



## ETUDE COMPARATIVE DE DEUX ALGORITHMES D'INVERSION DE LA TRANSFORMATION DE POISSON

F. SULTANI, H. LANTERI et C. AIME

Département d'Astrophysique  
U.R.A. 709 du CNRS  
Université de Nice — Sophia Antipolis  
Parc Valrose — 06108 Nice Cedex 2  
FRANCE

### RESUME

Deux méthodes de minimisation appliquées à un problème inverse à deux variables, intervenant en optique statistique, sont présentées. Dans la première méthode le bruit de mesure est considéré comme gaussien, ce qui conduit à la solution des moindres carrés, alors que dans la seconde méthode le bruit est supposé de nature poissonnienne. Cette hypothèse conduit à la classe des algorithmes EM.

### ABSTRACT

Two methods to solve an inverse problem, that intervenes in statistical optics, are presented. In the first case we assume that the noise is Gaussian. This assumption leads to the least square solution. In the second case, we assume that the noise has a Poisson statistics. That leads to the class of EM algorithms.

### I- INTRODUCTION

La transformation de Poisson (TP) [1] intervient en optique statistique pour les faibles flux lumineux où l'observation s'effectue en comptage de photons. Nous considérons le cas bidimensionnel. La TP est la transformation qui, appliquée à la densité de probabilité (DDP)  $P(\Omega_1, \Omega_2)$  d'une intensité lumineuse aléatoire, donne la répartition du nombre de photons détectés  $p(n, q)$ .

$$p(n, q) = \int_0^\infty \int_0^\infty K(\Omega_1, n, \Omega_2, q) P(\Omega_1, \Omega_2) d\Omega_1 d\Omega_2 \quad (1)$$

où  $p(n, q)$  représente la DDP en comptage de photons,  $K(n, \Omega_1, q, \Omega_2)$  le noyau d'intégration à variables séparables s'écrivant :

$$K(\Omega_1, n, \Omega_2, q) = \frac{\exp(-\Omega_1) \Omega_1^n}{n!} \frac{\exp(-\Omega_2) \Omega_2^q}{q!} \quad (2)$$

et  $P(\Omega_1, \Omega_2)$  la DDP à fort flux, quantité que l'on veut estimer. L'inversion de la relation (1) passe par une étape de discrétisation qui conduit à

$$p = HPH^t \quad (3)$$

où  $p$  est un tableau, assimilable à une matrice de dimension  $m \times m$ , représentant les données mesurées,  $H$  l'opérateur de la transformation ( $m \times m$ ),  $P$  le tableau des inconnues ( $m \times m$ ) et  $H^t$  la transposée de la matrice  $H$ .

La relation (3) peut être présentée sous une forme linéaire qui s'écrit :

$$p' = hP' \quad (4)$$

où  $p'$  et  $P'$  sont des vecteurs de dimension  $m^2 \times 1$  représentant respectivement les données mesurées et les inconnues et  $h$  la matrice de la transformation de dimension  $m^2 \times m^2$  qui est le produit de Kronecker [2] de la matrice  $H$  par elle-même. Une étude de la relation (4) a été présentée récemment par notre équipe [3].

L'objet de cette communication est l'étude d'un autre aspect de ce problème où l'on exploite la nature séparable du noyau d'intégration. La méthode proposée consiste à résoudre l'équation (3) en deux étapes. Dans une première étape, on pose

$$G = PH^t \quad (5)$$

En reportant la relation (5) dans la relation (3), on est donc conduit à résoudre :



$$p = HG \quad (6)$$

où les inconnues sont contenues dans  $G$ . Une fois l'estimation de  $G$  effectuée la deuxième étape va consister à estimer  $P$  par la relation suivante

$$G^t = HP^t \quad (7)$$

Nous proposons une étude de ce problème en maximisant l'estimateur de la vraisemblance.

Dans un souci de clarté nous développons notre étude en nous basant sur une relation tout à fait générale entre un objet  $X$  et sa transformée  $Y$  telle que :

$$Y = AX \quad (8)$$

équation tout à fait semblable aux relations (6) et (7). Dans la réalité, on dispose de  $Y$  qui est une version de  $Y$  entachée d'une erreur de mesure. Le problème est donc d'obtenir  $X$  à partir de  $Y$  compte tenu du modèle et connaissant la nature de bruit. Nous allons étudier l'effet de ce dernier point dans le cas de notre problème.

Deux modèles de formation de données vont être présentés. Dans le premier on considère que le bruit est de statistique gaussienne, ce qui conduit à la solution des moindres carrés et à la classe des algorithmes additionnels de type gradient. Dans le second cas on décrit l'effet de la statistique poissonnienne lors de la formation des données. Dans ce cas on est conduit à une autre classe d'algorithmes, multiplicatifs, basés sur le principe de Expectation Maximization (EM) [4].

La communication est organisée de la façon suivante. Les sections II et III sont consacrées à l'étude de la statistique gaussienne et poissonnienne du bruit. Et dans la section IV on présente l'application de ces deux méthodes à un modèle classique en optique statistique suivie de la conclusion.

## II- STATISTIQUE GAUSSIENNE

La vraisemblance de  $X$ ,  $L(X)$ , est une loi de probabilité conditionnelle qui s'écrit  $P(Y|X)$ . Maximiser cette loi, ou plutôt son logarithme, revient à estimer l'objet  $X$  qui serait le mieux en conformité avec les données  $Y$ . Par ailleurs, la loi conditionnelle  $P(Y|X)$  décrit la loi du bruit [5]. Si  $b$  est un bruit additif et indépendant du signal cela s'écrit :

$$\tilde{Y} = Y + b \quad (9)$$

où  $\tilde{Y}$  est une version entachée de bruit de  $Y$  et  $b$  est le

terme du bruit qui s'y ajoute. Si de plus le bruit est gaussien de moyenne nulle et d'écart type  $\sigma$ , alors la vraisemblance de  $X$  en tenant compte de l'indépendance des pixels  $i$  et  $j$ , s'écrit

$$L(X) = P(Y|X) = \prod_{ij} \exp\left(-\frac{(Y_{ij} - f_{ij})^2}{2\sigma^2}\right) \quad (10)$$

avec

$$f_{ij} = \sum_k A_{ik} X_{kj} \quad (11)$$

La solution  $X$  s'obtient en minimisant  $-\log(L(X))$ . Ce qui conduit à la solution des moindres carrés

$$\text{Min}(-\log(L)) \propto \text{Min} \sum_{ij} (Y_{ij} - f_{ij})^2 \quad (12)$$

La relation (12) n'est autre que la minimisation d'une distance Euclidienne.

$$\text{Min} \|Y - AX\|^2 \quad (13)$$

La minimisation de la relation (12) peut être effectuée par des algorithmes classiques de gradient, qui sont des algorithmes itératifs additifs. Nous avons estimé la solution  $X$  par l'algorithme de gradient à pas optimal qui s'écrit :

$$X_{k+1} = X_k + \alpha_k A^t (Y - AX_k) \quad (14)$$

où  $\alpha_k$  est le pas de descente [6].

L'application de cette méthode aux relations (6) et (7) conduit à :

$$G_{k+1} = G_k + \alpha_k H^t (p - HG_k) \quad (15)$$

Si  $G^*$  est une estimation de  $G$  au bout d'un certain nombre d'itérations, l'inconnue  $P$  peut être obtenue par :

$$P^t_{k+1} = P^t_k + \alpha_k H^t (G^* - HP^t_k) \quad (16)$$

La solution de maximum de vraisemblance peut présenter des instabilités lorsque le nombre d'itérations augmente. Ceci est essentiellement dû au caractère mal posé de notre problème inverse. Afin d'éviter des solutions instables on arrête les itérations avant l'apparition des instabilités. Les itérations sont arrêtées lorsque l'estimateur  $\chi^2$  [7] est inférieur à un seuil fixé.

$$\chi^2 = \sum_{ij} \frac{(\tilde{p}_{ij} - p_{ij})^2}{p_{ij}} \quad (17)$$

$p$  est obtenu par la relation (3) à partir de la solution  $P$ .

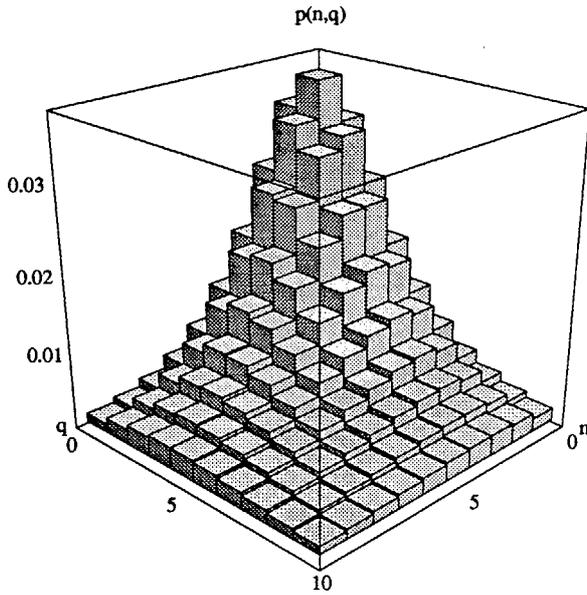


fig 1 : Illustration des données en comptage de photons,  $p(n,q)$ , d'une structure de speckle d'une étoile double de moyenne 2 photons par pixel et de rapport d'intensité de 1.4.

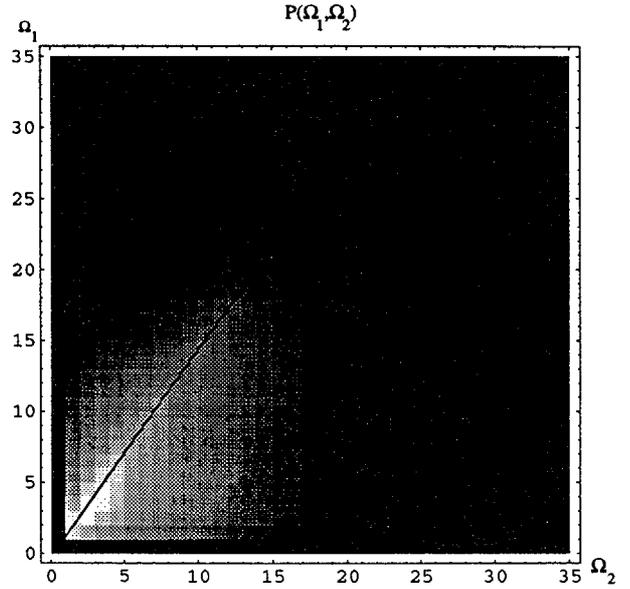


fig 2 : La densité de probabilité théorique,  $P(\Omega_1, \Omega_2)$  correspondant aux données  $p(n,q)$  montrée par la figure 1.

### III- STATISTIQUE POISSONNIENNE

Dans le cas de la statistique de Poisson l'estimateur de la vraisemblance de  $X$ , en tenant compte de l'indépendance des pixels  $i$  et  $j$ , s'écrit comme [5]

$$L(X) = P(Y | X) = \prod_{ij} \exp(-f_{ij}) \frac{f_{ij}^{Y_{ij}}}{Y_{ij}!} \quad (18)$$

avec  $f_{ij}$  donné par la relation (11).

On est conduit à minimiser la relation suivante :

$$\text{Min} \left( \sum_{ij} -f_{ij} + Y_{ij} \log f_{ij} + \text{const} \right) \quad (19)$$

La minimisation de la relation (19) par la méthode EM conduit à l'algorithme multiplicatif suivant :

$$X_{k+1} = X_k \bullet A^t * \frac{Y}{A * X_k} \quad (20)$$

où  $\bullet$  représente le produit élément par élément et  $*$  désigne le produit matriciel classique. Cet algorithme est couramment appelé algorithme de Lucy [9] ou Richardson [8] en astronomie. Mais ces deux auteurs ont établi cette relation en se basant sur la loi de Bayes sans faire aucune allusion sur la formation des données.

L'application de cette méthode aux relations (6) et (7) conduit à

$$G_{k+1} = G_k \bullet H^t * \frac{P}{H * G_k} \quad (21)$$

On désigne par  $G_0$  une estimation de  $G$  ce qui conduit à

$$P_{k+1}^t = P_k^t \bullet H^t * \frac{G_0}{H * P_k^t} \quad (22)$$

Dans les deux cas la solution  $P$  est obtenue par une transposition de la solution des relations (16) et (22).

### IV- RESULTATS ET CONCLUSION

Les données en comptage de photons utilisées comme mesures lors de cette étude sont simulées à partir d'un modèle classique en optique statistique (fig 1). Elles correspondent à la DDP de la répartition de brillance produite par la somme de deux structures de speckles [10,11]. En astronomie ceci a des applications pour la mesure des paramètres d'étoiles. Notre but est d'estimer le rapport d'intensité entre les deux composantes d'une binaire. Ce rapport d'intensité peut être estimé approximativement en évaluant la tangente de l'angle indiqué sur la figure 2 [12].

Nous avons proposé une étude de deux modèles de for-

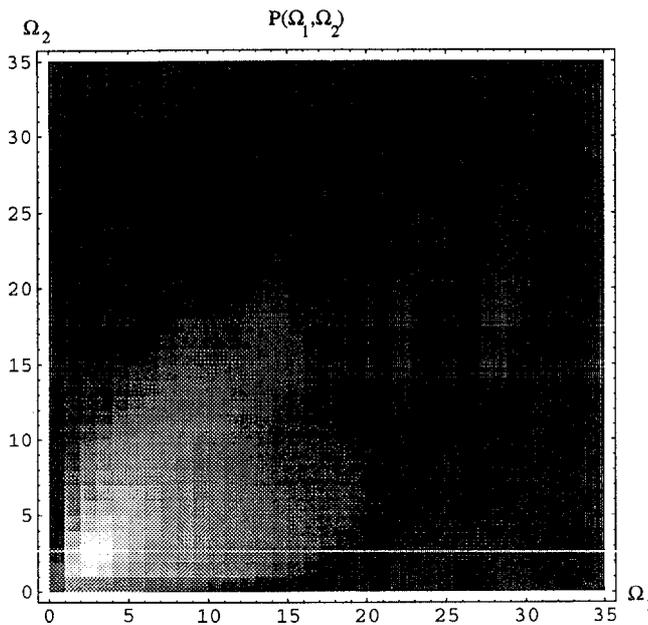


fig 3 : Le DDP reconstruite par la première méthode (bruit gaussien). Figure obtenue au bout de 800 itérations.

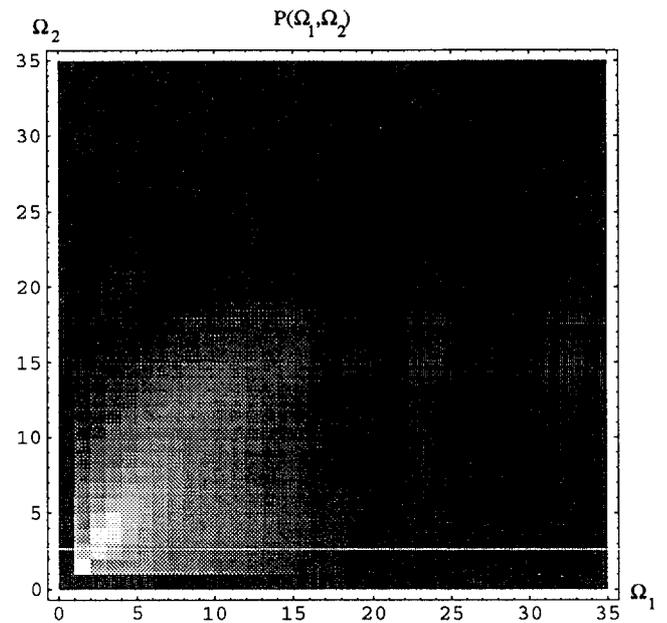


fig 4 : La solution obtenue par la méthode où le bruit est considéré comme poissonnien. Solution obtenue au bout de 800 itérations.

mation de données appliqués à un problème inverse à deux variables. Il est mis en évidence que la statistique du bruit joue un rôle important dans le choix du critère à minimiser. Si la statistique du bruit suit la loi de Gauss, la solution recherchée est une solution des moindres carrés qui est obtenue par un algorithme additif de type gradient. Alors que pour le cas Poissonnien, on doit minimiser un critère peu habituel qui conduit à l'algorithme multiplicatif de Lucy- Richardson. La première méthode, malgré une vitesse de convergence légèrement plus rapide ne garantit pas la positivité de la solution, on doit alors appliquer une contrainte de positivité après chaque itération.

Par contre la deuxième méthode à défaut d'être rapide a l'avantage de préserver la positivité de la solution, ce qui est non négligeable lorsque l'on traite une grandeur comme la DDP. La seule condition exigée pour satisfaire cette contrainte est de commencer le calcul avec un estimé initial positif.

Dans notre problème c'est la deuxième technique qui donne les meilleurs résultats.

### Remerciements

Les auteurs tiennent à remercier leur collègue Eric Aristidi pour son aide lors de la rédaction de cette communication.

### REFERENCES

- [1] : Mandel L, 1953 *Proc.Phys. Soc* **74**, 233.
- [2] : Andrews. H. C and Hunt. B. R, 1977 *Digital Image restoration*, Prentice-Hall Signal Processing Serie.
- [3] : Sultani. F, Lantéri. H et Aime. C, 1995 *CEPHAC-MCEA*.
- [4] : Shepp. L and Vardi Y, 1982 *IEEE.Trans. Med. Imaging* **2** 113-121.
- [5] : Nunez. J and Llacer. J, 1990 *IEEE. Trans. Med. Imaging* **9** 159-171.
- [6] : Minoux. M. 1990 *Programmation Mathématique* Dunod.
- [7] : Bédard. G, 1967 *J. O. S. A. A* **57** 1132-1135.
- [8] : Richardson. W. H, 1972 *J. O. S. A. A* **62** 55-59.
- [9] : Lucy. L, 1974 *Astronomical Journal* **79** 745-754.
- [10] : Aime C, 1993 *Trends in Optical Engineering* **1** 15-34 Kerala India.
- [11] : Goodman J.W, 1985 "Statistical Optics", J.Wiley & Son.
- [12] : Aristidi. E, 1992 *Thèse de Doctorat Université de Nice-Sophia Antipolis*.