

Approximations convexes de critères pour la synthèse de Fourier optique

Serge MEIMON¹, Laurent MUGNIER¹, Guy LE BESNERAIS²

¹Office National d'Études et de Recherches Aérospatiales
Département d'Optique Théorique et Appliquée
BP 72, F-92322 Châtillon cedex, France

²Office National d'Études et de Recherches Aérospatiales
Département Traitement de l'Information et Modélisation
BP 72, F-92322 Châtillon cedex, France

meimon@onera.fr, mugnier@onera.fr
Guy.Le_Besnerais@onera.fr

Résumé – La synthèse d'ouverture optique, ou interférométrie optique, permet d'atteindre de très hautes résolutions en couplant des télescopes de diamètres raisonnables. Elle conduit à résoudre un problème d'estimation où le modèle de bruit correspond à un critère d'attache aux données non-convexe. L'étude proposée ici consiste à approcher le mieux possible ce modèle de bruit tout en imposant un critère convexe. Cette approximation a été validée sur simulations puis sur des données expérimentales.

Abstract – Optical aperture synthesis allows one to reach very high resolutions by coupling telescopes of reasonable diameters. It leads to solve an estimation problem where the noise model yields a non-convex data likelihood criterion. In this study, we show how to approximate optimally this noise model while keeping the criterion convex. This approximation has been validated by simulations, and then on experimental data.

1 Introduction

La résolution d'un télescope est inversement proportionnelle à son diamètre; elle est limitée par la technologie actuelle à une dizaine de mètres pour des télescopes au sol, et par des contraintes d'encombrement et de poids pour des télescopes embarqués sur satellite. L'interférométrie Optique (IO) est une technique permettant de dépasser cette limite; elle consiste à faire interférer un réseau de télescopes dits élémentaires et à enregistrer des données dont la résolution peut atteindre celle d'un « grand » télescope contenant tous ces télescopes élémentaires.

Les applications de cette technique sont multiples. Dans cette communication, on s'intéresse à l'observation de l'espace depuis le sol en vue de traiter des données du Very Large Telescope Interferometer (VLTI) de l'European Southern Observatory (ESO).

2 Synthèse de Fourier

Les figures d'interférences formées à partir de chaque couple (T_i, T_j) de télescopes permettent de calculer un échantillon de la Transformée de Fourier (TF) de l'objet d'intérêt, à la fréquence $\nu_{ij} = T_i T_j / \lambda^a$. Le problème de Synthèse de Fourier consiste à reconstruire l'objet d'intérêt à partir de ces mesures de TF. Dans le domaine des ondes radios, on dispose de capteurs de champ, et on mesure directement la valeur complexe de la TF de l'objet. Dans le domaine optique, on mesure les

carrés des modules des échantillons de la TF, ainsi qu'un jeu de clôtures de phase [1, 2]. Si la couverture du plan (u, v) (plan des fréquences spatiales de l'objet) est suffisamment redondante, on peut déduire des clôtures de phase les arguments des échantillons de la TF de l'objet, à une translation de l'objet près, à toutes les fréquences spatiales mesurées. On se placera dans ce cas dans le reste de cette communication. La technique utilisée ici pour résoudre ce problème inverse consiste à minimiser un critère d'adéquation aux données sous contrainte de positivité.

3 Modèle de données

Cette section aborde le problème du choix du terme d'attache aux données. Les techniques de reconstruction en interférométrie ont d'abord été développées dans le domaine des ondes radio, en supposant les hypothèses suivantes [3] :

- chaque mesure complexe est entachée d'un bruit complexe additif circulaire,
- les bruits sur deux mesures différentes sont decorrélés.

Ces hypothèses conduisent au modèle $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$ avec \mathbf{H} l'opérateur linéaire complexe transformant un vecteur \mathbf{x} d'intensités lumineuses en un vecteur complexe \mathbf{y} contenant la TF de l'objet aux fréquences mesurées. On note N_D le nombre de données, ou de mesures. La distribution choisie pour le bruit étant le plus souvent normale centrée, l'anti-log-vraisemblance obtenue est de la forme :

$$J_r(\mathbf{x}) = \sum_{i=0}^{N_D-1} \frac{|\{\mathbf{H}\mathbf{x}\}_i - y_i|^2}{\sigma_i^2} \quad (1)$$

^aAprès projection perpendiculairement à la direction d'observation.

Une étude du bruit sur l'interféromètre fibré de l'IRCOM [4] propose le modèle suivant, pour les $i^{\text{èmes}}$ mesures en phase et module :

$$\begin{cases} y_{m,i} &= |\{\mathbf{H}\mathbf{x}\}_i| + b_{m,i} \\ y_{\varphi,i} &= \angle\{\mathbf{H}\mathbf{x}\}_i + b_{\varphi,i} \end{cases} \quad (2)$$

où les deux distributions du bruit sont gaussiennes centrées et indépendantes pour phase et module de chaque mesure, avec les propriétés suivantes :

$$\begin{cases} \sigma_{m,i} &= s_m |y_i| \\ \sigma_{\varphi,i} &= s_{\varphi} \sqrt{N_{b,i}} \end{cases} \quad (3)$$

avec $N_{b,i}$ la distance entre les deux télescopes lors de la $i^{\text{ème}}$ mesure y_i , rapportée à la distance minimale entre télescopes. Des travaux sont en cours au Jean-Marie Mariotti Center (JMMC) [5] pour obtenir un modèle analogue sur le VLTI. Classiquement, les paramètres de réglage s_m et s_{φ} sont réglés à quelques pour cent.

Ces lois mènent à une anti-log-vraisemblance de la forme suivante^b :

$$J_1(\mathbf{x}) = \sum_{i=0}^{N_D-1} \left(\frac{|\{\mathbf{H}\mathbf{x}\}_i| - |y_i|}{\sigma_{m,i}} \right)^2 + \sum_{i=0}^{N_D-1} \left(\frac{\angle\{\mathbf{H}\mathbf{x}\}_i - \angle y_i}{\sigma_{\varphi,i}} \right)^2 \quad (4)$$

Nous allons montrer maintenant que le critère J_1 est non-convexe. \mathbf{H} étant linéaire, il suffit de montrer que \tilde{J}_1 , défini par

$$\tilde{J}_1(z \in \mathbb{C}) = \frac{(|z| - |z_0|)^2}{\sigma_m^2} + \frac{(\angle z - \angle z_0)^2}{\sigma_{\varphi}^2} \quad (5)$$

est non convexe. Les complexes $z_1 = z_0 \exp[\frac{2i\pi}{3}]$ et $z_2 = z_0 \exp[\frac{-2i\pi}{3}]$ vérifient $\tilde{J}_1(\frac{z_1+z_2}{2}) > \frac{1}{2}(\tilde{J}_1(z_1) + \tilde{J}_1(z_2))$, ce qui permet d'infirmer la convexité de chacun des termes en phase et module de \tilde{J}_1 .

4 Un critère convexe

Une manière naturelle d'obtenir un critère approché convexe est de recomposer à partir des mesures en phase et module des pseudo-mesures complexes $y_i = y_{m,i} e^{i y_{\varphi,i}}$. On forme alors un bruit complexe \mathbf{B} pseudo-additif

$$\begin{aligned} y_i &= \{\mathbf{H}\mathbf{x}\}_i + B_i \\ B_i &= e^{i\angle\{\mathbf{H}\mathbf{x}\}_i} \times \\ &(|\{\mathbf{H}\mathbf{x}\}_i| + b_{m,i}) e^{i b_{\varphi,i}} - \{\mathbf{H}\mathbf{x}\}_i \end{aligned} \quad (6)$$

dont on va chercher une distribution approchée gaussienne. En première approximation, la variance de y_i est de l'ordre de $\sigma_{m,i}^2$ dans une direction et de $|y_i|^2 \sigma_{\varphi,i}^2$ dans l'autre. On peut donc envisager d'utiliser le critère J_r (Cf. eq 1) hérité des techniques radio^c en fixant par exemple : $\sigma_i = \sqrt{|y_i| \cdot \sigma_{\varphi,i} \cdot \sigma_{m,i}}$

^bCertains auteurs [4] préfèrent remplacer le terme en phases par un terme sur les phaseurs

$\angle\{\mathbf{H}\mathbf{x}\}_i - \angle y_i \longleftrightarrow \frac{\{\mathbf{H}\mathbf{x}\}_i}{|\{\mathbf{H}\mathbf{x}\}_i|} - \frac{y_i}{|y_i|}$ pour éviter des problèmes d'enroulement de 2π .

^cC'est la technique utilisée par [6].

On a représenté sur la figure 1 les lignes iso-critères pour le critère J_1 dans le plan complexe. La forme de ces courbes confirme la non-convexité, et conduit à chercher une approximation de J_1 dont les courbes iso-critère seraient des ellipses d'axes principaux bien choisis, plutôt que les cercles associés au critère J_r . Il est utile de définir le changement de base suivant : pour chaque mesure y_i , on définit les vecteurs directeurs $\vec{u}_{r,i}$ et $\vec{u}_{t,i}$ obtenus par rotation de la base canonique $(O, 1, i)$ d'un angle $\angle\{\mathbf{H}\mathbf{x}\}_i$. Dans cette base, les $2 \times N_D$ coordonnées de \mathbf{B} sont les $\mathcal{B} =$

$$\{B_{r,1}, B_{t,1}, \dots, B_{r,i}, B_{t,i}, \dots, B_{r,N_D}, B_{t,N_D}\}.$$

Deux techniques sont possibles :

- approcher J_1 par son développement limité en les $B_{r,i}, B_{t,i}$ à l'ordre deux autour du minimum ;
- estimer les moments de \mathcal{B} jusqu'à l'ordre deux et supposer nuls les moments d'ordres supérieurs.

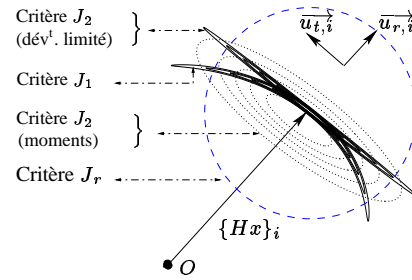


FIG. 1 – Lignes iso-critères pour une valeur de $\{\mathbf{H}\mathbf{x}\}_i$, dans le plan complexe.

Les deux techniques conduisent à des distributions de bruit gaussiennes, et par conséquent à des critères de la forme suivante :

$$J_2(\mathbf{x}) = (\mathbf{y} - \mathbf{H}\mathbf{x} - \mathbf{m}_B)^t C_B^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x} - \mathbf{m}_B) \quad (7)$$

On montre aisément que les coordonnées de \mathcal{B} sont statistiquement indépendantes. C'est l'intérêt du changement de base. Les matrices de covariance sont donc diagonales dans cette base, et le bloc diagonal 2×2 correspondant à la $i^{\text{ème}}$ mesure s'écrit :

- Par développement limité ($J_2(dl)$) :

$$C_{B,i} = \begin{pmatrix} \sigma_{m,i}^2 & 0 \\ 0 & |\{\mathbf{H}\mathbf{x}\}_i|^2 \sigma_{\varphi,i}^2 \end{pmatrix} \quad (8)$$

- Par estimation des moments ($J_2(em)$) :

$$C_{B,i} = \begin{pmatrix} \sigma_{r,i}^2 & 0 \\ 0 & \sigma_{t,i}^2 \end{pmatrix} \quad (9)$$

avec

$$\begin{cases} \sigma_{r,i}^2 &= \frac{|\{\mathbf{H}\mathbf{x}\}_i|^2}{2} (1 - \alpha_i)^2 + \sigma_{m,i}^2 \left(\frac{1 + \alpha_i^2}{2} \right) \\ \sigma_{t,i}^2 &= \frac{|\{\mathbf{H}\mathbf{x}\}_i|^2 + \sigma_{m,i}^2}{2} (1 - \alpha_i^2) \\ \alpha_i &= e^{-\sigma_{\varphi,i}^2} \end{cases}$$

La figure 1 représente quelques lignes iso-vraisemblance ou iso-critère pour les critères J_1 , J_r , $J_2(dl)$ et $J_2(em)$, dans le plan complexe, pour une valeur de $\{\mathbf{H}\mathbf{x}\}_i$. Dans le cas de l'estimation des moments, il faut tenir compte du fait que l'espérance des variables $B_{r,i}$ n'est pas nulle : $E\{B_{r,i}\} = \{\mathbf{H}\mathbf{x}\}_{r,i} (\alpha_i - 1)$. C'est à cause de ce biais \mathbf{m}_B que les centres des ellipses obtenues sur la figure 1 pour chacun des critères sont décalés selon $\vec{u}_{r,i}$. Si l'on prend $\sigma_{\varphi,i} \ll 1$ et $\sigma_{m,i} \ll \{\mathbf{H}\mathbf{x}\}_{r,i}$, ce biais tend vers 0 et la matrice de l'équation 9 tend vers celle de l'équation 8.

5 Simulations

On va maintenant comparer ces différents critères par rapport à la reconstruction d'une image astronomique. L'objet choisi est un modèle développé par P. Garcia et coll. [7] du micro-jet émis par l'étoile RU Lupus^d (cf. Fig. 2(a)).

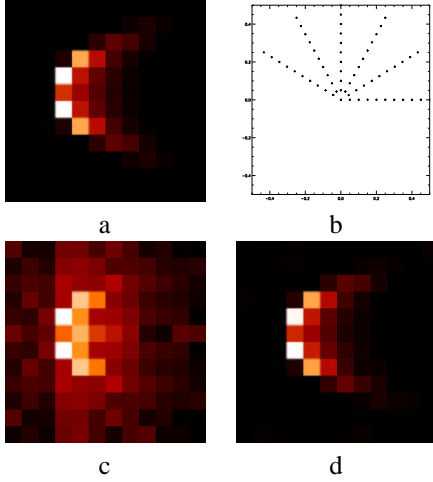


FIG. 2 – (a) Objet à reconstruire, (b) couverture (u,v), (c) « Dirty Map », (d) objet reconstruit avec $J_2(dm \text{ ou } em)$

Les échantillons de TF (les $\{\mathbf{H}\mathbf{x}\}_i$) sont générés en supposant une couverture du plan (u, v) de 61 fréquences correspondant au dispositif expérimental décrit dans [4](cf. 2(b)). Ces échantillons sont ensuite bruités avec une distribution de bruit du type de l'équation 3, avec $s_m = s_\phi = s = 3\%$, pour obtenir les y_i . On a représenté sur la figure 2(c) la « dirty map », calculée par $\tilde{\mathbf{X}} = \mathbf{H}^t \mathbf{y}$, et qui correspond à un objet dont les TF aux fréquences mesurées sont égales aux y_i , et dont la TF aux autres fréquences est nulle.

Sous-détermination On effectue 100 tirages de bruit, et on reconstruit l'objet en prenant successivement $J_2(em)$, $J_2(dl)$, J_r comme terme d'adéquation aux données. Le nombre de pixels (12×12) étant proche du nombre de mesures (61 mesures dans $\mathbb{C} = 122$ mesures dans \mathbb{R}), on reconstruit en régularisant uniquement par une contrainte de positivité. Les images obtenues sont très proches à l'œil, on n'a donc représenté qu'une reconstruction (figure 2(d)). L'utilisation du critère J_2 permet donc une amélioration de SNR^e de 0,5 dB en moyenne par rapport au critère J_r , mais les fluctuations entre deux tirages de bruits sont importantes (jusqu'à 1 dB).

Sur-détermination On se place dans un cas sur-déterminé, c'est-à-dire où le nombre d'informations est supérieur au nombre de pixels à reconstruire. Ceci est fait en imposant un support très restrictif à la reconstruction, pris égal au support réel de l'objet à reconstruire (*i.e.* 48 pixels). Les images obtenues quelles que soient les méthodes sont forcément excellentes, étant donnée la qualité des informations disponibles.

^détoile observée à une distance de 140 parsecs, à la longueur d'onde $\lambda = 1,282\mu\text{m}$, avec une taille de pixels de $1,7 \cdot 10^{-3} \text{ arcsec}$.

^e« Improvement of the Signal to Noise Ratio » de $\hat{\mathbf{x}}_2$ par rapport à $\hat{\mathbf{x}}_1 = \text{ISNR}(\hat{\mathbf{x}}_2/\hat{\mathbf{x}}_1) = 10 \log_{10} \left(\frac{\|\hat{\mathbf{x}}_1 - \mathbf{x}\|^2}{\|\hat{\mathbf{x}}_2 - \mathbf{x}\|^2} \right)$

Cependant les qualités *relatives* des reconstructions obtenues avec les différents critères sont très différentes. L'amélioration de SNR résultant de l'utilisation de J_2 plutôt que J_r est de 5 ± 1 dB. La figure 3 représente l'écart en module entre l'objet reconstruit et l'objet « vrai », pour J_r à gauche, et J_2 à droite. Il semble logique que le choix d'un bon critère d'attache aux données devienne plus sensible à mesure que le rapport données/inconnues augmente.

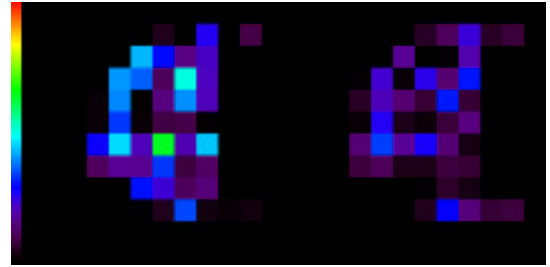


FIG. 3 – Erreur de reconstruction en module : avec J_r (gauche), avec $J_2(dm \text{ ou } em)$ (droite)

Données expérimentales Ces nouvelles techniques ont été validées sur des données expérimentales provenant de [4]. Un programme destiné à tester ces concepts sur des données provenant du VLTI est en cours d'élaboration.

6 Conclusion et perspectives

Les critères $J_2(em)$ et $J_2(dl)$ que nous avons développés sont convexes; ils plus fidèles au modèle de bruit; enfin, ils conduisent à des objets reconstruits meilleurs que le critère J_r dans un cas sous-déterminé, l'écart augmentant de façon notable lorsqu'on sur-détermine le problème. Leur utilisation dans le cadre de la reconstruction d'objets à quelques paramètres pourrait s'avérer extrêmement fructueuse.

L'étape suivante consiste à prendre en compte l'indétermination sur les phases mesurées résultant de la perturbation atmosphérique dans le cas où la couverture (u,v) est non redondante. Les méthodes actuelles, dites d'« auto-calibration », utilisent une minimisation alternée [8, 6] dont l'étape de reconstruction de l'objet à phase connue pourrait bénéficier des critères que nous proposons.

Remerciements

Les auteurs tiennent à remercier Eric Thiébaud pour son aide substantielle à la rédaction de ce manuscrit.

Références

- [1] Monnier (J. D.). – An introduction to closure phases. In : *Principles of Long Baseline Stellar Interferometry*, éd. par Lawson (P. R.), chap. 13, pp. 203–239. – Jet Propulsion Laboratory, 1999.
- [2] Le Besnerais (G.). – Problèmes de calibration en radio-astronomie par voie interférométrique. *Traitement du Signal*, vol. 10, n5-6, 1993, pp. 371–375.

- [3] Wohlleben (R.), Mattes (H.) et Krichbaum (Th.) (édité par). – *Interferometry in Radioastronomy and Radar Techniques*. – Kluwer Academic Publishers, 1991.
- [4] Delage (L.), Reynaud (F.) et Thiébaud (E.). – Imaging laboratory test on a fiber linked telescope array. *Opt. Commun.*, vol. 160, février 1999, pp. 27–32.
- [5] Jean-Marie Mariotti Center (Centre d'expertise). – <http://www-laog.obs.ujf-grenoble.fr/~jmhc/>.
- [6] Lannes (A.). – Weak-phase imaging in optical interferometry. *J. Opt. Soc. Am. A*, vol. 15, n4, avril 1998, pp. 811–824.
- [7] Garcia (P. J. V.), Cabrit (S.), Ferreira (J.) et Binette (L.). – Atomic T Tauri disk winds heated by ambipolar diffusion. II. observational tests. *Astron. Astrophys.*, vol. 377, octobre 2001, pp. 609–616.
- [8] Le Besnerais (G.). – *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.