



HAL
open science

Lois de Poisson mélangées multivariées : application à la détection en présence de speckles

Florent Chatelain, André Ferrari, Jean-Yves Tourneret

► To cite this version:

Florent Chatelain, André Ferrari, Jean-Yves Tourneret. Lois de Poisson mélangées multivariées : application à la détection en présence de speckles. *Traitement du Signal*, 2009, 26 (3), pp.227–238. 10.3166/ts.26.227-238 . hal-00372288

HAL Id: hal-00372288

<https://hal.science/hal-00372288>

Submitted on 15 Feb 2024

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.

Lois de Poisson mélangées multivariées : application à la détection en présence de speckles

F. Chatelain, A. Ferrari et J.-Y. Tourneret

F. Chatelain,
Gipsa-lab, Grenoble INP
961 rue de la Houille Blanche, BP 46,
38042 Saint Martin d'Hères
email : florent.chatelain@gipsa-lab.fr

A. Ferrari,
Lab. Fizeau, UNS/OCA/CNRS
Parc Valrose,
06108 Nice cedex 2.
email : ferrari@unice.fr

J.-Y. Tourneret,
IRIT/ENSEEIH/TSa,
2 rue Charles Camichel, BP 7122,
31071 Toulouse cedex 7.
email : jean-yves.tourneret@enseeiht.fr

Résumé

L'article étudie l'estimation des paramètres de l'amplitude complexe du front d'onde à partir de mesures d'intensités faible-flux. La corrélation spatiale de l'amplitude complexe du front d'onde est modélisée au sein d'une clique de l'image observée. Le cas général où cette amplitude complexe est de moyenne non nulle est également considéré. Un estimateur classique des moments, ainsi qu'un estimateur des moindres carrés non-linéaires prenant en compte les corrélations spatiales sont proposés. Dans le cadre de la détection d'exoplanètes par imagerie directe, les performances de ces estimateurs et des détecteurs associés sont étudiées et comparées. Les résultats obtenus mettent en évidence la nécessité de disposer d'un a priori plus fort sur la corrélation spatiale du front d'onde afin d'améliorer les performances d'estimation et de détection.

Mots clés : Speckle, corrélation spatiale, lois de Poisson mélangées multivariées, vraisemblance composite, méthode de moments, moindres carrés non linéaires, détection d'exoplanètes.

1 Introduction

On s'intéresse dans cet article à l'estimation des paramètres de l'amplitude complexe d'un front d'onde à partir de la mesure de l'intensité lumineuse. Nous ferons l'hypothèse que l'amplitude complexe est gaussienne circulaire. On supposera également que les mesures

sont obtenues par comptage de photons, ce qui est le cas pour de faibles durées d'exposition ou lorsque les objets observés sont de faible intensité. Cette problématique, liée à de très nombreux systèmes optiques, intervient notamment en imagerie astronomique. L'application ici visée est la détection d'exoplanètes par imagerie directe, problème qui suscite actuellement un intérêt important [3] en raison du développement de nouveaux instruments tels que ESO/VLT SPHERE ou Gemini Planet Imager.

Considérons tout d'abord le cas où la moyenne de l'amplitude complexe est nulle, ce qui est le cas lorsque l'amplitude complexe résulte de la somme d'un très grand nombre de phaseurs aléatoires indépendants. La distribution du nombre de photons comptés sur un pixel dans le cas d'une amplitude complexe de moyenne nulle est un résultat classique d'optique statistique [13] : l'intensité est distribuée suivant une densité exponentielle (speckles pleinement développés) et le nombre de photons suit une loi géométrique (ou loi de Bose-Einstein). Cette distribution est, par exemple, à la base de la technique de détection de planètes proposée par A. Labeyrie [18] et analysée de façon plus "formelle" dans [1, 11].

Toujours dans le cas univarié, considérons maintenant le cas où l'amplitude complexe n'est plus à moyenne nulle. C'est le cas par exemple en holographie ou dans l'application développée section 4. L'intensité est alors proportionnelle à une variable suivant une loi du Chi deux décentrée à deux degrés de liberté [16] (ou distribution de Rice [7]). Le calcul de la distribution du nombre de photons détectés est cette fois beaucoup plus difficile. Une expression de cette distribution a été obtenue sous forme de série de polynômes de Laguerre dans [7] et en utilisant la fonction hypergéométrique confluyente de Kumer dans [2].

L'exemple précédent montre clairement le niveau de complexité apporté par le fait que la moyenne de l'amplitude complexe n'est pas nulle. Or, cette hypothèse de moyenne non nulle est centrale dans de nombreux problèmes comme la modélisation des speckles laser en holographie [14] ou des images obtenues par optique adaptative [7], ou encore pour modéliser les images issues d'un système de détection de planètes extrasolaire par imagerie directe [2].

L'objectif de cet article est d'étendre les résultats précédents au cas multivarié et de proposer une stratégie d'estimation et de détection qui prenne alors en compte les corrélations spatiales du front d'onde. Il est organisé de la façon suivante :

- La deuxième partie présente le modèle statistique utilisé pour l'amplitude complexe du front d'onde, les mesures d'intensité et les mesures en comptage de photons.
- La troisième partie est consacrée à l'estimation des paramètres de l'amplitude complexe du front d'onde à partir de mesures de comptage de photons. Deux cas sont distingués : si la moyenne de l'amplitude complexe est nulle, l'estimation peut être obtenue par maximisation de la vraisemblance composite. Dans le cas contraire, il faudra recourir à une méthode de moments. Soulignons que le problème de l'identifiabilité des paramètres n'est pas abordé.
- Enfin, dans la dernière partie, nous nous proposons d'appliquer les résultats précédents au problème de la détection de planètes extrasolaires par imagerie directe.

2 Modèle Statistique

2.1 Loi des intensités

On considère le vecteur de taille $M \times 1$ noté $\psi(\vec{x}) = (\psi_1(\vec{x}), \dots, \psi_M(\vec{x}))^T$ des amplitudes complexes de M pixels situés dans le voisinage noté \mathcal{V} du point à la position \vec{x} . Afin de simplifier les notations, nous ferons dans la suite abstraction de la dépendance envers la

position \vec{x} . Le modèle général utilisé dans [2] suppose que sur le pixel k , $\psi_k(\vec{x})$ se décompose en la somme d'un terme déterministe et d'un terme gaussien circulaire centré. Ce modèle s'étend naturellement au cas multivarié en supposant que l'amplitude complexe ψ se décompose en la somme de deux termes :

- le premier terme qui est *déterministe* est lié à la réponse impulsionnelle du système. Ce terme jouera, comme on le verra dans la suite un rôle essentiel dans le niveau de complexité du problème.
- le second terme représente les imperfections du front d'onde. Il est modélisé de façon classique par un vecteur gaussien complexe circulaire centré [21].

Il en résulte que le vecteur ψ suit une loi gaussienne complexe circulaire : $\psi \sim \mathcal{N}_c(\mu(\theta), \Sigma(\theta))$, de moyenne $\mu(\theta)$ et de matrice de covariance complexe $\Sigma(\theta)$. Nous supposons donc dans la suite que le de paramètres θ est inconnu et l'on s'intéressera à son estimation.

Les composantes du vecteur (de taille $M \times 1$) des intensités $I = (I_1, \dots, I_M)^T$ sont définies par $I_i = |\psi_i|^2$, pour $i = 1, \dots, M$, et sont donc proportionnelles à des variables aléatoires suivant des lois du Chi deux décentrées.

Le calcul de la densité jointe de I est, lui, beaucoup plus délicat. Avant d'aborder ce problème, il est possible de s'intéresser au calcul des moments joints de I :

$$\mathbb{E} \left[\prod_{k \in \mathcal{E}} I_k \right] = \mathbb{E} \left[\prod_{k \in \mathcal{E}} |\psi_k|^2 \right] = \mathbb{E} \left[\prod_{k \in \mathcal{E}} \psi_k \psi_k^* \right] \quad (1)$$

où \mathcal{E} est un ensemble d'indices. Cette quantité peut être calculée en utilisant le développement des moments d'ordre supérieur d'un vecteur en fonction de ses cumulants [19] :

$$\mathbb{E} \left[\prod_{k \in \mathcal{L}} \psi_k^{\$k} \right] = \sum_{\{v_1, \dots, v_p\} \in \mathcal{P}_{\mathcal{L}}} \text{cum} [\psi_l^{\$l}, l \in v_1] \cdots \text{cum} [\psi_l^{\$l}, l \in v_p], \quad (2)$$

où $\mathcal{P}_{\mathcal{L}}$ désigne l'ensemble de toutes les partitions $\{v_1, \dots, v_p\}$ de \mathcal{L} et, $\$k$ dans $z^{\$k}$, indique si z est conjugué ou non. La nature de la distribution de ψ permet une simplification importante de (2) puisque les seuls cumulants non nuls d'un vecteur gaussien circulaire sont sa moyenne et sa covariance.

Ce résultat permet par exemple un calcul des deux premiers moments de l'intensité :

$$\begin{aligned} \mathbb{E} [|\psi_k|^2] &= \text{var}[\psi_k] + |\mathbb{E}[\psi_k]|^2, \\ \text{cov}[|\psi_k|^2, |\psi_l|^2] &= |\text{cov}[\psi_k, \psi_l^*]|^2 + \mathbb{E}[\psi_k^*] \mathbb{E}[\psi_l] \text{cov}[\psi_k, \psi_l^*] + \mathbb{E}[\psi_k] \mathbb{E}[\psi_l^*] \text{cov}[\psi_l, \psi_k^*], \end{aligned}$$

qui se simplifie en :

$$\mathbb{E}[I_k] = \Sigma_{k,k}(\theta) + |\mu(\theta)|^2, \quad (3)$$

$$\text{cov}[I_k, I_l] = |\Sigma_{k,l}(\theta) + \mu_k(\theta) \mu_l^*(\theta)|^2 - |\mu_k(\theta) \mu_l(\theta)|^2. \quad (4)$$

Quand $k = l$, (4) donne la variance de l'intensité déjà obtenue dans [1] :

$$\text{var}[I_k] = \Sigma_{k,k}^2(\theta) + 2\mu_k^2(\theta) \Sigma_{k,k}(\theta), \quad (5)$$

qui montre le phénomène classique d'amplification de la variance de l'image par la moyenne du front d'onde.

Une façon simple de caractériser la loi de I est de la relier à une loi de Wishart complexe non centrée, définie dans [20] ou [23]. En effet, si l'on définit la matrice suivante :

$$U = \psi\psi^H, \quad (6)$$

où H représente l'opérateur de transposition-conjugaison, la matrice U suit, par construction, une loi de Wishart complexe non centrée à 1 degré de liberté, ψ étant un vecteur gaussien complexe circulaire non centré. On a, évidemment, $I = \text{diag}(U)$. Ici $\text{diag}(A)$ désigne indifféremment le vecteur des éléments diagonaux de la matrice A , ou la matrice diagonale ayant les composantes du vecteur A sur sa diagonale. Donc :

$$I \sim \text{diag} \left(\mathcal{W}_c^d(1, \Sigma(\theta), \Sigma(\theta)^{-1}\mu(\theta)\mu(\theta)^H) \right). \quad (7)$$

Il est important de souligner que, pour $M > 1$, la matrice U est singulière et donc sa densité de probabilité n'existe pas. D'autre part cette caractérisation ne permet pas d'aboutir à une expression explicite de la densité de probabilité de I dans le cas général ($M > 1$). La fonction génératrice des moments de I s'obtient par contre directement en mettant à zero les variables appropriées de la fonction génératrice des moments de U :

$$L(s) = \mathbb{E} \left[e^{\sum_{k=1}^M s_k I_k} \right] = \frac{e^{\text{Trace}[\text{diag}(s)(I - \Sigma(\theta)\text{diag}(s))^{-1}\mu(\theta)\mu(\theta)^H]}}{\det[I - \Sigma(\theta)\text{diag}(s)]}. \quad (8)$$

2.2 Intensité photonisée

Le vecteur des intensités défini au paragraphe précédent correspond au cas où les images sont enregistrées sous l'hypothèse de fort-flux. Dans le cadre faible-flux, le bruit de Poisson lié à la photonisation des intensités doit être modélisé. Soit N le vecteur des nombres de photons associés au vecteur des intensités I . Alors, conditionnellement au vecteur I , les variables aléatoires N_i sont indépendantes et distribuées selon des lois de Poisson de moyennes I_i , pour tout $i = 1, \dots, M$. L'expression des masses de probabilité est alors :

$$\Pr(N = k) = \int \cdots \int_{(\mathbb{R}^+)^d} \prod_{\ell} \frac{(I_{\ell})^{k_{\ell}}}{k_{\ell}!} \exp(-I_{\ell}) p(I) dI, \quad (9)$$

où $k = (k_1, \dots, k_M) \in \mathbb{N}^M$ et $p(I)$ est la densité du vecteur I défini précédemment. Le vecteur N suit une loi multivariée appelée *loi de Poisson mélangée multivariée* dans la littérature [15, 12]. Cette opération de mélange de la loi de Poisson par la loi des intensités est connue en optique statistique sous le nom de transformation de Poisson-Mandel [13]. Les propriétés de cette transformation ont été étudiées en détail dans la littérature et nous nous limiterons ici aux quelques propriétés nécessaires dans le développement de cet article.

Il est bien sûr difficile d'obtenir une expression explicite des masses $\Pr(N = k)$ à partir de (9). Par contre, les moments joints du vecteur N peuvent se calculer directement en fonction des moments joints du vecteur des intensités I selon la formule suivante [12] :

$$\mathbb{E} \left[\prod_{k \in \mathcal{E}} N_k^{r_k} \right] = \sum_{j_1=0}^{r_1} \cdots \sum_{j_M=0}^{r_M} \prod_{k \in \mathcal{E}} S(r_k, j_k) \mathbb{E} \left[\prod_{k \in \mathcal{E}} I_k^{j_k} \right], \quad (10)$$

où $S(j, k)$ sont les nombres de Stirling de deuxième espèce [17] ($S(j, k)$ est le nombre de partitions d'un ensemble de j éléments en k sous-ensembles). Par exemple, cette relation

conduit pour les deux premiers moments à $\mathbb{E}[N_k] = \mathbb{E}[I_k]$, et $\text{cov}(N_k, N_l) = \text{cov}(I_k, I_l) + \delta(k - l)\mathbb{E}[I_k]$.

Soulignons enfin que la fonction génératrice des moments joue un rôle central pour la transformation (9). Ainsi, comme cela est montré dans [12], la fonction génératrice du vecteur N , notée $G(z) = \mathbb{E} \left[\prod_{k=1}^M z_k^{N_k} \right]$, a pour expression :

$$G(z) = L(z - 1). \quad (11)$$

La fonction génératrice du vecteur N s'obtient donc en substituant simplement s_k par $z_k - 1$ dans la fonction génératrice des moments de I .

3 Estimation des paramètres et détection

3.1 Cas centré : $\mu = 0$

Dans ce cas, le terme déterministe lié à la réponse impulsionnelle du système est supposé nul. Par conséquent, la loi de l'amplitude complexe ψ du front d'onde se ramène à une loi gaussienne complexe centrée. À partir de l'équation (8), on obtient immédiatement l'expression suivante de la fonction génératrice des moments, i.e. la transformée de Laplace, du vecteur des intensités I :

$$L(s) = \mathbb{E} \left[e^{\sum_{k=1}^M s_k I_k} \right] = [\det(I - \Sigma(\theta)\text{diag}(s))]^{-1}. \quad (12)$$

Il est important de noter que chaque intensité est distribuée selon une loi exponentielle. En effet, la transformée de Laplace de la composante I_k , notée L_k , se déduit directement de la transformée de Laplace du vecteur des intensités I puisque :

$$L_k(s_k) = \mathbb{E} \left[e^{s_k I_k} \right] = L(0, \dots, 0, s_k, 0, \dots, 0),$$

pour tout $1 \leq k \leq M$. À partir de (12), en développant le déterminant selon la k ème colonne, il vient finalement que la transformée de Laplace de la variable I_k a pour expression :

$$L_k(s_k) = (1 - \Sigma_{k,k}(\theta)s_k)^{-1},$$

où $\Sigma_{k,k}(\theta)$ est la k ème composante de la diagonale de la matrice de covariance $\Sigma(\theta)$. On retrouve donc l'expression classique de la transformée de Laplace d'une loi gamma de paramètre de forme égal à 1, i.e. une loi exponentielle, et de paramètre d'échelle $\Sigma_{k,k}(\theta)$. Par ailleurs, en raison de la propriété de multilinéarité du déterminant, l'application $P(s) = \det(I - \Sigma(\theta)\text{diag}(s))$ est un polynôme affine¹ par rapport au vecteur s . Par conséquent, la loi du vecteur des intensités I appartient à la famille des *lois gamma multivariées* introduite dans [4] et étudiée, en particulier, dans [6].

La fonction génératrice du vecteur des intensités photonisées N s'obtient maintenant directement à partir de l'expression de la fonction génératrice des moments du vecteur des intensités I (12) et de la relation (11). Ceci permet de montrer que cette fonction génératrice est la puissance négative d'un polynôme affine. La famille de lois de probabilité à laquelle

¹Un polynôme $P(s)$ est dit affine par rapport au vecteur $s = (s_1, \dots, s_M)$ si, pour tout $j = 1, \dots, M$, la fonction d'une seule variable $s_j \mapsto P(s)$ peut s'exprimer sous la forme $A^{(-j)}s_j + B^{(-j)}$, où $A^{(-j)}$ et $B^{(-j)}$ ne dépendent pas de s_j et sont des polynômes par rapport aux variables s_i pour tout $i \neq j$.

appartient la loi du vecteur N a alors été recensée et étudiée dans la littérature sous le nom de *loi négative multinomiale* [5]. Soulignons que les marges d'une telle loi multivariée, et donc les lois des intensités photonisés N_k pour tout $1 \leq k \leq M$, sont alors des lois négatives binomiales, famille à laquelle appartient la loi géométrique.

Un obstacle majeur afin d'estimer par maximum de vraisemblance les paramètres de ces lois négatives multinomiales, qui modélisent la distribution des observations N , est qu'il n'est pas possible d'obtenir une expression analytique explicite des masses de probabilité en dimension M quelconque. Ceci conduit à utiliser des méthodes de maximum de vraisemblance composite afin d'estimer les paramètres. Les vraisemblances composites sont, en effet, construites à partir d'évènements marginaux ou conditionnels pour lesquels la vraisemblance admet une expression explicite [24, et les références proposées]. La log-vraisemblance composite est alors obtenue comme une combinaison linéaire des log-vraisemblances associées à ces évènements. Le choix des évènements considérés doit alors être guidé par les deux critères suivants :

1. les paramètres à estimer doivent être identifiables à partir de la vraisemblance composite,
2. un compromis entre les performances des estimateurs du maximum de vraisemblance composite ainsi obtenus, et, le coût calculatoire de la méthode doit être recherché.

Dans le cas qui nous intéresse, il s'avère ainsi possible d'obtenir une expression explicite des masses de probabilités bivariées d'une loi négative multinomiale de dimension M quelconque [9, 10]. En particulier, on trouve l'expression suivante des masses de probabilités des paires du vecteur des intensités photonisées N :

$$\Pr(N_i = n_i, N_j = n_j) = a_{ij}^{n_i} b_{ij}^{n_j} [(1 - a_{ij})(1 - b_{ij}) - c_{ij}] \sum_{k=0}^{\min(n_i, n_j)} \binom{n_i}{k} \binom{n_j}{k} \left(\frac{c_{ij}}{a_{ij} b_{ij}} \right)^k, \quad (13)$$

pour tout $1 \leq i, j \leq M$, avec

$$\begin{aligned} a_{ij} &= [\Sigma_{i,i}(\theta) + \Sigma_{i,i}(\theta)\Sigma_{j,j}(\theta) - \Sigma_{i,j}(\theta)^2] / \Delta_{ij}, \\ b_{ij} &= [\Sigma_{j,j}(\theta) + \Sigma_{i,i}(\theta)\Sigma_{j,j}(\theta) - \Sigma_{i,j}(\theta)^2] / \Delta_{ij}, \\ c_{ij} &= [\Sigma_{i,i}(\theta)\Sigma_{j,j}(\theta) - \Sigma_{i,i}(\theta)\Sigma_{j,j}(\theta) + \Sigma_{i,j}(\theta)^2] / \Delta_{ij}^2, \\ \Delta_{ij} &= 1 + \Sigma_{i,i}(\theta) + \Sigma_{j,j}(\theta) + \Sigma_{i,i}(\theta)\Sigma_{j,j}(\theta) - \Sigma_{i,j}(\theta)^2. \end{aligned}$$

Dans le cas centré, les paramètres à estimer correspondent aux paramètres θ de la matrice de covariance $\Sigma(\theta)$ du front d'onde complexe gaussien. La matrice de covariance $\Sigma(\theta)$ peut alors se déduire, moyennant quelques hypothèses non restrictives dans nos applications, de la matrice de covariance du vecteur des intensités photonisés N grâce aux relations (10) et (4). Les paramètres du modèle sont donc bien identifiables à partir de la vraisemblance composite associée aux paires du vecteur N . Cette méthode du maximum de vraisemblance composite associée aux paires a été étudiée dans [9, 10]. Les performances des estimateurs ainsi obtenus pour ces modèles s'avèrent alors sensiblement supérieures à des méthodes de moments classiques.

3.2 Cas non centré : $\mu \neq 0$

Le problème de l'estimation du vecteur θ est dans ce cas difficile puisqu'il n'existe pas d'expression analytique explicite des masses multivariées de la distribution du vecteur N

pour $M > 1$. La méthode du maximum de vraisemblance composite associée aux paires, utilisée dans le cas centré, n'est donc pas envisageable. On se propose donc d'estimer ce vecteur par des méthodes de moments.

Estimer les paramètres inconnus par une méthode des moments nécessite de connaître une expression théorique explicite de ces moments. Dans notre cas, une expression des moments joints du vecteur N est donnée par (10) en fonction des moments joints du vecteur des intensités I . Les moments joints du vecteur I peuvent être obtenus en utilisant les équations (1) et (2) où les résultats généraux sur les moments des matrices de Wishart non centrées [23].

Nous allons dans la suite de cette section rappeler un certain nombre de résultats relatifs à l'estimation par la méthode des moments. Nous supposons pour cela que nous disposons de n réalisations indépendantes du vecteur N , notées $N^{[k]}$, pour tout $k = 1, \dots, n$. Considérons une fonction $h(\cdot) : \mathbb{R}^M \rightarrow \mathbb{R}^L$ et la statistique associée : $s_n = \frac{1}{n} \sum_{i=1}^n h(N^{[i]})$. Si l'on considère, par exemple, $M = 2$ et $h(x) = (x_1, x_1^2, x_1x_2, x_2, x_2^2)$, la statistique utilisée pour l'estimation sera composée de la moyenne des deux pixels, de leur puissance et de leur corrélation.

Soient

$$\begin{aligned} f(\theta) &= \mathbb{E}[s_n] &= \mathbb{E}[h(N^{[1]})], \\ C(\theta) &= n \operatorname{cov}[s_n] &= \operatorname{cov}[h(N^{[1]})], \end{aligned} \quad (14)$$

les expressions théoriques de la moyenne et de la covariance associée à la statistique s_n en fonction du vecteur des paramètres à estimer θ . Les deux estimateurs suivants sont alors étudiés [8] :

$$\begin{aligned} \hat{\theta}_1 &= g(s_n) \quad \text{où} \quad g(f(\theta)) = \theta, \\ \hat{\theta}_2 &= \arg \min_x \frac{1}{2} (f(x) - s_n)^T C(x)^{-1} (f(x) - s_n). \end{aligned} \quad (15)$$

L'estimateur $\hat{\theta}_1$ est un estimateur classique des moments tandis que l'estimateur $\hat{\theta}_2$ est obtenu par une méthode des moindres carrés non linéaires (MCNL). Ces estimateurs sont convergents et leurs performances asymptotiques s'expriment en fonction de $C(\theta)$, et des matrices jacobiniennes $G(\theta)$ et $F(\theta)$ des fonctions g et f en $f(\theta)$ et θ respectivement [22].

$$\lim_{n \rightarrow \infty} n \mathbb{E}[(\hat{\theta}_1 - \theta)(\hat{\theta}_1 - \theta)^T] = G(\theta)C(\theta)G(\theta)^T, \quad (16)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} n \mathbb{E}[(\hat{\theta}_2 - \theta)(\hat{\theta}_2 - \theta)^T] &= (F(\theta)C(\theta)^{-1}F(\theta)^T)^{-1} \\ &= B(\theta). \end{aligned} \quad (17)$$

L'intérêt de l'estimateur MCNL, plus complexe à calculer, est justifié par la propriété asymptotique suivante [22] :

$$G(\theta)C(\theta)G(\theta)^T \geq B(\theta), \quad (18)$$

où $A_1 \geq A_2$ signifie que la différence $A_1 - A_2$ est une matrice semi-définie positive. La matrice $B(\theta)$ représente donc une borne asymptotique inférieure de la variance de tout estimateur déduit de la statistique s_n .

4 Application à la détection d'exoplanètes

4.1 Présentation du modèle

Nous nous limiterons ici à un modèle *extrêmement simplifié* de ce problème. En particulier nous ne prendrons pas en compte les divers bruits de lecture, les effets de rotation de champ

pendant la durée de pose, la dépendance temporelle des speckles (atmosphériques et quasi statiques), ...

Dans le cas de la détection de planète par imagerie directe, l'intensité mesurée sur un pixel résulte de la somme de l'intensité de l'éventuelle planète avec l'intensité de l'étoile diffractée au niveau de la planète. Ce dernier terme représente la limitation principale de la détection de planètes par imagerie directe et son atténuation nécessite l'utilisation d'un coronographe.

En ce qui concerne l'étoile, ψ résulte de la propagation du front d'onde de la lumière de l'étoile observée à travers l'atmosphère puis l'optique adaptative et le coronographe du télescope. La partie aléatoire de ψ représente les turbulences résiduelles de l'atmosphère non corrigées par l'optique adaptative et propagées à travers le coronographe. La moyenne de ψ est liée à la réponse impulsionnelle du coronographe et aux aberrations statiques du système optique. Nous supposons dans la suite que la moyenne de ψ est connue.

L'intensité lumineuse en provenance de la planète suit un modèle similaire. Si l'on considère que l'effet du coronographe est négligeable au niveau de la planète, la partie déterministe est simplement donnée par la réponse du télescope translatée à l'emplacement de cette planète. Dans la mesure où on se limite au voisinage immédiat de la planète, la partie non corrigée du front d'onde est négligeable par rapport à l'onde parfaite. On considère donc que l'intensité de la planète se résume à la réponse impulsionnelle du système translaté. Nous supposons également que cette réponse impulsionnelle est connue.

Étoiles et planètes étant incohérentes, l'intensité totale résulte de la somme des deux contributions. Dans la mesure où l'on teste la présence d'une planète en \vec{x} à partir des mesures dans un voisinage \mathcal{V} de \vec{x} , on peut écrire :

$$I(\vec{x}) = \text{diag} \left(\psi(\vec{x})\psi(\vec{x})^T \right) + \alpha m(\vec{x}), \quad (19)$$

où le vecteur $m(\vec{x})$ de dimension $M \times 1$ est connu mais la constante $\alpha \geq 0$ est inconnue. D'après (9), le modèle photonisé de (19) conduira à la somme de deux variables aléatoires indépendantes : la première est distribuée suivant une loi de Poisson de moyenne $\alpha m(\vec{x})$ est la seconde distribuée suivant le modèle décrit dans la partie 2. Les paramètres inconnus sont donc l'intensité de la planète α , et la covariance de l'amplitude complexe $\Sigma(\theta)$.

Afin de limiter la dimension du vecteur $\psi(\vec{x})$ et donc la complexité de la méthode d'estimation, le voisinage spatial \mathcal{V} étudié est réduit à la clique, décrite figure 1, de $M = 5$ pixels centrée sur le pixel à la position \vec{x} . Si l'on considère le cas de VLT SPHERE, l'échantillonnage spatial (dans la bande I) correspond à 2 pixels par largeur à mi hauteur de la tâche d'Airy $m(\vec{x})$. On peut ainsi considérer que si une planète est présente au niveau du pixel central de la clique de la figure 1, son effet en dehors de la clique sera négligeable, ce qui justifie pleinement ce choix pour la détection.

La structure de corrélation associée à cette clique est décrite figure 2. Nous supposons dans la suite que l'autocorrélation de l'amplitude complexe décroît suivant une loi géométrique. Dans ce cas, un raisonnement géométrique, amène à poser :

$$\rho_x = \rho_y = \rho, \quad \gamma_x = \gamma_y = \rho^2 \quad \text{et} \quad \nu = \eta = \rho^{\sqrt{2}} \quad \text{où} \quad \rho \in \mathbb{C}. \quad (20)$$

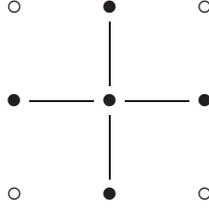


FIG. 1 – Clique de 5 pixels

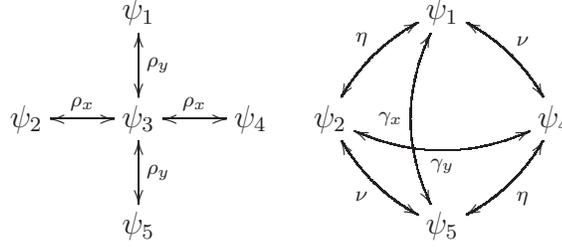


FIG. 2 – Structure de la corrélation du vecteur ψ .

La matrice de covariance complexe $\Sigma(\theta)$ s'écrit alors :

$$\Sigma(\theta) = \sigma^2 \begin{pmatrix} 1 & \rho\sqrt{2} & \rho & \rho\sqrt{2} & \rho^2 \\ (\rho^*)\sqrt{2} & 1 & \rho & \rho^2 & \rho\sqrt{2} \\ \rho^* & \rho^* & 1 & \rho & \rho \\ (\rho^*)\sqrt{2} & (\rho^*)^2 & \rho^* & 1 & \rho\sqrt{2} \\ (\rho^*)^2 & (\rho^*)\sqrt{2} & \rho^* & (\rho^*)\sqrt{2} & 1 \end{pmatrix}. \quad (21)$$

Le domaine de définition du paramètre complexe ρ correspond aux valeurs de ce paramètre pour lesquelles la matrice hermitienne $\Sigma(\theta)$ est définie positive. Il est difficile d'obtenir une expression analytique des valeurs propres de $\Sigma(\theta)$. Cependant, ces dernières peuvent être évaluées numériquement en fonction des valeurs du paramètre ρ . La figure 3 représente en blanc dans la plan complexe les valeurs du paramètre ρ pour lesquelles la matrice $\Sigma(\theta)$ est bien une matrice de covariance. On peut remarquer que l'intervalle réel $[0, 1[$ appartient bien au domaine de définition du paramètre ρ .

En définitive, le vecteur des paramètres inconnus est $\theta = (\alpha, \sigma^2, \rho)$, le paramètre α étant nul en l'absence de planète, strictement positif sinon. Cet article étudie finalement un test d'hypothèses binaires permettant la détection de la planète et défini comme suit :

$$\begin{aligned} H_0 & \text{ (absence de planète) : } \alpha = 0, \\ H_1 & \text{ (présence de planète) : } \alpha > 0. \end{aligned} \quad (22)$$

4.2 Estimateurs

La moyenne de l'amplitude complexe du front d'onde étant non nulle dans ce cas, le vecteur θ sera estimé par les méthodes de moments présentées plus haut. On obtient pour tout $1 \leq i \leq M$ et $1 \leq k < l \leq M$:

$$\mathbb{E}[N_i] = \sigma^2 + 2\mu_i^2 + \alpha m_i, \quad (23)$$

$$\text{var}[N_i] = \sigma^4 + \sigma^2(4\mu_i^2 + 1) + 2\mu_i^2 + \alpha m_i, \quad (24)$$

$$\text{cov}(N_k, N_l) = |\Sigma(\theta)_{k,l}|^2 + 4\mu_k\mu_l\Re(\Sigma(\theta)_{k,l}), \quad (25)$$

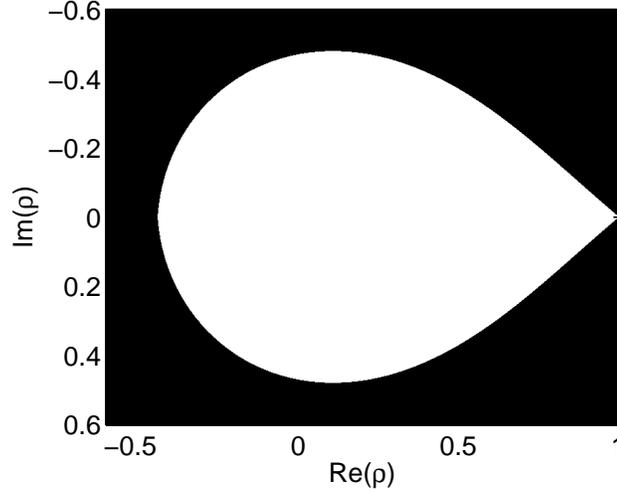


FIG. 3 – Domaine (en blanc) de définition du paramètre $\rho \in \mathbb{C}$ dans le plan complexe.

où les $\Sigma(\theta)_{k,l}$ sont les éléments de la matrice de covariance (21) du vecteur ψ , et $\Re(z)$ est la partie réelle de z .

4.2.1 MCNL

Nous allons considérer un estimateur MCNL, noté $\hat{\theta}_2$, du vecteur $\theta = (\alpha, \sigma^2, \rho)^T$ défini en (15) reposant sur la statistique $\mathbf{s}_n = \frac{1}{n} \sum_{i=1}^n h(N^{[i]})$ associée à la fonction

$$h(x) = \left((x_i)_{1 \leq i \leq M}, (x_i^2)_{1 \leq i \leq M}, (x_i x_j)_{1 \leq i < j \leq M} \right)^T. \quad (26)$$

Les moments d'ordre 1, d'ordre 2 et les moments joints d'ordre 2 sont donc pris en compte. Les équations (23, 24, 25) assurent alors que les paramètres réels α , σ^2 et $\rho \in \mathbb{C}$ à estimer sont identifiables à partir de l'expression de ces moments.

4.2.2 Moments

L'estimateur MCNL est comparé avec l'estimateur des moments $\hat{\theta}_1$ du vecteur $\tilde{\theta} = (\alpha, \sigma^2)^T$ basé sur la statistique $\tilde{\mathbf{s}}_n = \frac{1}{n} \sum_{i=1}^n \tilde{h}(N^{[i]})$ associée à la fonction

$$\tilde{h}(x) = \left((x_i)_{1 \leq i \leq M}, (x_i^2)_{1 \leq i \leq M} \right)^T. \quad (27)$$

Cet estimateur des moments ne prend donc pas en compte les corrélations inter-pixels. En posant $\tilde{f}(\tilde{\theta}) = \mathbb{E}[\tilde{\mathbf{s}}_n]$, on trouve alors à partir des équations (23, 24) l'expression suivante d'une fonction $\tilde{g}(y) = (\tilde{g}_1(y), \tilde{g}_2(y))^T$ de $\mathbb{R}^{2M} \rightarrow \mathbb{R}^2$ telle que $\tilde{g}(\tilde{f}(\tilde{\theta})) = \tilde{\theta}$:

$$\begin{aligned} \tilde{g}_2(y) &= \frac{1}{M} \sum_{i=1}^M \frac{-4\mu_i^2 + \sqrt{\Delta_i(y)}}{2}, \\ \tilde{g}_1(y) &= \frac{1}{M} \sum_{i=1}^M \frac{y_i - 2\mu_i - \tilde{g}_2(y)}{m_i}, \end{aligned} \quad (28)$$

où $\Delta_i(y) = (16\mu_i^4 - 4(2(\mu_i^2 - \mu_i) + y_i + y_i^2 - y_{i+M}))$. L'estimateur des moments défini en (15) est alors $\hat{\theta}_1 = \tilde{g}(\tilde{\mathbf{s}}_n)$. Par ailleurs la statistique $\tilde{\mathbf{s}}_n$ étant la restriction de \mathbf{s}_n aux moments d'ordre 1 et 2, la variance asymptotique de cet estimateur est supérieure à celle de l'estimateur MCNL $\hat{\theta}_2$ selon (18).

4.3 Simulations

Plusieurs simulations ont été menées afin d’apprécier l’apport de la modélisation multi-variée introduite, et les performances des estimateurs retenus. Dans les simulations de cet article les paramètres sont fixés aux valeurs suivantes :

- La moyenne μ des amplitudes complexes est un vecteur dont tous les termes sont identiques :

$$\mu = 2e^{i\pi/4}(1, 1, 1, 1, 1)^T \quad (29)$$

- La signature de la planète m est un vecteur dont tous les termes sont identiques :

$$m = 2(1, 1, 1, 1, 1)^T \quad (30)$$

- la variance des amplitudes complexes est $\sigma^2 = 10$
- l’intensité de la planète est $\alpha = 0.1$ en présence de planète, $\alpha = 0$ sinon

Soulignons que ces valeurs sont données à titre d’exemple et ne correspondent pas à une configuration physique particulière.

La figure 4 représente les variances asymptotiques théoriques des estimateurs $\hat{\theta}_1$ (Moment) et $\hat{\theta}_2$ (MCNL) du paramètre α en fonction du paramètre ρ pour $0 \leq \rho \leq 1$. Cette

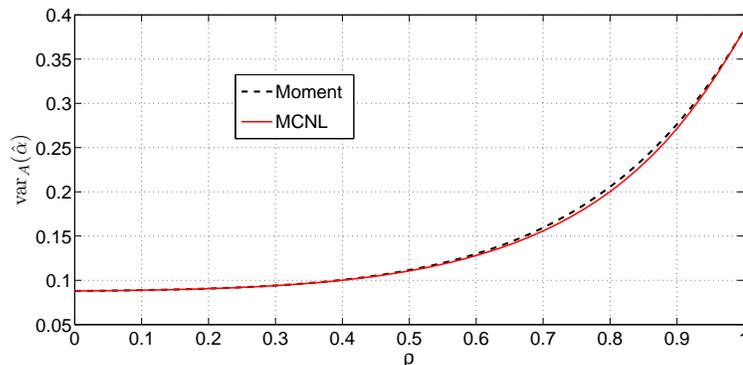


FIG. 4 – Variances asymptotiques des estimateurs du paramètre α ($\alpha = 0.1$) en fonction du paramètre de corrélation ρ .

figure montre que les performances des deux méthodes sont très proches, voire identiques dans les cas limites $\rho = 0$ et $\rho = 1$. On peut également constater que lorsque la corrélation spatiale du front d’onde croît, i.e. lorsque ρ tend vers 1, la variance asymptotique des estimateurs augmente. En effet, les pixels voisins de la clique n’apportent que peu d’information supplémentaire par rapport au pixel central lorsque les observations sont fortement corrélées.

Afin de représenter les courbes Caractéristiques Opérationnelles du Récepteur (COR) -probabilité de détection en fonction de la probabilité de fausse alarme- empiriques associées au problème de détection présenté en (22), des données synthétiques ont été générées : 50000 échantillons de $n = 1000$ cliques (poses de 7 sec. pendant 2 h.) ont été générées en présence de planète ($\alpha = 0.1$) et 50000 sans planète ($\alpha = 0$). Le paramètre de corrélation est $\rho = 0.7$ dans ces simulations. Les courbes COR empiriques obtenues sont représentées figure 5(a), pour les détecteurs basés sur les estimateurs $\hat{\theta}_1$ (Moment) et $\hat{\theta}_2$ (MCNL). Ces résultats confirment ceux obtenus sur la performance des estimateurs, à savoir que les détecteurs basés sur ces deux estimateurs ont des performances identiques. De plus la courbe COR théorique obtenue sous l’hypothèse de normalité asymptotique de l’estimateur $\hat{\theta}_2$ est aussi représentée (Asymptotique MCNL) figure 5(b). Cette courbe correspond parfaitement à

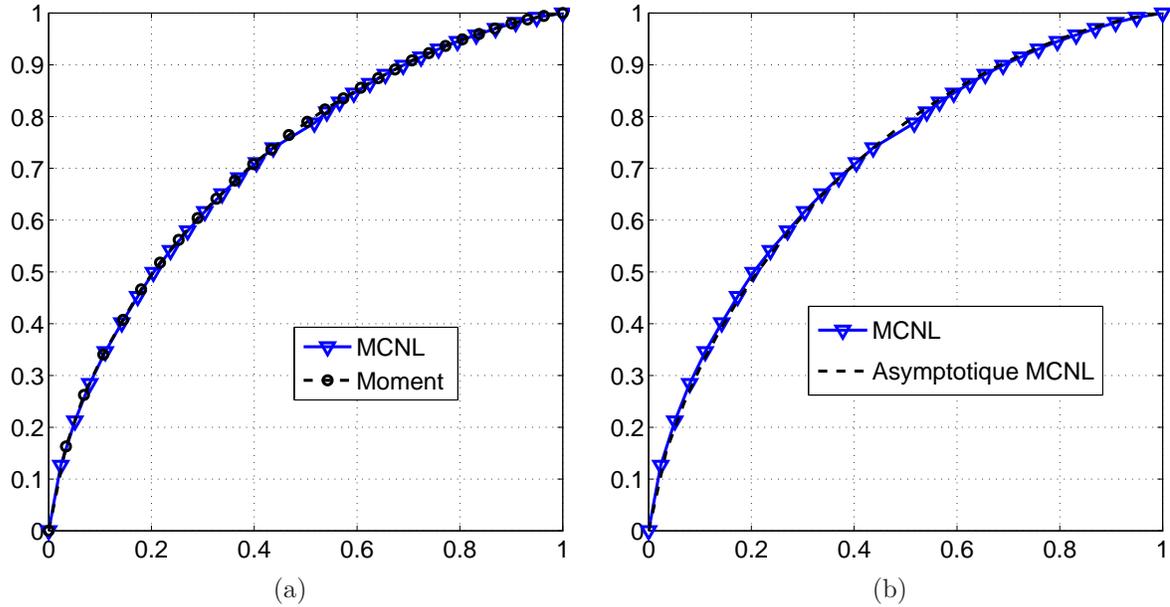


FIG. 5 – Courbes COR ($n = 1000$, $\rho = 0.7$, $\alpha = 0.1$ en présence de planète 0 sinon)

la courbe empirique obtenue pour des échantillons de $n = 1000$ cliques. En définitive, les simulations effectuées ne permettent pas de mettre en évidence une amélioration significative des performances de détection lors de l'estimation conjointe de $(\alpha, \sigma^2, \rho)^T$ (estimateur $\hat{\theta}_2$) par rapport au cas plus simple où seul $(\alpha, \sigma^2)^T$ (estimateur $\hat{\theta}_1$) est estimé.

5 Conclusion

Dans ce travail, nous nous sommes intéressés à l'étude de familles de lois de Poisson mélangées multivariées. Cette famille est centrale pour la modélisation de la distribution statistique des photons, par exemple au foyer d'un télescope. Des méthodes d'estimation appropriées, les moindres carrés non linéaires, ont alors été étudiées afin de prendre en compte les dépendances spatiales du front d'onde lorsque ce dernier n'est pas à moyenne nulle.

En dépit des propriétés d'optimalité asymptotique théorique de ces méthodes, les performances obtenues pour la détection ne sont pas significativement meilleures que celles obtenues pour des méthodes reposant sur les seules statistiques univariées des données observées, et donc ne prenant pas en compte les dépendances spatiales du front d'onde. Afin d'améliorer ces performances de détection, il serait nécessaire d'introduire un a priori plus fort sur les paramètres modélisant la corrélation spatiale. Cet a priori pourrait être issu d'une estimation desdits paramètres à partir de données spécifiques issues, par exemple, de l'analyseur de front d'onde, ou encore, d'une modélisation statistique précise de la distribution de ces paramètres.

Références

- [1] C. Aime. Analysis of the technique of dark speckles for detection of exo-planets. *Journal of Optics A : Pure and Applied Optics*, pages 411–421, 2000.

- [2] C. Aime and R. Soummer. Influence of speckle and Poisson noise on exoplanet detection with a coronagraph. In *EUSIPCO-04*, pages 509–512, Vienna, Austria, Sept. 2004. Elsevier.
- [3] C. Aime and F. Vakili, editors. *Direct Imaging of Exoplanets : Science & Techniques*. International Astronomical Union Colloquium 200, Cambridge University Press, 2006.
- [4] S. K. Bar-Lev, D. Bshouty, P. Enis, G. Letac, I-Li Lu, and D. Richards. The diagonal multivariate natural exponential families and their classification. *J. of Theoret. Probab.*, 7(4) :883–929, Oct. 1994.
- [5] P. Bernardoff. Which negative multinomial distributions are infinitely divisible? *Bernoulli*, 9(6), 2003.
- [6] P. Bernardoff. Which multivariate Gamma distributions are infinitely divisible? *Bernoulli*, 12(1) :169–189, 2006.
- [7] V.F. Canales and M.P. Cagigal. Photon statistics in partially compensated wave fronts. *Journal of the Optical Society of America*, 16 :2550—2554, oct 1999.
- [8] F. Chatelain, A. Ferrari, and J.-Y. Tourneret. Parameter estimation for multivariate mixed poisson distributions. In *IEEE ICASSP*, volume 3, pages 684–687, Toulouse, May. 2006.
- [9] F. Chatelain and J.-Y. Tourneret. Composite likelihood estimation for multivariate mixed Poisson distributions. In *Proc. IEEE-SP Workshop Stat. Signal Processing*, Bordeaux, France, July 2005.
- [10] Florent Chatelain, Sophie Lambert-Lacroix, and Jean-Yves Tourneret. Pairwise likelihood estimation for multivariate mixed Poisson models generated by gamma intensities. *Statistics and Computing*. To appear (published online in Sept. 2008).
- [11] A. Ferrari and C. Aime. Étude statistique de la détection d’exo-planètes en imagerie courte pose. In *Actes du 18ème colloque GRETSI*, pages 669–672, Toulouse, 2001. <http://hdl.handle.net/2042/13384>.
- [12] A. Ferrari, G. Letac, and J.-Y. Tourneret. Multivariate mixed Poisson distributions. In *EUSIPCO*, Vienna, Austria, Sept. 2004. Elsevier.
- [13] J. Goodman. *Statistical Optics*. Wiley, New York, 1985.
- [14] J.W. Goodman. Statistical properties of laser speckle pattern. In J.C. Dainty, editor, *Laser Speckle and Related Phenomena*, pages 9–75. Springer-Verlag, Berlin, 1975.
- [15] J. Grandell. *Mixed Poisson Processes*. Chapman and Hall, 1997.
- [16] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 2. John Wiley, New York, 2nd edition, 1995.
- [17] N. L. Johnson, S. Kotz, and A. W. Kemp. *Univariate discrete distributions*. John Wiley, New York, 2nd edition, 1992.
- [18] A. Labeyrie. Images of exo-planets obtainable from dark speckles in adaptive telescopes. *Astronomy and Astrophysics*, 298 :544–548, 1995.
- [19] P McCullagh. *Tensor methods in statistics*. Monographs on statistics and applied probability. Chapman and Hall, London, 1987.
- [20] Robb J Muirhead. *Aspects of multivariate statistical theory*. Wiley series in probability and mathematical statistics. Wiley, New York, 1982.

- [21] Bernard Picinbono. *Théorie des signaux et des systèmes avec problèmes résolus*. Dunod Université, Paris, 1989.
- [22] B. Porat and B. Friedlander. Performance analysis of parameter estimation algorithms based on high-order moments. *International Journal of adaptive control and signal processing*, 3 :191–229, 1989.
- [23] J.-Y. Tournéret, A. Ferrari, and G. Letac. The noncentral Wishart distribution : Properties and application to speckle imaging. In *Proc. IEEE-SP Workshop Stat. Signal Processing*, Bordeaux, France, jul 2005.
- [24] C. Varin. On composite marginal likelihoods. *Advances in Statistical Analysis*, 92(1) :1–28, fev. 2008.