$

$$\Rightarrow \mathrm{E}_f\{-\Sigma^{-t}(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})^t\Sigma^{-t}\} + \Sigma^{-t} = 0.$$

Because $\Sigma$ is symmetric, we have: $\Sigma^{-t}=\Sigma^{-1}$, so

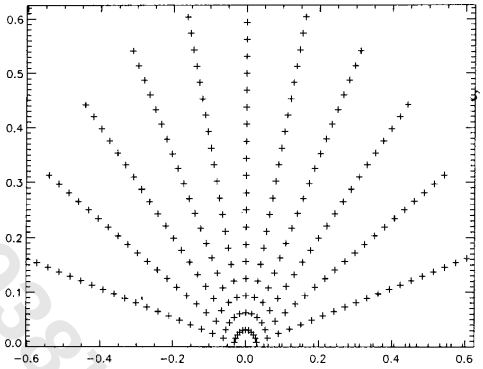



Fig. 8. Experimental frequency coverage.

$$\frac{\partial \delta(f,g)}{\partial \Sigma} = 0$$

$$\Rightarrow \Sigma^{-t}\mathrm{E}_f\{(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})^t\}\Sigma^{-t} = \Sigma^{-t}$$

$$\Rightarrow \mathrm{E}_f\{(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})^t\} = \Sigma\Sigma^{-t}\Sigma$$

$$\Rightarrow \Sigma = \mathrm{E}_f\{(\mathbf{X}-\bar{\mathbf{X}})(\mathbf{X}-\bar{\mathbf{X}})^t\}$$

$$\Rightarrow \Sigma = \mathrm{Var}(\mathbf{X}),$$

which concludes the proof.

### 4. Two-Dimensional Case

$$g(x,y) = \frac{1}{\sqrt{2\pi\det\Sigma}}\exp -\frac{1}{2}P,$$

$$P(\bar{x},\bar{y}) = \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}^t \Sigma^{-1} \begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}$$

is such that

PROOF COPY [57938] 014511JOA

PROOF COPY [57938] 014511JOA



$$\bar{x} = \mathrm{E}_f\{x\},$$

$$\bar{y} = \mathrm{E}_f\{y\},$$

$$\Sigma = \mathrm{E}_f\left\{\begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}\begin{bmatrix} x-\bar{x} \\ y-\bar{y} \end{bmatrix}^t\right\}. \quad \text{(A2)}$$

## ACKNOWLEDGMENTS

The data processed in Subsection 5.B were supplied by Laurent Delage and François Reynaud. The authors express their special thanks to Eric Thiébaut for his support and for letting them use his minimization software.

Corresponding author Serge Meimon can be reached at Serge.Meimon@onera.fr.

## REFERENCES

1. L. Delage, F. Reynaud, and E. Thiébaut, "Imaging laboratory test on a fiber linked telescope array," Opt. Commun. **160**, 27–32 (1999).
2. A. Quirrenbach *et al.*, "PRIMA: study for a dual beam instrument for the VLT interferometer," in *Astronomical Interferometry*, 3350 (1998).
3. T. J. Cornwell and P. N. Wilkinson, "A new method for making maps with unstable radio interferometers," Mon. Not. R. Astron. Soc. **196**, 1067–1086 (1981).
4. J. W. Goodman, *Statistical Optics* (Wiley, 1985).
5. D. M. Titterington, "General structure of regularization procedures in image reconstruction," Astron. Astrophys. **144**, 381–387 (1985).
6. J. Idier, ed., *Approche Bayésienne pour les Problèmes Inverses* (Hermès, 2001).
7. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024–2036 (1989).
8. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C* (Cambridge U. Press, 1988).
9. A. Lannes, "Weak-phase imaging in optical interferometry," J. Opt. Soc. Am. A **15**, 811–824 (1998).
10. P. J. V. Garcia, S. Cabrit, J. Ferreira, and L. Binette, "Atomic T Tauri disk winds heated by ambipolar diffusion. II. Observational tests," Astron. Astrophys. **377**, 609–616 (2001).
11. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images using object and point spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
12. E. Thiébaut, "Optimization issues in blind deconvolution algorithms," in *Astronomical Data Analysis*, 4847 (2002).
13. G. Le Besnerais, "Méthode du maximum d'entropie sur la moyenne, critères de reconstruction d'image et synthèse d'ouverture en radio-astronomie," Thèse de doctorat (Université de Paris-Sud, Orsay, 1993).
14. S. C. Meimon, L. M. Mugnier, and G. Le Besnerais, "A novel method of reconstruction for weak-phase optical interferometry," in *New Frontiers in Stellar Interferometry*, W. A. Traub, ed., Proc. SPIE 5491, 909–919 (2004).

APC: #9
APC: #10
APC: #12

PROOF COPY [57938] 014511JOA

## F.2 Optics Letters 2005

July 15, 2005 / Vol. 30, No. 14 / OPTICS LETTERS 1809

# Reconstruction method for weak-phase optical interferometry

**Serge C. Meimon and Laurent M. Mugnier**

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

**Guy Le Besnerais**

*Département Traitement de l'Information et Modélisation, Office National d'Études et de Recherches Aérospatiales, BP 72, F-92322 Châtillon Cedex, France*

Received January 24, 2005

Current optical interferometers are affected by unknown turbulent phases on each telescope. We account for this lack of phase information by introducing system aberration parameters, and we solve the image reconstruction problem by minimizing an original joint criterion in the aberrations and in the object. We validate this method by means of simulations. Tests on experimental data are under way. © 2005 Optical Society of America

*OCIS codes:* 120.3180, 100.3020, 100.3190, 010.1330, 110.6770.

Optical interferometry allows one to reach the angular resolution that a 100 m telescope would provide using several 10 m telescopes. The interferograms of current instruments are affected by turbulence, which corrupts the recorded object phases, so one is led to form quantities that are turbulence independent such as phase closures. To cope with the missing phase information, we introduce phase calibration parameters to be estimated jointly with the observed object, and we propose a weak-phase interferometric sample alternating reconstruction device (WISARD) to perform this estimation. This algorithm combines, within a Bayesian framework, an alternating estimation of the object and phase parameters (in the spirit of self-calibration algorithms proposed by radioastronomers,[1] a recently developed noise model suited to optical interferometry data,[2] and an edge-preserving regularization[3] to deal with the sparsity of the data typical of optical interferometry. Our work is dedicated to processing data from three-telescope interferometers, yet the method can be easily extended to data from interferometers of more than three telescopes. We consider throughout this Letter an interferometric array of three telescopes $T_1$, $T_2$, and $T_3$, which are pointed at the same monochromatic source at $n_a$ different moments. The three frequencies measured at one moment $i$ are given by

$$\nu^i(12) = \frac{\overline{T_1}\widehat{T_2}}{\lambda}, \qquad \nu^i(23) = \frac{\overline{T_2}\widehat{T_3}}{\lambda}, \qquad \nu^i(31) = \frac{\overline{T_3}\widehat{T_1}}{\lambda}.$$

The basic observable of an interferometer without turbulence is the complex visibility, which can be measured from the fringe pattern obtained for each couple of telescopes belonging to the same array. We will call $V_0^i(12)$ the visibility corresponding to frequency $\nu^i(12)$ and arrange all these visibilities in a vector $\mathbf{V}_0$. With $X$ the unknown object and $H$ the Fourier operator, we have

$$\mathbf{V}_0(X) = HX. \tag{1}$$

We also define the visibility amplitudes and phases as $A = |HX|$ and $\varphi_0 = \arg HX$, respectively.

When there is atmospheric turbulence, the instantaneous complex visibilities are affected by random optical path differences $\alpha^i(j)$ at each telescope $T_j$. One cannot measure $\mathbf{V}_0$, only its turbulence-corrupted value $\mathbf{V}$:

$$\begin{aligned} V^i(12) &= V_0^i(12)\exp i[\alpha^i(2) - \alpha^i(1)], \\ V^i(23) &= V_0^i(23)\exp\{i[\alpha^i(3) - \alpha^i(2)]\}, \\ V^i(31) &= V_0^i(31)\exp\{i[\alpha^i(1) - \alpha^i(3)]\}. \end{aligned} \tag{2}$$

If we define $\varphi = \arg \mathbf{V}$, it follows that

$$\begin{aligned} \varphi^i(12) &= \varphi_0^i(12) + \alpha^i(2) - \alpha^i(1), \\ \varphi^i(23) &= \varphi_0^i(23) + \alpha^i(3) - \alpha^i(2), \\ \varphi^i(31) &= \varphi_0^i(31) + \alpha^i(1) - \alpha^i(3). \end{aligned} \tag{3}$$

Actually, even the corrupted visibility phases $\varphi = \arg \mathbf{V}$ are difficult to measure, because they fluctuate in a fast and random way. However, the visibility amplitudes are not affected and can still be measured: $A = |\mathbf{V}_0| = |\mathbf{V}|$. Moreover, we find that the turbulent phasers cancel out[4] in

$$\begin{aligned} \mathcal{C}^i &= \arg V^i(12) + \arg V^i(23) + \arg V^i(31), \\ &= \varphi^i(12) + \varphi^i(23) + \varphi^i(31), \\ &= \varphi_0^i(12) + \varphi_0^i(23) + \varphi_0^i(31), \end{aligned}$$

or in a vector formulation $\mathcal{C} = \Omega\mathbf{V}_0 = \Omega\mathbf{V}$. $\Omega$ is the so-called closure operator.

0146-9592/05/141809-3/$15.00 © 2005 Optical Society of America

**1810** OPTICS LETTERS / Vol. 30, No. 14 / July 15, 2005

We will now design a metric that expresses the likelihood of the observables as a function of $X$ and the noise statistics. The observable are time-averaged closures $\mathcal{C}_{mes}$ and visibility amplitudes $A_{mes}$ (actually, most interferometers provide squared visibilities from which we extract amplitudes). We assume here that only standard deviation on closures and amplitudes will be supplied and that all the noises are uncorrelated, zero mean, and Gaussian. The model consequently reads

$$\mathcal{C}_{mes} = \mathbf{\Omega}(HX) + \text{Gaussian noise},$$

$$A_{mes} = |HX| + \text{Gaussian noise}. \tag{4}$$

The data likelihood yielded by this model is not convex[5] and has local minima, which makes minimization difficult.[6] Yet, there is another way to state the data model: It is possible to account for missing phase information through myopic aberration parameters. Because of random delays $\alpha$ we measure one closure per instant, instead of three phases. In other words, two pieces of information per triple are missing. We can account for this missing information by introducing two aberration parameters in the system.

Let us consider system (3). We suppose we have measured virtually all the corrupted phases $\varphi$ and introduce the turbulent optical path delays through unknown aberration parameters $\beta$, which generate the null space of the closure operator $\Omega$. If we define $\varphi = \arg \mathbf{V}$, it follows that

$$\begin{aligned}
\varphi^i_{mes}(12) &= \varphi^i_0(12) + \beta^i_1 + \text{Gaussian noise},\\
\varphi^i_{mes}(23) &= \varphi^i_0(23) + \beta^i_2 + \text{Gaussian noise},\\
\varphi^i_{mes}(31) &= \varphi^i_0(31) - \beta^i_1 - \beta^i_2 + \text{Gaussian noise},
\end{aligned} \tag{5}$$

$$\varphi_{mes} = \arg HX + \beta + \text{Gaussian noise}. \tag{6}$$

If we set $\varphi^i_{mes}(12) = \varphi^i_{mes}(23) = \varphi^i_0(23) = \mathcal{C}^i_{mes}/3$, the sum of the three lines of system (5) is equal to

$$\mathcal{C}^i_{mes} = \varphi^i_0(12) + \varphi^i_0(23) + \varphi^i_0(31) + \text{Gaussian noise},$$

which is exactly the first equation of system (4). With Eq. (6) and the second equation of system (4), we define

$$\mathbf{V}_{mes} = A_{mes} e^{i\varphi_{mes}},$$

$$\mathcal{V}(X,\beta) = |HX| \exp[i(\arg HX + \beta)], \tag{7}$$

and gather $\mathbf{V}_{mes} = \mathcal{V}(X,\beta) + N$, with $N$ being a complex noise. $\mathcal{V}(X,\beta)$ is the noiseless $\beta$-corrupted complex data model. Myopic data model (7) is equivalent to direct data model (4).

We have shown[5] that model (7) yields a nonconvex data-likelihood criterion. Although it is possible to use circular approximations of the noise distribution $N$ (Ref. 7) that are convex, they are not well adapted for optical interferometry. We hence use the elliptic convex approximation described in Ref. 5, which better fits the noise statistics and can be stated as

$$J_{ell}(X,\beta) = \|\mathbf{V}_{mes} - \mathcal{V}(X,\beta)\|^2_{C_b}. \tag{8}$$

It can be easily shown that the criterion $J_{ell}(X,\beta)$ is the sum of $n_a$ terms, each involving only the measurements obtained at instant $i$. Hence, each $J^i_{ell}$ depends only on two aberration parameters $\beta^i_1$ and $\beta^i_2$:

$$J_{ell}(X,\beta) = \sum_{i=1\ldots n_a} J^i_{ell}(X,\beta^i_1,\beta^i_2). \tag{9}$$

This is a consequence of the fact that, by definition, if the time between two instants is much greater than the turbulence coherent time (around 10 ms), aberrations at two different instants are statistically independent. Hence, $J_{ell}(X,\beta)$ is convex with respect to $X$, separable with respect to $\beta$, and designed from an approximated noise model that accurately fits the true one.

Here, we describe our interferometric data reduction algorithm WISARD (for three or four pupil array cases). Although the criterion we designed is convex for given aberrations, it is not convex for the whole unknown set, i.e., for both aberrations and objects. To obtain a good solution with gradient-based minimization algorithms, it is useful, as empirically witnessed, on the one hand to estimate a reasonable initial guess from the visibility data and on the other hand to incorporate the data gradually, starting with low frequencies and ending with high frequencies. To make use of the remarkable properties of $J_{ell}(X,\beta)$, we minimize it following an alternating pattern; i.e., we optimize it given aberrations with respect to the object and then optimize the aberrations for the current object.

The object step is a convex functional minimization under a positivity constraint. It is performed by Broyden–Fletcher–Goldfarb–Shanno method software, OP-VMLMB designed by Eric Thiébaut.[8] Since $J_{ell}(X,\beta)$ is separable for a given $X$, instead of optimizing for $2n_a$ aberration parameters, we perform $n_a$ separate optimization of two parameters that can be done simply by an exhaustive search on a fine grid. The pattern of WISARD is described in Fig. 1.

Due to the poor spectral coverage, the object reconstruction, even with known aberrations, is an ill-posed inverse problem and must be regularized (see Refs. 9 and 10 for reviews on regularization). Whereas quadratic regularization tends to soften the

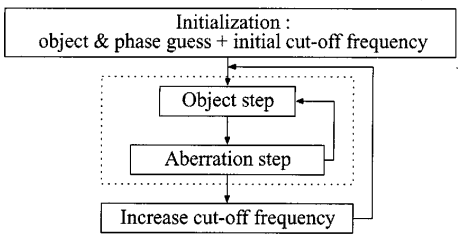



Fig. 1. WISARD algorithm loop.

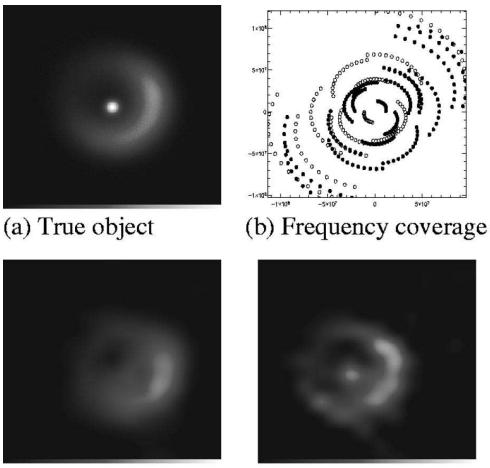



Fig. 2. Original object and reconstruction maps.

edges of the reconstructed map, a quadratic-linear regularization (or L1–L2) is a good trade-off between obtaining a clean image and retrieving peaks.[11,12] Among the several versions of an L1–L2 functional, we use the isotropic criterion developed by one of the authors.[3] We took part in an international blind reconstruction contest, which aimed at comparing the performance of five different algorithms designed for synthesis imaging. Only the L2-regularized map [Fig. 2(c)] was submitted to this contest,[13] and our method has evolved toward a less-case-dependent procedure since then. We present here the new results obtained by using L1–L2 regularization.

The data sets were produced by Christian Hummel using the data reduction software OYSTER[14] and simulating a six-station Navy prototype optical interferometer (NPOI). The image of the star with an asymmetric shell shown in Fig. 2 was provided to Christian Hummel by Peter Tuthill. As shown in Fig. 2, we retrieved satisfactorily the global structures of both objects. The central dot is not reconstructed with the quadratic regularization but is retrieved with L1–L2 regularization, although it is softer and slightly wider than the original.

Although the reconstructed maps are satisfactory, we hope to get even closer by adapting WISARD to arrays of more than three telescopes. Another crucial aspect is the extensive study of the criterion shape, mainly its behavior as a function of UV coverage. Finally, we hope we can further improve the reconstructions by using other regularization methods.

We express our special thanks to Eric Thiébaut for fruitful discussions, for his support, and for letting us use his minimization software. We are grateful to Peter Lawson for coordinating the beauty contest. This work was partially supported by a European Commission Joint Research Action under contract RII3-CT-2004-001566. S. C. Meimon's e-mail address is meimon@onera.fr.

**References**

1. T. J. Cornwell and P. N. Wilkinson, Mon. Not. R. Astron. Soc. **196**, 1067 (1981).
2. S. Meimon, L. M. Mugnier, and G. Le Besnerais, "A convex approximation of the likelihood in optical interferometry," J. Opt. Soc. Am. A (to be published).
3. L. M. Mugnier, T. Fusco, and J.-M. Conan, J. Opt. Soc. Am. A **21**, 1841 (2004).
4. F. Roddier, Opt. Commun. **60**, 145 (1986).
5. S. Meimon, L. Mugnier, and G. Le Besnerais, in *19ième Colloque sur le Traitement du Signal et des Images*, J.-M. Chassery and C. Jutten, eds. (GRETSI, 2003).
6. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes in C* (Cambridge U. Press, 1988).
7. A. Lannes, J. Opt. Soc. Am. A **15**, 811 (1998).
8. E. Thiebaut, Proc. SPIE **4847**, 174 (2002).
9. D. M. Titterington, Astron. Astrophys. **144**, 381 (1985).
10. G. Demoment, IEEE Trans. Acoust., Speech, Signal Process. **37**, 2024 (1989).
11. C. Bouman and K. Sauer, IEEE Trans. Image Process. **2**, 296 (1993).
12. P. J. Green, IEEE Trans. Med. Imaging **9**, 84 (1990).
13. P. R. Lawson, W. D. Cotton, C. A. Hummel, J. D. Monnier, M. Zhao, J. S. Young, H. Thorsteinsson, S. Meimon, L. Mugnier, G. Le Besnerais, E. Thiébaut, and P. G. Tuthill, Proc. SPIE **5491**, 886 (2004).
14. http://www.sc.eso.org/~chummel/oyster/oyster.html.

# Table des figures


1 Deux types d'interféromètres optiques (d'après Faucherre *et al.* (1989)). . . . . 6

1.1 L'expérience des trous d'Young. . . . . 11
1.2 Diffraction de Fraunhofer à l'infini. . . . . 12
1.3 Influence de la configuration pupillaire sur la formation d'images[Cassaing (1997)]. . . . . 16
1.4 Pupilles, fonction de transfert et réponse impulsionnelle. . . . . 17
1.5 Un interféromètre à deux télescopes . . . . . 19
1.6 Modulation des franges par la dispersion spectrale, pour $2\Delta k/\bar{k} = 3\%$ et $2\Delta k/\bar{k} = 6\%$ . . . . . 22
1.7 (a) Positions et bases, (b) Bases et rotation de la Terre, (c) Couverture fréquentielle. . . . . 24
1.8 Effet du repliement de phase. . . . . 26
1.9 Régularisation entropique. . . . . 35
1.10 Comparaison des fonctions de coût $-f1$ et $f_{L1L2}$. . . . . 42

2.1 Courbes iso-valeur de $j_i^{\text{data}}(y)$, pour $\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$. . . . . 46
2.2 Bases orthonormées du plan complexe $\mathbb{C}$. . . . . 47
2.3 Courbes isocritère pour les modèles , pour de bruit optique ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$) et radio ($\sigma_{y,i} = \sqrt{a^d \sigma_{a,i}\sigma_{\phi,i}}$). . . . . 49
2.4 Interprétation graphique des écarts-type de l'approximation optimale développée au premier ordre ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$). . . . . 53
2.5 Courbes isocritère pour les distributions « vraie », « circulaire » et « optimale » ($\sigma_{a,i} = 0{,}1$ et $\sigma_{\phi,i} = 0{,}6\,rad$). . . . . 54
2.6 Modèle de DSP. . . . . 60
2.7 Modèle de Micro-jet de l'étoile Ru Lupus (Garcia *et al.*, 2001) . . . . . 62
2.8 Couverture fréquentielle obtenue avec 3 des télescopes du VLTI, pour 6 nuits d'observation de la même source, avec 20 points de mesure par nuit. . . . . 63
2.9 Histogramme de répartition de l'AEQM . . . . . 64
2.10 Résultats de simulation. A gauche, reconstruction avec approximation elliptique, à droite avec approximation circulaire. Meilleur cas (AEQM=9,35dB) : (a) et (b) ; Pire cas (AEQM=0,56dB) : (c) et (d) ; Cas moyen (AEQM=4,16dB) : (e) et (f). $256 \times 256$ pixels, de 0.2 mas. de côté. . . 66
2.11 Dispositif expérimental. Delage *et al.* (1999) . . . . . 67
2.12 Couverture fréquentielle expérimentale. . . . . 68
2.13 Configuration pupillaire. . . . . 68
2.14 Objet vrai (gauche) et objet reconstruit (droite). Niveaux des contours : $10\%, 20\%, \ldots, 100\%$ du maximum. . . . . 70
2.15 Objet vrai (gauche) et objet reconstruit (droite). D est le diamètre de la composante principale, utilisé dans le tableau 2.2. . . . . 70

3.1 Statistique $p$ d'une variable aléatoire $a$ de carré $s$ gaussien et approximation gaussienne $p_g$. . . . . 76

4.1 Boucle d'auto-calibration . . . . . 96
4.2 Influence sur le critère d'un tip-tilt global . . . . . 102
4.3 Critère $\mathcal{J}_{ell}$ en fonction des deux aberrations myopes à déterminer. . . . . 105
4.4 Visibilités absolues en fonction de la fréquence radiale. . . . . 108
4.5 Initialisation circulaire. . . . . 109
4.6 Reconstruction par fréquences croissantes. . . . . 109
4.7 Structure de l'algorithme WISARD. . . . . 111
4.8 Estimations aux moindres carrés avec un modèle équivalent à 3 télescopes. . . . 113

5.1 Reconstruction avec WISARD pour la phase préliminaire de l'Imaging Beauty Contest. $60mas \times 60mas$. . . . . 116
5.2 Erreur de reconstruction due au champ reconstruit. . . . . 117
5.3 Objet vrai et couverture $u-v$ pour le fichier de données 1. Ces données simulent l'observation d'un objet correspondant au modèle de LkHa 101, de taille 12,1 × 12,1 mas, avec le Navy Prototype Optical Interferometer (NPOI). La majorité des mesures sont à faible rapport signal à bruit[Lawson *et al.* (2004)]. . . . . 118
5.4 Objet vrai et couverture $u-v$ pour le fichier de données 1. Ces données simulent l'observation d'un objet correspondant à un modèle d'étoile double, de taille 30 × 30 mas, avec le Navy Prototype Optical Interferometer (NPOI). La majorité des mesures sont à fort rapport signal à bruit[Lawson *et al.* (2004)]. . . . . 118
5.5 Objet original et reconstructions soumises à l'Imaging Beauty Contest, obtenues avec WISARD1, régularisation : L2 et positivité. . . . . 119
5.6 Amélioration des performances de WISARD. Les reconstructions (b) et (c) s'appuient sur une modélisation à 3 télescopes, les reconstructions (d), (e) et (f) sur une modélisation à 6 télescopes (voir section 4.5) . . . . . 120
5.7 Illustration des performances de WISARD. Les images (c), (d) et (e) simulent les images que formeraient des télescopes monolithiques limités par la diffraction, de diamètres 10m, 66m et 100m. . . . . 122
5.8 Vue générale de l'interféromètre IOTA . . . . . 123
5.9 Vue schématique de IOTA . . . . . 124
5.10 Schéma de principe de l'interféromètre IOTA, en configuration « North delayed » 125
5.11 Schéma de principe de l'interféromètre IOTA, en configuration « South delayed » 126
5.12 Pupille virtuelle de l'interféromètre IOTA en configuration A5-B5-C0. . . . . 128
5.13 Comparaison des FTO obtenues pour les configurations pupillaires classique (à gauche) et optimisée (à droite). . . . . 129
5.14 Visibilités en module, pour un disque uniforme de diamètre 30 mas, en bande H ($\lambda = 1,65$ microns). . . . . 129
5.15 Modélisation de la structure de $\chi$ du cygne [Perrin *et al.* (2004)] . . . . . 131
5.16 Diagramme HR de l'amas globulaire M3 [Renzini (1988)] . . . . . 132
5.17 Modèle laminaire d'étoile MIRA [Perrin *et al.* (2004)] . . . . . 133
5.18 $\chi$ Cygni. Couverture $u - v$, IOTA mai 2005. . . . . 134
5.19 $\chi$ Cygni. Modules de visibilités, bande H, IOTA mai 2005. . . . . 135
5.20 Histogramme de répartition des clôtures mesurées. . . . . 135
5.21 Reconstruction paramétrique de $\chi$ Cygni. . . . . 136
5.22 Erreur sur les visibilités carrées d'un disque uniforme de diamètre 16 mas, avec un filtre en bande H de largeur 350 nm. . . . . 137

5.23 Comparaison des mesures de visibilités en module de $\chi$ du Cygne avec un modèle de disque uniforme de diamètre 16 mas, bande H. . . . . 138
5.24 Visibilités mesurées et visibilités estimées pour l'étoile $\chi$ du Cygne. . . . . 138
5.25 $\chi$ du Cygne. Champ reconstruit :40mas x 40 mas. Reconstructions originale et sur-échantillonnée. La tache brillante concentre 13% du flux. Regularisation : rappel quadratique vers un disque uniforme de diamètre 16 mas . . . . . 139

# Liste des tableaux

1.1 Interféromètres au sol de plus de 3 télescopes . . . . . 25

2.1 Influence du paramètre de régularisation sur l'AEQM . . . . . 65
2.2 Positions relatives et flux des 3 sources les plus faibles par rapport à la plus brillante. . . . . 71

3.1 Nombre de bases, d'observables de phase indépendants, degré d'information de phase et nombre minimal d'informations manquantes . . . . . 79