

Étude des performances asymptotiques en interférométrie des tavelures

Yoann CAVALLIN, André FERRARI

Laboratoire d'astrophysique LUAN
CNRS/UNSA, Parc Valrose, 06108 Nice cedex 2, France
cavallin@unice.fr, ferrari@unice.fr

Résumé – Les avantages liés à l'utilisation d'interféromètres en astrophysique sont conditionnés par les performances d'estimation du contraste des franges. Cet article s'attache au calcul de la borne de Cramér-Rao sur la variance du contraste. Ce calcul est possible si on suppose que l'image suit une loi Gaussienne. Par ailleurs il requiert les moments du 1er et 2nd ordre de l'image, pour lesquels une expression analytique est obtenue. Le problème est ensuite posé pour une image observée avec des détecteurs à comptage de photons.

Abstract – Advantages related to the use of interferometers in astrophysics are conditioned by the performances of estimation of the fringes contrast. This paper is devoted to the computation of the Cramér-Rao lower bound on the variance of the contrast. This computation is tractable assuming Gaussian image distribution. Moreover it requires the 1st and 2nd order moment of the image, for which analytical expressions are obtained. The problem is then stated for image observed with photon-counting detectors.

1 Introduction

L'interférométrie stellaire consiste à faire interférer des rayons lumineux issus d'une même source, Figure 1, afin d'obtenir des informations sur cette source. Dans le cadre de l'imagerie à haute résolution angulaire, le système optique recueillant l'onde lumineuse se compose de plusieurs télescopes indépendants, 2 dans le cas de l'interféromètre du GI2T (Observatoire de la Côte d'Azur).

En l'absence de turbulence atmosphérique, le signal obtenu dans le plan image avec un interféromètre se présente sous la forme d'une tâche d'Airy frangée. Le contraste des franges d'interférences donne accès au module au carré de la transformée de Fourier de l'objet, à la fréquence spatiale associée à la longueur de base de l'interféromètre. La mesure du contraste permettra par exemple de déterminer le diamètre angulaire de la source observée.

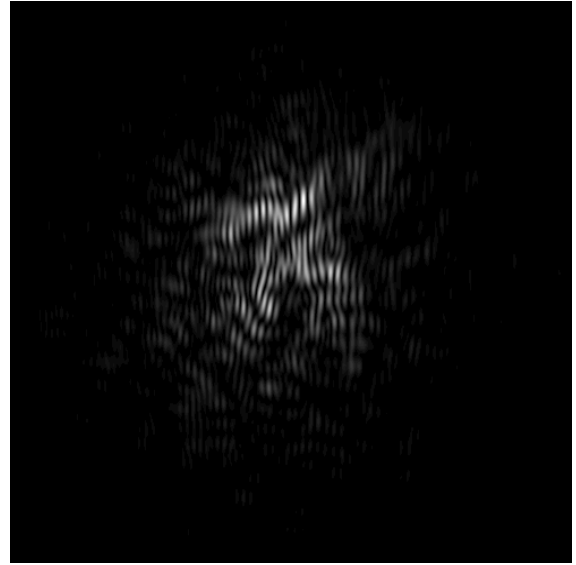


Figure 2: Image interférométrique tavelée.

Toutefois le front d'onde rayonné par la source est perturbé par la turbulence atmosphérique, et présente ainsi une phase aléatoire lorsqu'il arrive sur le télescope. L'image interférométrique, alors considérée comme un processus aléatoire non stationnaire (spatialement), se compose de petites tâches appelées tavelures, Figure 2. Ces tavelures possèdent une durée de vie nettement inférieure au temps de pose en imagerie classique : elles sont donc moyennées, engendrant une perte d'information à haute résolution angulaire. La mesure du contraste est alors impossible.

Grâce à l'interférométrie des tavelures, [1], technique qui repose sur le traitement de la densité spectrale moyenne d'images à court temps de pose, il sera possible de préserver l'informa-

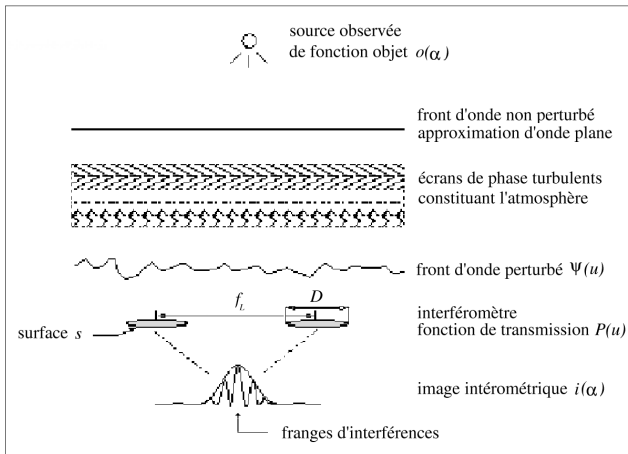


Figure 1: Obtention des franges d'interférences.

tion HF. Le module du contraste des franges peut ainsi être estimé, à partir d'une estimation paramétrique liée à ce spectre. Des méthodes d'estimation existent, mais les bornes n'ont pas encore été déterminées. Nous nous proposons dans cet article de calculer la borne de Cramér-Rao (BCR) sur la variance de l'estimé.

Après une présentation du formalisme utilisé en interférométrie des tavelures (Section 2), nous développerons les expressions analytiques des moments du 1er et 2nd ordre de l'image (Section 3), dont la connaissance est nécessaire pour le calcul de la BCR sous l'hypothèse Gaussienne. Ensuite la BCR sera déterminée lorsque l'image est acquise avec une caméra CCD classique (Section 4). Enfin nous poserons le problème du calcul de la BCR lorsque l'image est acquise avec des détecteurs à comptage de photons (Section 5), celle-ci étant alors distribuée suivant une loi de Gauss-Poisson, [9].

2 Estimation du contraste

Nous considérerons par la suite, pour simplifier les calculs, une image à une seule dimension. L'image formée au foyer d'un interféromètre est la convolution entre la fonction objet $o(\alpha)$, c'est-à-dire le flux émergent de la source observée, et la réponse impulsionnelle $s(\alpha)$ de l'ensemble atmosphère + télescope, [2]:

$$i(\alpha) = o(\alpha) * s(\alpha), \quad s(\alpha) = \frac{1}{2s} \left| \int \psi(u) P(u) e^{-2j\pi\alpha u} du \right|^2 \quad (1)$$

voir Figure 1 pour les notations. u est une variable réduite sans dimension, exprimée en unité de longueur d'onde λ du rayonnement, il en est de même pour la surface s .

L'amplitude complexe $\psi(u)$ de l'onde est modélisée par une distribution complexe Gaussienne, [2]. En se plaçant dans l'hypothèse de champ proche, seules les fluctuations de phase de $\psi(u)$ sont prises en compte, décrites par le moment du second ordre $B_\psi(f) = \mathbb{E}[\psi(u)\psi^*(u+f)]$.

En supposant que $O(f)$, transformée de Fourier de $O(\alpha)$, a une variation négligeable à travers les supports des pics franges et speckle, la densité spectrale moyenne (Figure 3) pour une source résolue admet comme modèle, [2]:

$$\mathbb{E}[|I(f)|^2] = \theta_1^2 B^2(f) + \theta_1^2 \frac{\sigma_c}{2s} T_0(f) + \theta_2^2 \frac{\sigma_c}{4s} [T_0(f+f_L) + T_0(f-f_L)] \quad (2)$$

Le vecteur $\theta = [\theta_1, \theta_2]^t = [|O(0)|, |O(f_L)|]^t$ constitue le vecteur paramètre à estimer. Pour une source ponctuelle, $\theta = [1, 1]^t$. La turbulence atmosphérique est représentée par la constante σ_c et la fonction de transfert optique à long temps de pose $B(f) = B_\psi(f)T_0(f)$, où $T_0(f) = \frac{1}{s} \int P(u)P^*(u+f)du$ est l'autocorrélation normalisée de la pupille, de telle sorte que $s = \int T_0(f)df$. Le contraste des franges d'interférences, soit le carré de la visibilité, est défini par $C_\theta = \theta_2^2/\theta_1^2$. Il est actuellement estimé par un rapport entre l'énergie des contributions de franges et l'énergie de la contribution speckle:

$$\hat{C}_\theta = \frac{\int_f \hat{\theta}_2^2 \frac{\sigma_c}{4s} [T_0(f+f_L) + T_0(f-f_L)] df}{\int_f \hat{\theta}_1^2 \frac{\sigma_c}{2s} T_0(f) df} = \frac{\hat{\theta}_2^2}{\hat{\theta}_1^2} \quad (3)$$

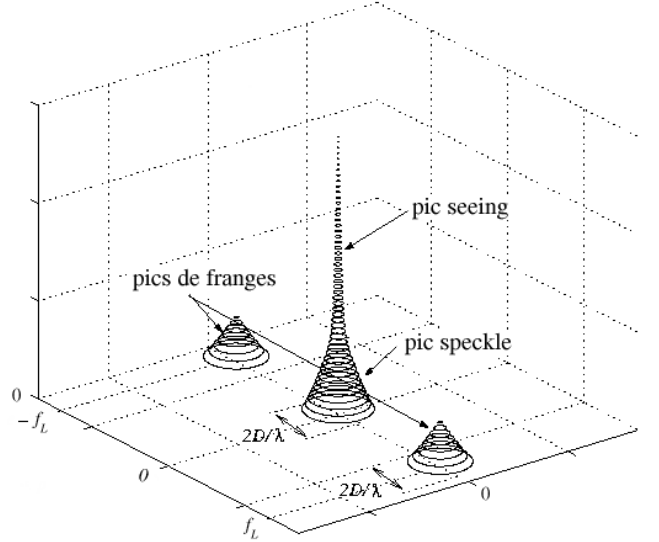


Figure 3: Modèle de densité spectrale moyenne.

3 Propriétés statistiques de l'image

Considérons une image interférométrique i_N définie sur un intervalle $[0, \ell]$ et composée de N échantillons (N pixels), telle que $\{i_N = i(nT_e), n = 0, \dots, N-1\}$, où $T_e = \ell/N$ est la période d'échantillonnage. Les expressions analytiques du vecteur moyenne \mathbf{m}_N et de la matrice de covariance \mathbf{R}_N de l'image sont déterminées dans [7]:

$$[\mathbf{m}_N]_n = \theta_1 b(nT_e); \quad [\mathbf{R}_N]_{k,l} = r((k-l)T_e) \quad (4)$$

$$r(\tau) = \frac{\sigma_c}{2s} t_0(\tau) [\theta_1^2 + \theta_2^2 \cos(2\pi f_L \tau)] \quad (5)$$

où $t_0(\tau)$ et $b(nT_e)$ sont les transformées de Fourier inverses de $T_0(f)$ et $B(f)$. Elles sont calculées à partir de l'expression (1) de $s(\alpha)$, qui s'écrit également:

$$s(\alpha) = \frac{1}{2s} \iint \psi(u)\psi^*(u+f)P(u)P^*(u+f)e^{2j\pi\alpha f} du df \quad (6)$$

En notant que $B_\psi(f) \approx 0$ lorsque $f \approx f_L$, [2], le calcul de la moyenne est immédiat. Celui de la covariance, plus conséquent, fait intervenir 2 produits:

- le moment du 4ème ordre des fluctuations de phase, qui s'écrit, pour un processus Gaussien, [3]:

$$M_\psi(f, \xi, \nu) = B_\psi(f)B_\psi(f+\nu) + B_\psi(\xi)B_\psi(\xi+\nu) \quad (7)$$

- la somme de 16 termes représentant les aires d'intersection de 4 pupilles composées chacune de 2 télescopes:

$$\int P(u)P^*(u+f)P(u+\xi)P^*(u+f+\xi+\nu) du \quad (8)$$

D'après l'approximation précédente sur $B_\psi(f)$, 24 termes sur les 32 constituant la covariance ne présentent pas d'intersection entre les télescopes. Ainsi seuls 8 termes sont non nuls. Outre les définitions de $T_0(f)$ et $B(f)$, leur développement repose essentiellement sur 2 approximations: $B_\psi^2(f) \approx \sigma_c \delta(f)$ lorsque le diamètre du télescope est grand, [8]; de la même manière il est considéré que $B_\psi(f)B_\psi(f+\rho) \approx \sigma_c \delta(f)\delta(\rho)$. Par ailleurs

ce calcul est proche des travaux de P. Berio, [8], sur la covariance de la transformée de Fourier de l'image.

Il est important de souligner que la matrice de covariance \mathbf{R}_N est une matrice de Toeplitz de dimension $N \times N$.

4 Borne de Cramér-Rao sur la variance du contraste

Considérons une image distribuée selon une loi Gaussienne: $i_N \sim \mathcal{N}(\mathbf{m}_N, \mathbf{R}_N)$. Outre le fait que cette hypothèse se justifie lors d'une acquisition en long temps de pose, elle reste valable pour les poses courtes, d'après la théorie du maximum d'entropie, lorsque seuls les moments du 1er et 2nd ordre sont connus. Or pour un processus Gaussien, la matrice 2×2 d'information de Fisher (MIF) correspondant à la BCR(θ) est, [4]:

$$[\mathbf{I}(\theta)]_{ij} = \frac{1}{2} \text{tr} \left[\mathbf{R}_N^{-1} \frac{\partial \mathbf{R}_N}{\partial \theta_i} \mathbf{R}_N^{-1} \frac{\partial \mathbf{R}_N}{\partial \theta_j} \right] \quad (9)$$

$$+ \left[\frac{\partial \mathbf{m}_N}{\partial \theta_i} \right]^t \mathbf{R}_N^{-1} \left[\frac{\partial \mathbf{m}_N}{\partial \theta_j} \right] \quad (10)$$

expression qui est divisée en 2 parties.

(9) est la MIF $\tilde{\mathbf{I}}(\theta)$ pour $\tilde{i}_N \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_N)$, un processus Gaussien, ici stationnaire d'après (4). $\tilde{\mathbf{I}}(\theta)$ est alors calculée en utilisant l'approximation asymptotique de la BCR, [4], ce qui revient à approcher la vraisemblance de \tilde{i}_N par la vraisemblance de Whittle, [5]:

$$[\tilde{\mathbf{I}}(\theta)]_{ij} \approx \frac{\ell}{2} \int_{-\frac{F_e}{2}}^{\frac{F_e}{2}} \frac{\partial \ln S(f, \theta)}{\partial \theta_i} \frac{\partial \ln S(f, \theta)}{\partial \theta_j} df \quad (11)$$

Cette expression tient compte du fait que $T_e \neq 1$. $S(f, \theta)$ est la densité spectrale de puissance de \tilde{i}_N , telle que:

$$S_N(f, \theta) = T_e \sum_{n=-N+1}^{N-1} r(nT_e) \exp\left(-2j\pi f \frac{n}{N}\right) \quad (12)$$

$$S(f, \theta) = \lim_{N \rightarrow +\infty} S_N(f, \theta) \quad (13)$$

Pour N suffisamment grand, le théorème de Shanon appliqué à l'image discrète est vérifié. Ainsi $S(f, \theta)$ caractérise la transformée de Fourier de $r(\tau)$:

$$S(f, \theta) = \frac{\sigma_c}{2s} \left[\theta_1^2 T_o(f) + \frac{\theta_2^2}{2} \{T_o(f+f_L) + T_o(f-f_L)\} \right] \quad (14)$$

L'expression (11) présente l'avantage de n'utiliser que le spectre de l'image pour calculer $\tilde{\mathbf{I}}(\theta)$. D'après les domaines de définition de $T_0(f)$ et $T_0(f \pm f_L)$, voir Figure 3, le calcul de $[\tilde{\mathbf{I}}(\theta)]_{11}$ donne:

$$[\tilde{\mathbf{I}}(\theta)]_{11} \approx \frac{\ell}{2} \int_{-D/\lambda}^{D/\lambda} \left[\frac{\partial \ln S(f, \theta)}{\partial \theta_1} \right]^2 df = \frac{4\ell D}{\lambda \theta_1^2} \quad (15)$$

Celui de $[\tilde{\mathbf{I}}(\theta)]_{22}$ est similaire, et les termes croisés sont nuls, ainsi

$$\tilde{\mathbf{I}}(\theta) = \begin{pmatrix} \frac{4\ell D}{\lambda \theta_1^2} & 0 \\ 0 & \frac{4\ell D}{\lambda \theta_2^2} \end{pmatrix} \quad (16)$$

L'expression (10) représente la contribution du vecteur moyenne $\mathbf{m}_N = \theta_1 \mathbf{b}_N$. Puisque \mathbf{m}_N ne dépend que de θ_1 , (10)

contribue uniquement à $[\mathbf{I}(\theta)]_{11}$ et vaut:

$$\left[\frac{\partial \mathbf{m}_N}{\partial \theta_1} \right]^t \mathbf{R}_N^{-1} \left[\frac{\partial \mathbf{m}_N}{\partial \theta_1} \right] = \mathbf{b}_N^t \mathbf{R}_N^{-1} \mathbf{b}_N \quad (17)$$

Il est donc nécessaire de disposer d'une expression analytique de la matrice \mathbf{R}_N^{-1} pour calculer $\mathbf{b}_N^t \mathbf{R}_N^{-1} \mathbf{b}_N$. Pourtant très souvent cette expression n'est pas disponible. Une alternative est d'approcher \mathbf{R}_N^{-1} en utilisant l'équivalence entre les matrices de Toeplitz et les matrices circulantes, [6]. En effet la matrice de Toeplitz \mathbf{R}_N est asymptotiquement équivalente à une matrice circulante \mathbf{C}_N , i.e:

$$\lim_{N \rightarrow +\infty} |\mathbf{R}_N| = \lim_{N \rightarrow +\infty} |\mathbf{C}_N| \quad (18)$$

De plus, si \mathbf{R}_N est une séquence de matrices de Toeplitz hermitiennes, alors $\mathbf{R}_N^{-1} \approx \mathbf{C}_N^{-1}$. Or l'inverse d'une matrice circulante peut s'écrire $\mathbf{C}_N^{-1} = \mathbf{U}^* \mathbf{\Psi}^{-1} \mathbf{U}$, avec

$$\psi_k = \sum_{m=-N+1}^{N-1} c_m e^{-2j\pi km/N}, \quad \mathbf{\Psi} = \{\psi_k \delta_{k,j}\}, \quad (19)$$

$$\mathbf{U} = N^{-1/2} \left\{ e^{-2j\pi km/N} \right\} \quad (20)$$

D'après l'équivalence entre \mathbf{R}_N et \mathbf{C}_N , les valeurs propres ψ_k de \mathbf{C}_N , ne sont autre que le spectre discret de \tilde{i}_N , soit:

$$\psi_k = \sum_{m=-N+1}^{N-1} r(mT_e) e^{-2j\pi km/N} = \frac{S_N(k/\ell, \theta)}{T_e} \quad (21)$$

En tenant compte de cette approximation, le développement de $\mathbf{b}_N^t \mathbf{R}_N^{-1} \mathbf{b}_N$ donne lieu au résultat suivant:

$$\mathbf{b}_N^t \mathbf{R}_N^{-1} \mathbf{b}_N \approx \frac{1}{N} \sum_{k=0}^{N-1} \frac{B^2(k/\ell)}{S_N(k/\ell, \theta)} \quad (22)$$

Si l'on suppose $B^2(k/\ell)$ et $S_N(k/\ell, \theta)$ constants sur l'intervalle $1/\ell$, on peut approcher (22) par son intégrale de Riemann associée, pour N grand:

$$\mathbf{b}_N^t \mathbf{R}_N^{-1} \mathbf{b}_N \approx \frac{\ell}{N} \int \frac{B^2(f)}{S(f, \theta)} df \approx \frac{2sT_e}{\theta_1^2} \quad (23)$$

Il en résulte que pour N grand, T_e tendant alors vers 0, la contribution de (10) devient négligeable par rapport à celle de (9). Intuitivement, il semble cohérent de négliger l'influence de la moyenne dans le calcul de la MIF, puisque la moyenne dépend seulement de θ_1^2 alors que la covariance dépend conjointement de θ_1^2 et θ_2^2 .

Finalement, la MIF de i_N se réduit à

$$\mathbf{I}(\theta) = \begin{pmatrix} \frac{4\ell D}{\lambda \theta_1^2} & 0 \\ 0 & \frac{4\ell D}{\lambda \theta_2^2} \end{pmatrix} \quad (24)$$

$\mathbf{I}(\theta)$ correspond aux BCR sur les variances de θ_1 et θ_2 , indépendamment. La BCR sur la variance du contraste, défini par $C_\theta = \theta_2^2/\theta_1^2$, se déduit alors de (24) par une transformation de paramètres, [4]:

$$BCR(C_\theta) = \left[\frac{\partial C_\theta}{\partial \theta} \right] \mathbf{I}^{-1}(\theta) \left[\frac{\partial C_\theta}{\partial \theta} \right]^t = \frac{\lambda C_\theta^2}{2\ell D} \quad (25)$$

De la même manière, le calcul de la BCR pour des fonctions à deux dimensions donne $C_\theta^2/4s\ell$.

Notons que (25) décroît avec ℓ . Cette décroissance est due au modèle simplifié utilisé pour décrire les propriétés statistiques de l'image. En pratique l'image (Figure 2) est contenue dans un disque appelé *seeing disc*, de taille λ/r_0 , où le paramètre de Fried r_0 caractérise l'état de la turbulence atmosphérique. Il est directement relié à σ_c par la relation $\sigma_c = 0.342r_0^2$. En exprimant la taille de l'image en fonction de la taille du *seeing disc*, de sorte que $\ell = \eta\lambda/r_0$, la BCR de C_θ peut se mettre sous la forme:

$$BCR(C_\theta) = \frac{C_\theta^2}{2\eta} \left[\frac{D}{r_0} \right]^{-1} \quad (26)$$

où la quantité D/r_0 est proportionnelle au nombre de tavelures présentes dans l'image, [2]. Au plus D/r_0 sera grand, au plus la borne sera faible. Cela traduit le fait qu'une image constituée d'un grand nombre de tavelures contient plus d'information, et présente alors de meilleures performances dans l'estimation de paramètres.

5 Photonisation de l'image

À l'inverse de la détection classique d'images utilisant une caméra CCD, la photonisation consiste à acquérir les images au moyen d'une caméra à comptage de photons. Le processus de photonisation peut être modélisé par la théorie semi-classique de la détection photoélectrique, [10]. Cette théorie permet notamment de déterminer la probabilité $P(K_n)$ de détecter K_n photoévénements sur le n -ième pixel de l'image.

Lorsque l'image est déterministe, K_n suit une loi de Poisson de moyenne $\bar{K}_n = \alpha i_n$, où α est le rendement quantique et i_n l'intensité intégrée du n -ième pixel de l'image. C'est aussi le cas lorsque l'image est non déterministe, toutefois \bar{K}_n est alors conditionné par la valeur de i_n , en l'occurrence aléatoire. Ainsi K_n suivra une loi de Poisson conditionnelle, également appelée formule de Mandel, [10]. Pour une image distribuée suivant une loi conjointe Gaussienne $p_i(i_1, \dots, i_N)$ avant photonisation, telle que $i \sim \mathcal{N}(\mathbf{m}_N, \mathbf{R}_N)$, les K_n suivront une loi conjointe de Gauss-Poisson, [9]. La distribution conjointe des N photoévénements constituant l'image s'écrira alors:

$$P(K_1, \dots, K_N) = \int_i \prod_{n=1}^N \frac{(\alpha i_n)^{K_n}}{K_n!} \exp(-\alpha i_n) \times p_i(i_1, \dots, i_N) di_1 \dots di_N \quad (27)$$

Malheureusement il n'existe pas d'expression analytique plus accessible que (27), ainsi le calcul de la BCR devient très vite complexe. Seule la fonction caractéristique des K_n est disponible, [9]. Toutefois, les expressions du vecteur moyenne et de la matrice de covariance de (27) sont connues, [10]:

$$\mathbf{E}[P(K_1, \dots, K_N)] = \alpha \mathbf{m}_N \quad (28)$$

$$\text{cov}[P(K_1, \dots, K_N)] = \alpha \mathbf{m}_N + \alpha^2 \mathbf{R}_N \quad (29)$$

En première approximation, nous serions tentés de calculer la BCR de l'image photonisée de la même manière que pour l'image non photonisée (Section 4), c'est-à-dire en considérant une image distribuée suivant une loi Gaussienne ayant pour moments les expressions (28) et (29). Mais un problème supplémentaire se pose ici : la photonisation introduit une non-stationnarité dans l'expression de la covariance, par le terme $\alpha \mathbf{m}_N$. Or cela ne nous permet pas d'utiliser les approximations effectuées précédemment pour le calcul de la MIF.

Pourtant le calcul de la BCR incluant la photonisation ne peut s'appuyer que sur la fonction caractéristique de (27), ou, à défaut, sur les expressions de ses moments. Toutefois, dans les deux cas, la BCR semble ne pas admettre de solutions analytiques sans l'apport d'hypothèses supplémentaires sur la distribution de l'image.

6 Conclusion

Dans le cadre de l'estimation du contraste des franges d'interférences d'une image interférométrique, nous avons calculé la borne de Cramér-Rao sur la variance du contraste. Ce calcul a nécessité l'hypothèse d'une image distribuée suivant une loi Gaussienne, dont nous avons auparavant déterminé les expressions analytiques des moments du 1er et 2nd ordre. La borne résultante est certainement optimiste car elle n'inclut pas certains effets défavorables telles que les erreurs instrumentales. Elle constitue néanmoins une référence dans l'estimation du contraste.

La prise en compte de la photonisation des images a été envisagée, mais elle donne lieu à des expressions analytiques trop complexes pour parvenir au calcul de la BCR.

References

- [1] A. Labeyrie. *Progress in Optics*, vol. XIV, pp. 49-87, E. Wolf, 1976.
- [2] F. Roddier. *Progress in Optics*, vol. XIX, pp. 282-376, E. Wolf, 1981.
- [3] B. Picinbono. *Random Signals and Systems*, Prentice Hall, 1993.
- [4] S. Kay. *Fundamentals of Statistical Signal Processing - Estimation Theory*, Prentice Hall, 1993.
- [5] P. Whittle. "Gaussian estimation in stationary time series", *Bul. Int. Ins. Stat.*, vol. 39, pp. 105-129, 1962.
- [6] U. Grenander and G. Szegő. *Toeplitz Forms and their Applications*. University of California press, 1958.
- [7] Y. Cavallin and A. Ferrari. "Fringes contrast estimation in Michelson stellar interferometry : asymptotic performances," *Astronomy & Astrophysics*, submitted on april 2003.
- [8] P. Berio et al. "Spectrally resolved Michelson stellar interferometry. I. Exact formalism in the multispeckle mode," *JOSA A*, vol. 16, pp. 872-881, 1999.
- [9] J. Grandell. *Mixed Poisson Processes*. Chapman and Hall, 1997.
- [10] J.W. Goodman. *Statistical Optics*. Wiley Interscience, 1985.