PARAMETER ESTIMATION FOR MULTIVARIATE MIXED POISSON DISTRIBUTIONS

Florent Chatelain[†]*, André Ferrari*[*] *and Jean-Yves Tourneret*[†]

[†] IRIT/ENSEEIHT/TéSA, 2 rue Charles Camichel, BP 7122, 31071 Toulouse cedex 7, France
[*] LUAN, Université de Nice-Sophia-Antipolis, 06108 Nice cedex 2, France
florent.chatelain@enseeiht.fr, ferrari@unice.fr, jean-yves.tourneret@enseeiht.fr

ABSTRACT

Estimating the parameters of multivariate distributions whose densities or masses cannot be expressed in tractable closed-form is a challenging problem. This paper concentrates on a family of such discrete distributions referred to as multivariate mixed Poisson distributions (MMPDs). These distributions are interesting for modeling correlations between adjacent pixels of active and astronomical images. Several estimators of MMPD parameters are investigated. These estimators include a composite likelihood estimator and a non-linear least squares estimator.

1. INTRODUCTION

This communication addresses the problem of estimating the statistical properties of wavefront amplitudes from intensity measurements. It assumes that the wavefront amplitude results from the sum of incoherent complex Gaussian components. Moreover it considers that the resulting intensities are very low and are recorded using a photocounting camera. This model arises in many optical systems encountered in active and astronomical imagery:

- Active imaging consists of forming an image of a scene which has been illuminated with laser light. The main advantages of this technique are to allow night vision and to improve image resolution for a given aperture size (see [1] for more details). When the intensity level of the reflected light is very low, the observed images are corrupted by two sources of noise: speckle noise and Poisson noise. The speckle fluctuations are classically modeled by a Gamma distribution of order L which results from the sum of L incoherent waves with *zero mean* Gaussian distributed complex amplitudes.
- Increasing interest has been shown in the astronomical community for the direct imaging of extrasolar planets. As explained in [2], the complex amplitude of a wave in the focal plane of a telescope is the sum of a *deterministic term* proportional to the wave amplitude in absence of turbulence and a wavefront amplitude (associated to the speckles) distributed as a zero mean complex circular Gaussian distribution. The speckle fluctuations are modeled in this case by a Rice distribution. The Poisson noise arises from the short exposure times and the low intensity of the observed objects.

This paper is organized as follows. Section 2 discusses the statistical properties of the observed data. Section 3 is devoted to the estimation of the wavefront parameters. In the general case, two methods of moments are compared: a "classical" approach and an optimal one which is based on a non-linear least squares minimization. In the special case were the wavefronts are zero mean, the estimation can also be achieved by maximizing a composite likelihood of the measurements. Section 4 compares the performance of the different estimators. Conclusions are reported in section 5.

2. SIGNAL MODEL AND PROBLEM FORMULATION

2.1. Intensity Distribution

We assume that the $M \times 1$ vectors containing the complex amplitudes $\boldsymbol{\psi}(k)$ of the L incoherent waves are independent circular Gaussian vectors:

$$\boldsymbol{\psi}(k) \sim \mathcal{N}_c(\mu(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})), \tag{1}$$

where $k = 1, ..., L$, $\boldsymbol{\mu}(\boldsymbol{\theta}) \in \mathbb{R}^M$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is an $M \times M$ covariance matrix. The resulting $M \times 1$ vector of intensities $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_M)^T$ has components:

$$\lambda_q = \sum_{k=1}^{L} |\psi_q(k)|^2, \qquad q = 1, ..., M. \tag{2}$$

Eq. (2) shows that λ_k is proportional to a random variable distributed according to a noncentral χ^2 distribution with $2L$ degrees of freedom [3]. Note that:

- the case $\boldsymbol{\mu}(\theta) = 0$ leads to the Gamma distribution of order L encountered in active imaging,
- the case $L = 1$ leads to the Rice distribution (or equivalently a noncentral χ^2 distribution with 2 degrees of freedom) mentioned above.

The multivariate distribution of $\boldsymbol{\lambda}$ is more complicated to derive. It can be obtained by noting that $\boldsymbol{\lambda}$ is the diagonal of the $M \times M$ matrix $\sum_{k=1}^{L} \boldsymbol{\psi}(k)\boldsymbol{\psi}(k)^H$ which has a noncentral Wishart distribution [4]. Consequently, the distribution of $\boldsymbol{\lambda}$ is the ad-hoc marginal of this distribution[1]:

$$\boldsymbol{\lambda} \sim \text{"diagonal of"}\ \mathcal{W}_m(2L, \boldsymbol{\Sigma}(\boldsymbol{\theta})/2, \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{\mu}(\boldsymbol{\theta})\boldsymbol{\mu}(\boldsymbol{\theta})^t).$$

This approach has been adopted in [5] and has allowed to derive a general formula to compute the moments of $\boldsymbol{\lambda}$. However, whereas the moment generating function of $\boldsymbol{\lambda}$ has a simple closed form expression, the computation of its probability distribution function (pdf) is untractable in the general case.

This work was supported by the CNRS under MathSTIC Action No. 80/0244.

[1] This result allows to derive identifiability conditions for model (1,2). These conditions are not developed here for a problem of space.

2.2. Photocount distribution

The vector of intensities $\boldsymbol{\lambda}$ described before corresponds to the case where the image has been recorded under a high flux assumption. However, for low-flux objects or short exposure time, the photo-counting effect has to be considered. Denote as N_i the number of photons associated to the intensity λ_i. Conditioned upon the vector of intensities $\boldsymbol{\lambda}$, the random variables $N_i, i = 1, \ldots, M$ are independent and distributed according to Poisson distributions with means λ_i. In this case, the probability masses of $\boldsymbol{N} = (N_1, \ldots, N_M)$ are defined as:

$$\Pr(\boldsymbol{N} = \boldsymbol{k}) = \int_{(\mathbb{R}^+)^d} \cdots \int \prod_{\ell} \frac{(\lambda_\ell)^{k_\ell}}{k_\ell!} \exp(-\lambda_\ell) p(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \quad (3)$$

where $\boldsymbol{k} = (k_1, ..., k_d) \in \mathbb{N}^d$ and $p(\boldsymbol{\lambda})$ is the pdf of $\boldsymbol{\lambda}$ defined in section 2.1. The distribution of $\boldsymbol{N}$ is a multivariate mixed Poisson distribution. Eq. (3) is also known as the Poisson-Mandel transform of $p(\boldsymbol{\lambda})$ in the semiclassical theory of photodetection [6].
Tractable expressions of $\Pr(\boldsymbol{N} = \boldsymbol{k})$ defined in (3) are obviously difficult to obtain. However, many interesting properties regarding the distribution of $\boldsymbol{N}$ can be derived in the monovariate [7] and in the multivariate cases [8]. In particular, the joint moments of $\boldsymbol{N}$ can be computed as follows:

$$\mathsf{E}\left[\prod_{k=1}^{d} N_k^{r_k}\right] = \sum_{j_1=0}^{r_1} \cdots \sum_{j_d=0}^{r_d} \prod_{k=1}^{d} S(r_k, j_k) \mathsf{E}\left[\prod_{k=1}^{d} \lambda_k^{j_k}\right], \quad (4)$$

where $S(j, k)$ are the Stirling numbers of the second kind [9] (the Stirling number $S(r_k, j_k)$ is the number of partitions of j_k elements obtained from a set with r_k elements).

3. ESTIMATION ALGORITHMS

This section addresses the problem of estimating the parameter vector $\boldsymbol{\theta}$ defined in (1) from the observations of n independent sequences $\boldsymbol{N}^{[i]}, i = 1, ..., n$ distributed according to (3). This problem is challenging since there is no tractable expression for the intensity pdf $p(\boldsymbol{\lambda})$ and for the masses $\Pr(\boldsymbol{N} = \boldsymbol{k})$ in the general case.

3.1. Estimation in the central case

When $\boldsymbol{\mu}(\boldsymbol{\lambda}) = 0$, it has been demonstrated in [8] that the distribution of $\boldsymbol{\lambda}$ belongs to the multivariate Gamma distribution family. Consequently, the Poisson-Mandel transform of $\boldsymbol{\lambda}$ belongs to the negative multinomial family (the reader is invited to consult [8],[10] for more details). The general expression of the multinomial distribution associated to the intensities (2) is unmanageable for realistic problems when $M > 2$. However, an expression of the bivariate distribution $\Pr(N_j = k_j, N_l = k_l)$ can be derived. This bivariate distribution was used to estimate the parameter vector $\boldsymbol{\theta}$ for longitudinal count data [11] and for active images [12]. The proposed strategy consisted of maximizing an appropriate composite likelihood function $l(\boldsymbol{\theta})$ summarizing all information regarding the pairs (N_j, N_l):

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{1 \le j < l \le M} \log\left(\Pr(N_j^{[i]} = k_j^i, N_l^{[i]} = k_l^i)\right).$$

As long as the pairwise likelihood depends on $\boldsymbol{\theta}$, this algorithm yields a consistent estimator denoted as $\hat{\boldsymbol{\theta}}_n^{CL}$ under appropriate regularity conditions.

3.2. Estimation in the noncentral case

Unfortunately, the problem is much more complicated when $\boldsymbol{\mu}(\boldsymbol{\theta}) \neq 0$, since the likelihood of the pair (N_j, N_l) is not manageable for practical problems (see [13] or [4] for an expression of this likelihood). Instead, this paper proposes to estimate the parameters by using moment methods.

3.2.1. Algorithm framework

Estimating the unknown parameters of a given time series by matching its estimated and theoretical moments is a classical strategy. Of course, this approach requires to have an explicit expression of the theoretical moments as a function of the unknown parameters. In our case, this expression can be obtained by combining (4) with the general expression of the joint moments of $\boldsymbol{\lambda}$ derived in [5].
The last part of this section recalls some important results on parameter estimation by using moment methods. Consider a function $\boldsymbol{h}(.) : \mathbb{R}^M \to \mathbb{R}^L$ and the size L statistic defined as:

$$\boldsymbol{s}_n = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{h}(\boldsymbol{N}^{[i]}), \quad (5)$$

and denote as:

$$\mathsf{E}[\boldsymbol{s}_n] = \boldsymbol{f}(\boldsymbol{\theta}) = \mathsf{E}[\boldsymbol{h}(\boldsymbol{N}^{[1]})], \quad (6)$$

$$n\mathsf{cov}[\boldsymbol{s}_n] = \boldsymbol{C}(\boldsymbol{\theta}) = \mathsf{cov}[\boldsymbol{h}(\boldsymbol{N}^{[1]})]. \quad (7)$$

Note that an appropriate choice of $\boldsymbol{h}(.)$ leads to statistics $\boldsymbol{s}_n$ composed of empirical moments. In particular, we will focus in the sequel on the two estimators:

$$\hat{\boldsymbol{\theta}}_n^1 = \boldsymbol{g}(\boldsymbol{s}_n) \text{ where } \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{\theta})) = \boldsymbol{\theta}, \quad (8)$$

$$\hat{\boldsymbol{\theta}}_n^2 = \arg\min_{x} \frac{1}{2} (\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{s}_n)^t \boldsymbol{C}(\boldsymbol{x})^{-1} (\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{s}_n). \quad (9)$$

As mentioned above, the expressions of $\boldsymbol{f}(\boldsymbol{\theta})$ and $\boldsymbol{C}(\boldsymbol{\theta})$ are obtained from (6,7) by combining (4) with [5].
The asymptotic performance of estimators $\hat{\boldsymbol{\theta}}_n^1$ and $\hat{\boldsymbol{\theta}}_n^2$ can be derived by imitating the results of [14] derived in the context of time series analysis ($n = 1$ and $M \to \infty$). A key point of these proofs is the assumption $\boldsymbol{s}_n \overset{a.s.}{\to} \boldsymbol{f}(\boldsymbol{\theta})$ which is verified herein by applying the strong law of large numbers to (5). As a result, the asymptotic mean square error of $\hat{\boldsymbol{\theta}}_n^1$ and $\hat{\boldsymbol{\theta}}_n^2$ can be derived:

$$\lim_{n \to \infty} n\mathsf{E}[(\hat{\boldsymbol{\theta}}_n^1 - \boldsymbol{\theta})^2] = \boldsymbol{G}(\boldsymbol{\theta}) \boldsymbol{C}(\boldsymbol{\theta}) \boldsymbol{G}(\boldsymbol{\theta})^t, \quad (10)$$

$$\lim_{n \to \infty} n\mathsf{E}[(\hat{\boldsymbol{\theta}}_n^2 - \boldsymbol{\theta})^2] = \boldsymbol{B}(\boldsymbol{\theta}) = (\boldsymbol{F}(\boldsymbol{\theta}) \boldsymbol{C}(\boldsymbol{\theta})^{-1} \boldsymbol{F}(\boldsymbol{\theta})^t)^{-1}, \quad (11)$$

where $\boldsymbol{G}(\boldsymbol{\theta})$ and $\boldsymbol{F}(\boldsymbol{\theta})$ are the Jacobian matrices of the vectors $\boldsymbol{g}(\boldsymbol{\theta})$ and $\boldsymbol{f}(\boldsymbol{\theta})$. Another important theorem in the framework of moment methods is:

$$\boldsymbol{G}(\boldsymbol{\theta}) \boldsymbol{C}(\boldsymbol{\theta}) \boldsymbol{G}(\boldsymbol{\theta})^t \geq \boldsymbol{B}(\boldsymbol{\theta}), \quad (12)$$

where $\geq$ means that the difference between the two matrices is positive definite. Since $\boldsymbol{B}(\boldsymbol{\theta})$ uses only the statistical properties of $\boldsymbol{s}_N$, $\boldsymbol{B}(\theta)$ provides a lower bound on the asymptotic variances of all estimators constructed from functions $\boldsymbol{f}(.)$ as described above.

3.2.2. Second order moment method

The following function $\boldsymbol{h}$ will be considered in what follows the:

$$\boldsymbol{h}(\boldsymbol{x}) = (x_1, \ldots, x_M, x_1^2, \ldots, x_M^2, x_1x_2, \ldots, x_{M-1}x_M), \quad (13)$$

i.e. $\hat{\boldsymbol{\theta}}_n^1$ and $\hat{\boldsymbol{\theta}}_n^2$ rely only on first and second order moments of the observations. Extension to correlations with lags higher than one is straightforward and will be discussed in section 4.

The estimation procedure described in the previous section requires to determine the first and second-order moments of the sample statistic (6, 7). The mean $\boldsymbol{f}(\boldsymbol{\theta})$ depends on:

$$\mathsf{E}[N_k] = \mathsf{E}[\lambda_k], \quad (14)$$
$$\mathsf{E}[N_kN_q] = \mathsf{E}[\lambda_k\lambda_q] + \delta(k-q)\mathsf{E}[\lambda_k]. \quad (15)$$

The variance $\boldsymbol{C}(\boldsymbol{\theta})$ depends on $\mathsf{cov}[N_k, N_l]$ and on:

$$\mathsf{cov}(N_r, N_sN_t) = \mathsf{E}[N_rN_sN_t] - \mathsf{E}[N_r]\mathsf{E}[N_sN_t],$$
$$\mathsf{cov}(N_rN_s, N_tN_u) = \mathsf{E}[N_rN_sN_tN_u] - \mathsf{E}[N_rN_s]\mathsf{E}[N_tN_u].$$

When the indexes r, s, t and u are different, the following results are obtained:

$$\mathsf{E}[N_rN_sN_t] = \mathsf{E}[\lambda_r\lambda_s\lambda_t],\ \mathsf{E}[N_r^2N_s] = \mathsf{E}[\lambda_r^2\lambda_s] + \mathsf{E}[\lambda_r\lambda_s],$$
$$\mathsf{E}[N_r^3] = \mathsf{E}[\lambda_r^3] + 3\mathsf{E}[\lambda_r^2] + \mathsf{E}[\lambda_r],$$
$$\mathsf{E}[N_rN_sN_tN_u] = \mathsf{E}[\lambda_r\lambda_s\lambda_t\lambda_u],$$
$$\mathsf{E}[N_r^2N_tN_u] = \mathsf{E}[\lambda_r\lambda_t\lambda_u] + \mathsf{E}[\lambda_r^2\lambda_t\lambda_u],$$
$$\mathsf{E}[N_r^3N_u] = \mathsf{E}[\lambda_r\lambda_u] + 3\mathsf{E}[\lambda_r^2\lambda_u] + \mathsf{E}[\lambda_r^3\lambda_u],$$
$$\mathsf{E}[N_r^2N_u^2] = \mathsf{E}[\lambda_r\lambda_u] + \mathsf{E}[\lambda_r^2\lambda_u] + \mathsf{E}[\lambda_r\lambda_u^2] + \mathsf{E}[\lambda_r^2\lambda_u^2],$$
$$\mathsf{E}[N_r^4] = \mathsf{E}[\lambda_r^4] + 7\mathsf{E}[\lambda_r^2] + 6\mathsf{E}[\lambda_r^3] + \mathsf{E}[\lambda_r^4].$$

The next step consists of expressing intensity moments versus the mean $\boldsymbol{\mu}(\boldsymbol{\theta})$ and the covariance $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ of the Gaussian variables ψ_k. This can be done by using the general result on moments of noncentral Wishart matrices given in [5]. An alternative is to replace λ_k by (2) in $\mathsf{E}[\prod_r \lambda_r]$ and to expand the products. The higher order moments of ψ_k are then computed by using the classic expansion of a moment in terms of its cumulants. This expansion is simple here since all cumulants of ψ_k of order greater than 2 are equal to zero.

4. SIMULATION RESULTS

The simulations conducted in this paper have been obtained with the following parameters:

$$L = 1,\ \Sigma(\boldsymbol{\theta})_{k,l} = \sigma^2\rho^{|k-l|}, \mu(\boldsymbol{\theta})_k = \mu, \quad (16)$$

and $\boldsymbol{\theta} = (\mu, \rho, \sigma^2)$. The mean $\mathsf{E}[N_k]$ and the second order moments $\mathsf{E}[N_k^2]$, $\mathsf{E}[N_kN_{k+1}]$ for this model are:

$$1 \le k \le M:\ f(\boldsymbol{\theta})_k = \sigma^2 + \mu^2,$$
$$M < k \le 2M:\ f(\boldsymbol{\theta})_k = \mu^4 + 2\sigma^4 + 4\sigma^2\mu^2 + \sigma^2 + \mu^2,$$
$$2M < k \le 3M-1:\ f(\boldsymbol{\theta})_k = (\mu^2+\sigma^2)^2 + \rho^2\sigma^4 + 2\rho\sigma^2\mu^2.$$

The function $\boldsymbol{g}(.) : \mathbb{R}^{3M-1} \to \mathbb{R}^3$ required to compute $\hat{\boldsymbol{\theta}}_n^1$ is obtained after expressing μ, σ^2 and ρ as functions of $x_k = f(\boldsymbol{\theta})_k$:

$$g(x)_1 = \sqrt{\bar{x} + 2\bar{x}^2 - \bar{y}},\ g(x)_2 = \bar{x} - g(x)_1, \quad (17)$$
$$g(x)_3 = \frac{-g(x)_1 + \sqrt{g(x)_1^2 + \bar{z} - \bar{x}^2}}{g(x)_2}. \quad (18)$$

with

$$\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_i,\ \bar{y} = \frac{1}{M}\sum_{i=M+1}^{2M} x_i,\ \bar{z} = \frac{1}{M-1}\sum_{i=2M+1}^{3M-1} x_i.$$

4.1. Central case

The active imaging model corresponds to $\boldsymbol{\mu}(\boldsymbol{\theta}) = 0$ or equivalently to the unknown parameter vector $\boldsymbol{\theta} = (\rho, \sigma^2)$. Figure 1 compares the performances of the composite likelihood (CL) estimator with the two moment estimators defined in (8) and (9) (referred to as Moment and NLLS estimators, respectively). The parameters used in these simulations are $\rho = 0.8$ and $\sigma^2 = 2$ and the number of Monte Carlo runs is 500. Figure 1 shows that the empirical MSEs of $\widehat{\rho}$ (computed from the 500 Monte Carlo runs) are in good agreement with the theoretical ones given in (10), (11) and derived in [12] for the CL estimator. Figure 1 also shows that the NLLS estimator for parameter ρ is outperformed by the CL estimator. The results obtained for parameter σ^2 are not presented her for brevity. However, it turns out that the MSEs of $\widehat{\sigma^2}$ are very similar for the NLLS and CL methods.

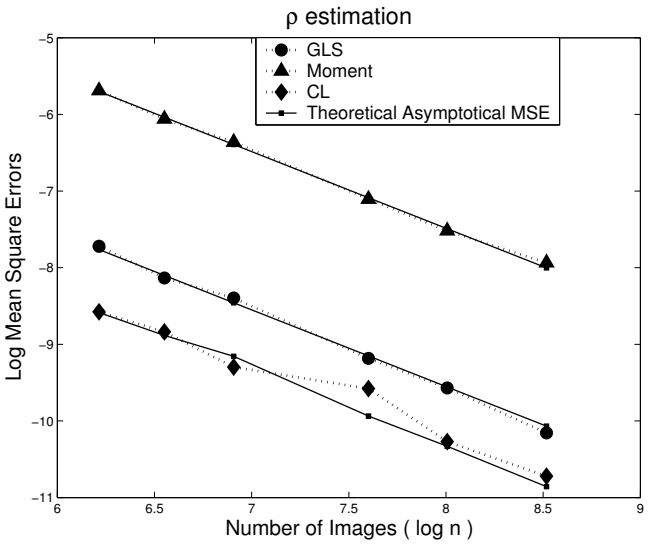

Fig. 1. Log MSE of ρ (500 Monte Carlo runs)

4.2. Noncentral case

The model used for exoplanet imaging requires $\boldsymbol{\mu}(\boldsymbol{\theta}) \neq 0$. As explained above, the CL estimator cannot be derived in this context. This section compares the performance of the NLLS estimator with that obtained with the Moment estimator.

Figure 2 shows the corresponding empirical MSEs computed with 500 Monte Carlo runs (the parameters are $\mu = 2$, $\sigma^2 = 2$ and $\rho = 0.8$). The theoretical asymptotic MSEs for both methods (provided in (11) and (10)) are also displayed. The empirical MSEs are clearly close to their asymptotic theoretical values. Moreover, it is possible to appreciate the better performance of the NLLS method. The performance obtained for parameters μ and σ^2 has not been presented here for brevity. However, the MSEs for these parameters are very close for both estimation methods.

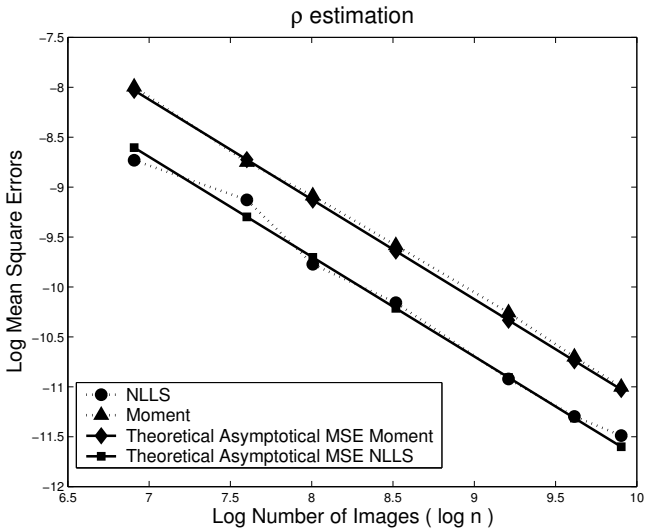

Fig. 2. Log MSE of ρ (500 Monte Carlo runs)

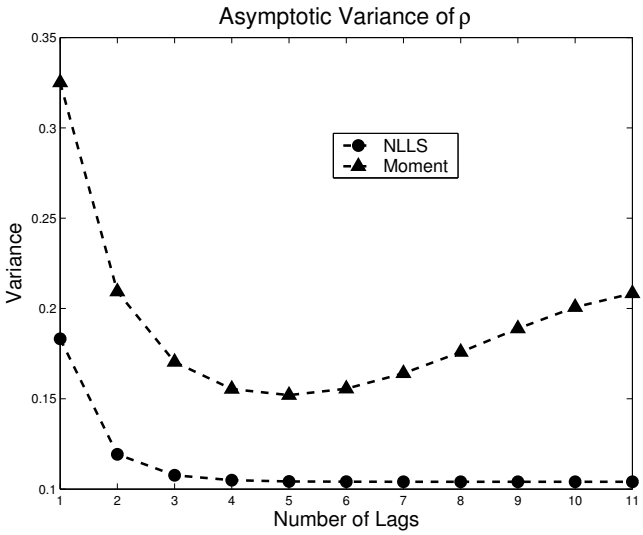

Fig. 3. Asymptotic MSE of ρ versus the number of moments (NLLS estimator)

Since the empirical and asymptotic MSEs for the Moment and NLLS estimators match in the previous experiment (i.e. for these values of n), the theoretical expressions of the unknown parameter MSEs can be used to determine the number of moments required to assess a given performance. In particular, second order moments involving lags > 1 (in addition to moments corresponding to lag 1) might improve the estimation performance. The asymptotic MSEs for parameter ρ are depicted on figure 3 as a function of the number of lags (for instance, when the number of lags is 3, the NLLS estimator considers the following moments $\mathsf{E}[N_k]$, $\mathsf{E}[N_k^2]$, $\mathsf{E}[N_k N_{k+1}]$, $E[N_k N_{k+2}]$ and $\mathsf{E}[N_k N_{k+3}]$) for a fixed value of M. This figure shows that the asymptotic MSE for parameter ρ decreases when the number of moments increases (as expected). However, the NLLS estimator complexity is an increasing function of the number of moments. The usual tradeoff between efficiency and computational cost might be used to select the appropriate number of moments used in the estimation. The behavior of the moment estimator versus the number of lags differs significantly from the NLLS estimator. Indeed, the MSEs for parameter ρ are not a decreasing function of the number of lags. Figure 3 shows that there is an optimal value of the number of lags (equal to 5) yielding a minimum MSE.

5. CONCLUSIONS

The parameters of multivariate mixed Poisson distributions can be estimated by the classical method of moments, by minimizing an appropriate non-linear least squares criterion or by maximizing a composite likelihood function. This paper compared the performance of these estimators for active images and astronomical images. The non-linear least squares estimator showed good properties for both classes of images. The application to real images is currently under investigation.

6. REFERENCES

[1] F. Goudail, N. Roux, and P. Réfrégier, "Performance parameters for detection in low-flux coherent images," *Optical Letters*, vol. 28, no. 2, pp. 81–83, 2003.

[2] C. Aime and R. Soummer, "Influence of speckle and Poisson noise on exoplanet detection with a coronograph," in *EUSIPCO-04* (L. Torres, E. Masgrau, and M. A. Lagunas, eds.), (Vienna, Austria), pp. 509–512, Elsevier, Sept. 2004.

[3] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2. New York: John Wiley, 2nd ed., 1995.

[4] R. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley and Sons, 1982.

[5] J.-Y. Tourneret, 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), July 2005.

[6] J. Goodman, *Statistical Optics*. New York: Wiley, 1985.

[7] D. Karlis and E. Xekalaki, "Mixed Poisson distributions," *International Statistical Review*, vol. 73, no. 1, pp. 35–58, 2005.

[8] A. Ferrari, G. Letac, and J.-Y. Tourneret, "Multivariate mixed Poisson distributions," in *EUSIPCO-04* (F. Hlawatsch, G. Matz, M. Rupp, and B. Wistawel, eds.), (Vienna, Austria), pp. 1067–1070, Elsevier, Sept. 2004.

[9] N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate discrete distributions*. New York: John Wiley, 2nd ed., 1992.

[10] P. Bernardoff, "Which negative multinomial distributions are infinitely divisible?," *Bernoulli*, vol. 9, no. 6, 2003.

[11] R. Henderson and S. Shimakura, "A serially correlated gamma frailty model for longitudinal count data," *Biometrika*, vol. 90, no. 2, pp. 355–366, 2003.

[12] 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.

[13] G. Letac and H. Massam, "A tutorial on non central Wishart distributions." `http://www.lsp.ups-tlse.fr/Fp/Letac/Wishartnoncentrales.pdf`.

[14] B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments," *International Journal of adaptive control and signal processing*, vol. 3, pp. 191–229, 1989.