DIRECTION-OF-ARRIVAL AND FREQUENCY ESTIMATION USING POISSON-GAUSSIAN MODELING

Frédéric Dublanchet[1,2], *Jérôme Idier*[1] *and Patrick Duwaut*[2]

[1]Laboratoire des Signaux et Systèmes (CNRS–Supélec–UPS), Plateau de Moulon,
91192 Gif-sur-Yvette cedex, France
[2]Équipe du Traitement des Images et du Signal, 6 av. du Ponceau,
95014 Cergy-Pontoise cedex, France
email : dublanchet@lss.supelec.fr, idier@lss.supelec.fr

ABSTRACT

We address the problem of identification of sinusoidal components from observed data, which is fundamental for array signal processing and spectral line decomposition. Joint detection and estimation are proposed in a unified Bayesian framework, so that no preliminary estimate of the number of signals is required. All unknown quantities are estimated from a unique regularized "stochastic" likelihood function, including the number of sources and statistical parameters. The impulsive solution is modeled as a continuous Poisson-Gaussian process. A powerful iterative technique is proposed to maximize the posterior likelihood. Simulation results show that the method behaves particularly well for small data sets, even for a single experiment.

1. INTRODUCTION

Evaluation of directions of arrival (DOAs) for array signal processing and spectral line decomposition share a common problem of identifying sinusoidal components. Considering the DOA estimation problem, under the basic model of p narrow-band planes waves impinging on a N-sensors uniform linear array from a set $\boldsymbol{\theta}$ of p distincts DOAs, the K noisy observed data are usually modeled as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) + \mathbf{n}(t) \ , \ t = 1, \ldots, K \qquad (1)$$

where columns of matrix $\mathbf{A}(\boldsymbol{\theta})$ are the so-called "steering vectors". In the sequel, the additive noise $\mathbf{n}(t)$ is supposed to be a white stationary, complex-valued circular Gaussian process with zero mean and variance σ_{n}^2, independent from the signal amplitude vector $\mathbf{x}(t)$ and independent from snapshot to snapshot. The problem is commonly presented as a three-fold issue:

- A *detection* step, *i.e.,* finding the number p;
- A *localization* step, *i.e.,* finding the p DOAs in $\boldsymbol{\theta}$;
- An *estimation* step, *i.e.,* estimating the amplitude $\mathbf{x}$ of each component.

As regards localization and estimation, the two last decades have witnessed the impressive development of "high resolution" methods, either based upon eigendecomposition of the data covariance matrix (such as ESPRIT, MUSIC), or upon maximum likelihood estimation, such as the well-known Stochastic Maximum Likelihood (SML) method [1]. As regards detection methods, the most commonly used are Akaike's information criterion (AIC) and Rissanen's minimum description length (MDL). Both globally suffer from a lack of robustness, especially in the realistic conditions of a finite and possibly small number of samples.

From a methodological point of view, one may wonder why detection and localization have always been coped as two separate issues. From intuitive considerations, trying to evaluate the number of components before locating them may rather seem an inefficient division. For sake of overall efficiency, we propose a challenging approach, designed as a unique and coherent localization procedure.

Our inspiration originates from the field of seismic deconvolution, which basically amounts to estimating sparse spike trains from scanned echoes. This issue is not very different from estimating DOAs from noisy data. The seismic wavelet echoed at a given depth plays a comparable role to the steering vector in the direction of a given source. Yet, both problems are addressed using rather different terminology and methodology. In particular, finding the number of separate echoes is usually not splitted from localization in seismic data processing. On the other hand, the ill-posed character of seismic deconvolution is acknowledged and most modern deconvolution techniques resort to regularization tools. The spiky nature of the unknown signal is the basic prior information which is taken into account. More specifically, Mendel and coworkers introduced a Bayesian approach based on a discrete Bernoulli-Gaussian (BG) prior model [2].

Efficient algorithmic strategies have been developped for BG deconvolution [3]. We have recently adapted the BG methodology to the analysis of superimposed complex sinusoids embedded in additive Gaussian noise [4]. However, it is intrinsically limited by the discrete nature of the Bernoulli part. In order to overcome this restriction, we propose an original extension to a continuous Poisson-Gaussian (PG) prior model. To our knowledge, the work of Kwakernaak is the unique precursor in the field of PG deconvolution [5].

This work was supported by the French organism Direction de la Recherche et de la Technologie (DRET).

2. POISSON-GAUSSIAN PRIOR MODEL

Prior models for all unknown quantities are chosen in accordance with structural knowledge and qualitative information, *i.e.,* the expected impulsive structure of the solution. In this respect, a Poisson process with uniform intensity λ seems a natural choice for any set $\boldsymbol{\theta}$ of p DOAs. Accordingly, the total number $N(\Delta)$ of events occuring during the interval Δ obeys the Poisson probability law: $\Pr(N(\Delta)=p) = \lambda^p \delta^p e^{-\lambda\,\delta}/p!$, where δ is the length of Δ. Now let us consider the problem of defining a proper prior likelihood function for any configuration $\boldsymbol{\theta}$ of DOAs. Since we cannot discriminate among the DOAs, $\boldsymbol{\theta}$ must be considered as a set rather than a vector. As a consequence, we obtain the following expression for the probability density of $\boldsymbol{\theta}$, conditionally to the presence of p sources:

$$f(\boldsymbol{\theta}\,|\,N(\Delta)=p) \;\begin{array}{l}=\\=\end{array}\; \begin{array}{ll} p!/\delta^p & \text{if } p=\dim\boldsymbol{\theta} \\ 0 & \text{otherwise.}\end{array} \qquad (2)$$

In a fairly natural way, we define the a priori likelihood of any angular sequence $\boldsymbol{\theta}$ as the product $L(\boldsymbol{\theta}, N(\Delta)=p) \triangleq f(\boldsymbol{\theta}\,|\,N(\Delta)=p)\ \Pr(N(\Delta)=p)$. Provided we intrinsically assume that p is the size of $\boldsymbol{\theta}$, the notation can remain implicit w.r.t. p and we get:

$$L(\boldsymbol{\theta}) = \lambda^p e^{-\lambda\,\delta}. \qquad (3)$$

Conditionally to $\boldsymbol{\theta}$, amplitudes vector $\mathbf{x}$ is assumed an independent complex circular Gaussian vector with zero mean and variance σ^2. In other words, we choose a very simple structure for the covariance matrix of source amplitudes:

$$\mathbf{S} \triangleq \mathrm{E}\left[\mathbf{x}(t)\mathbf{x}^\dagger(t)\,|\,\boldsymbol{\theta}\right] = \sigma^2\mathbf{I}. \qquad (4)$$

This is a slight difference with SML, since we prefer a simple white assumption for the amplitudes rather than estimating the covariance matrix based on averaging. Because one of our main concerns is to devise an efficient method even for small data sets (especially for the case of one snapshot only, as usually encountered in spectral analysis problem), we substituted the structured form (4) for the usual approach.

3. ESTIMATION STRATEGY

Let $\mathcal{H} = (\lambda, \sigma^2, \sigma_{\mathrm{n}}^2)$ gather the three hyperparameters. In order to provide a fully unsupervised method, estimation of $\mathcal{H}$ addressed in Section 5. For sake of clarity, we first consider that $\mathcal{H}$ is known or previously estimated, and the dependence of the likelihood functions w.r.t. $\mathcal{H}$ remains implicit.

In [5], Kwakernaak proposed to maximize the *joint* posterior likelihood function $L(\boldsymbol{\theta}, \mathbf{x}\,;\,\mathbf{y})$ w.r.t. both $\mathbf{x}$ and $\boldsymbol{\theta}$, but he also reported that such maximum a posteriori (MAP) estimator leads to an overestimated number of detected pulses. Indeed, it is clear from later contributions related to BG deconvolution [2, 3] that a sequential estimation scheme is far more reliable. Such an approach relies on the *marginal* MAP (MMAP) solution $\widehat{\boldsymbol{\theta}}$ as the maximizer of the marginal posterior likelihood:

$$L(\boldsymbol{\theta}\,;\,\mathbf{y}) = \int_{\mathbb{C}^p} L(\boldsymbol{\theta}, \mathbf{x}\,;\,\mathbf{y})\,d\mathbf{x} = p(\mathbf{y}\,|\,\boldsymbol{\theta})L(\boldsymbol{\theta})/p(\mathbf{y}), \qquad (5)$$

$$\text{where} \quad p(\mathbf{y}\,|\,\boldsymbol{\theta}) = \frac{1}{\pi^{NK}|\mathbf{B}|^K}\exp -\sum_{k=1}^{K}\mathbf{y}_k^\dagger\mathbf{B}^{-1}\mathbf{y}_k \qquad (6)$$

with the covariance matrix $\mathbf{B}$:

$$\mathbf{B}(\boldsymbol{\theta}) \triangleq \mathrm{E}\left[\mathbf{y}(t)\mathbf{y}^\dagger(t)\,|\,\boldsymbol{\theta}\right] = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\mathbf{A}^\dagger(\boldsymbol{\theta}) + \sigma_{\mathrm{n}}^2\mathbf{I}. \qquad (7)$$

At this point, remember that SML maximizes (6) [1], and no prior term is assumed w.r.t. $\boldsymbol{\theta}$. On the other hand, the Poisson prior (3) only depends on the size of $\boldsymbol{\theta}$, which is a fixed quantity in the framework of SML. In other words, the Poisson prior likelihood introduces no distortion in the localization step compared to SML. According to the MMAP approach, the localization problem boils down to the minimization of the regularized criterion:

$$\mathcal{C}(\boldsymbol{\theta}) = \sum_{k=1}^{K}\mathbf{y}_k^\dagger\mathbf{B}^{-1}\mathbf{y}_k + K\ln|\mathbf{B}| - p\ln\lambda. \qquad (8)$$

We prove in [6] that the criterion (8) has a global minimum $\widehat{\boldsymbol{\theta}} \triangleq \arg\min_{\boldsymbol{\theta}}\mathcal{C}(\boldsymbol{\theta})$ if and only if $\lambda \leq 1$.

As recommended in the field of BG deconvolution, the localization step is followed by MAP estimation of the associated complex amplitudes through maximization of $L(\widehat{\boldsymbol{\theta}}, \mathbf{x}\,|\,\mathbf{y})$ w.r.t. $\mathbf{x}$. Since the localization step provides an estimated dimension $\widehat{p} = \dim\widehat{\boldsymbol{\theta}}$, the estimation step reduces to a simple linear Gaussian problem. It is straightforward to derive the following optimal estimate of source amplitudes, for each snapshot: $\forall t\ :\ \widehat{\mathbf{x}}(t) = \sigma^2\mathbf{A}^\dagger(\widehat{\boldsymbol{\theta}})\,\mathbf{B}^{-1}(\widehat{\boldsymbol{\theta}})\,\mathbf{y}(t)$.

4. OPTIMIZATION STRATEGY

According to a global detection-localization approach, criterion (8) has to be minimized w.r.t the components of $\boldsymbol{\theta}$ but also to its dimension p. For, a powerful technique is proposed which performs alternate maximization over discrete and continuous sets, respectively based on Single Most Likely Replacement (SMLR) [3] and on gradient descent.

(i) The SMLR step enables to perform "jumps" between criteria in a numerically efficient way, through basic variation in the current set of detected sources. Its objective is not only to select a criterion, *i.e.,* to fix p, but also to provide relatively accurate DOA estimates on a grid.

(ii) The second step performs local minimization out of the grid in the neighborhood of the discrete solution. The dimension p of $\boldsymbol{\theta}$ is now held fixed. The role of this second step is to let the p angles drift away from their initial positions, provided it ensures a further decrease of the criterion $\mathcal{C}(\boldsymbol{\theta})$.

4.1. Minimization over a discrete grid

Let us consider the general form $\mathcal{D} \triangleq (c_1, \ldots, c_\ell, \ldots, c_L)$ for a discrete grid composed of L points on domain Δ. The parameter L may take arbitrary large values, greater than the number of sensors, in order to allow high resolution analysis. Without prior knowledge about the location of the DOAs, the discrete angles will be chosen equally spaced over domain Δ. Assuming that we restrain the search of the minimizer of $\mathcal{C}(\boldsymbol{\theta})$ to the set $\mathcal{P}_{\mathcal{D}}$ of subsets of $\mathcal{D}$, the numerical problem arises in terms of combinatorial exploration, very

close to MMAP BG deconvolution. Exact optimization of $\mathcal{C}(\boldsymbol{\theta})$ over $\mathcal{P}_{\mathcal{D}}$ would require 2^L evaluations of the criterion, which is computationally intractable for large values of L. Instead, SMLR only performs partial combinatorial exploration in an iterative way. Each iteration only explores a set of *neighboring* sequences of the current sequence. The algorithm is stopped when no neighboring sequence is more likely than the current one, which is chosen as the final one.

Let $\boldsymbol{\theta}_0$ denote the current sequence, which is supposed to belong to $\mathcal{P}_{\mathcal{D}}$. We define the neighborhood of $\boldsymbol{\theta}_0$ as the set of L sequences $\boldsymbol{\theta}_\ell$, $\ell = 1, \ldots, L$, that differ from $\boldsymbol{\theta}_0$ by one entry, *i.e.,* whether $\boldsymbol{\theta}_\ell = \boldsymbol{\theta}_0 \cup \{c_\ell\}$ or $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_\ell \setminus \{c_\ell\}$. For sake of readability, subscript ℓ (resp. 0) will refer to any quantity related to $\boldsymbol{\theta}_\ell$ (resp. $\boldsymbol{\theta}_0$). Let also $\mathbf{a}_\ell$ stand for the corresponding steering vector $\mathbf{a}(c_\ell)$, and the scalar ε_ℓ take the value 1 (resp. -1) when c_ℓ is added to (resp. removed from) $\boldsymbol{\theta}_0$. As a straightforward variation of [3], it can be shown that the criterion is iteratively computable using:

$$\mathcal{C}_\ell = \mathcal{C}_0 - \rho_\ell^{-1} \sum_{k=1}^{K} \left| \mathbf{y}_k^\dagger \mathbf{B}_0^{-1} \mathbf{a}_\ell \right|^2 + K \ln \left(\varepsilon_\ell \sigma^2 \rho_\ell \right) - \varepsilon_\ell \ln \lambda,$$

where

$$\begin{aligned} \mathbf{y}_k^\dagger \mathbf{B}_0^{-1} \mathbf{a}_\ell &= \sigma_{\mathrm{n}}^{-2} \mathbf{y}_k^\dagger \mathbf{a}_\ell - \sigma_{\mathrm{n}}^{-2} \sigma^2 \mathbf{y}_k^\dagger \mathbf{A}_0 \mathbf{C}_0^{-1} \mathbf{A}_0^\dagger \mathbf{a}_\ell, \\ \rho_\ell &= \varepsilon_\ell \sigma^{-2} + \sigma_{\mathrm{n}}^{-2} - \sigma_{\mathrm{n}}^{-2} \sigma^2 \mathbf{a}_\ell^\dagger \mathbf{A}_0 \mathbf{C}_0^{-1} \mathbf{A}_0^\dagger \mathbf{a}_\ell, \end{aligned}$$

are expressed in terms of the inverse of $p \times p$ matrix $\mathbf{C} \triangleq \sigma^2 \mathbf{A}^\dagger \mathbf{A} + \sigma_{\mathrm{n}}^2 \mathbf{I}$ at $\boldsymbol{\theta}_0$. The update of matrix $\mathbf{C}_0^{-1}$ is only needed once the optimal neighbor sequence is determined. All quantities $\mathbf{y}_k^\dagger \mathbf{a}_\ell$ and $\mathbf{y}_k^\dagger \mathbf{A}_0$ enter the $L \times K$ matrix $\mathbf{y}^\dagger \mathbf{A}(\mathcal{D})$, which can be computed at the beginning of the procedure by discrete Fourier transform of the snapshots. In the same way, $\mathbf{A}_0^\dagger \mathbf{a}_\ell$ enter the $L \times L$ matrix $\mathbf{A}(\mathcal{D})^\dagger \mathbf{A}(\mathcal{D})$, which has an explicit form easily derived.

The initial sequence is chosen as the empty set, which is the most favourable choice from the numerical viewpoint since $\mathbf{C}(\emptyset) = \sigma_{\mathrm{n}}^2 \mathbf{I}$. Moreover, it avoids the search of "a good initial solution", in contrast with many existing methods. Finally, each iteration involves $\mathcal{O}(KLp^2)$ complex operations. Convergence is reached in a finite (practically small) number of iterations since $\boldsymbol{\theta}$ spans a finite set.

4.2. Local Minimization

Local minimization is based upon a gradient-type algorithm

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} - \mu^{(m)} \left. \frac{\partial \mathcal{C}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta} = \boldsymbol{\theta}^{(m)}}$$

where superscript (m) refers the mth iteration and $\mu^{(m)}$ stands for a steplength. It is initialized by the outcome of the SMLR procedure.

In order to maintain low computational requirements, we replace the original expression (8) of $\mathcal{C}(\boldsymbol{\theta})$ by an equivalent expression involving $p \times p$ matrix $\mathbf{C}(\boldsymbol{\theta})$ rather than $N \times N$ matrix $\mathbf{B}(\boldsymbol{\theta})$. In addition, since the prior term does not depend on $\boldsymbol{\theta}$, the gradient may be expressed as:

$$\begin{aligned} \frac{\partial \mathcal{C}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= 2K \Re \Big\{ \mathrm{diag} \Big[\sigma^2 \mathbf{D}^\dagger \mathbf{A} \mathbf{C}^{-1} \\ &\quad - \sigma_{\mathrm{n}}^{-2} \mathbf{C}^{-1} \mathbf{A}^\dagger \widehat{\mathbf{R}} \left(\mathbf{I} - \sigma^2 \mathbf{A} \mathbf{C}^{-1} \mathbf{A}^\dagger \right) \mathbf{D} \Big] \Big\} \end{aligned}$$

where $\widehat{\mathbf{R}} \triangleq \mathbf{y}\mathbf{y}^\dagger / K$ is the empirical estimate of the sample covariance matrix. The evaluation of gradient essentially involves the calculus of inverse matrix $\mathbf{C}^{-1}$ and some additional matrix products, so that each iteration requires $\mathcal{O}(p^3 + (K+N)p^2 + KNp)$ complex multiplications.

5. ESTIMATION OF HYPERPARAMETERS

5.1. MGL approach

The problem of estimating the hyperparameters from the available data is now addressed. For this purpose, several techniques have been designed in the related context of BG deconvolution. We have selected a generalized marginal likelihood approach [3] which consists in estimating *simultaneously* $\boldsymbol{\theta}$ and the hyperparameters set $\mathcal{H}$ through maximization of the Marginal Generalized Likelihood (MGL) function $p(\mathbf{y} \,|\, \boldsymbol{\theta}\,;\, \sigma^2, \sigma_{\mathrm{n}}^2)\, L(\boldsymbol{\theta}\,;\, \lambda)$, which is the joint probability density of $(\boldsymbol{\theta}, \mathbf{y})$. As a function of $\boldsymbol{\theta}$, it is proportional to the marginal posterior likelihood $L(\boldsymbol{\theta}\,;\, \mathbf{y})$, so the MMAP estimate $\widehat{\boldsymbol{\theta}}$ is also the maximizer of L w.r.t. $\boldsymbol{\theta}$ when $\mathcal{H}$ is held constant. Accordingly, we shall extend the notation $\mathcal{C}$ introduced in (8) to designate the negative logarithmic form of MGL:

$$\mathcal{C}(\boldsymbol{\theta}\,;\, \mathcal{H}) = \sum_{k=1}^{K} \mathbf{y}_k^\dagger \mathbf{B}^{-1} \mathbf{y}_k + K \ln |\mathbf{B}| - p \ln \lambda + \lambda\, \delta. \qquad (9)$$

Minimization of $\mathcal{C}(\boldsymbol{\theta}\,;\, \mathcal{H})$ w.r.t. $\boldsymbol{\theta}$ and $\mathcal{H}$ may be suboptimally performed using the iterative two-step scheme

$$\widehat{\boldsymbol{\theta}}_i = \arg\min_{\boldsymbol{\theta}} \mathcal{C}(\boldsymbol{\theta}\,;\, \widehat{\mathcal{H}}_{i+1}) \qquad (10)$$

$$\widehat{\mathcal{H}}_{i+1} = \arg\min_{\mathcal{H}} \mathcal{C}(\widehat{\boldsymbol{\theta}}_i\,;\, \mathcal{H}) \qquad (11)$$

often referred to as a "Block component method".The first step (10) identifies with the supervised procedure described in section 4. The step (11) amounts to a 3-D optimization.

5.2. Estimation of parameter λ

Estimation of intensity λ can be achieved separately since only the prior term $L(\boldsymbol{\theta}\,;\, \lambda)$ depends on λ. Differentiating (9) w.r.t. λ results in the following MMGL estimate: $\widehat{\lambda} = \widehat{p}/\delta$. Note that λ and $\widehat{\lambda}$ are not dimensionless quantities, but rather numbers of events per angular sector. Accordingly, the numerical value of $\widehat{\lambda}$ depends on the angular unit. On the other hand, the behaviour of supervised estimation of $\boldsymbol{\theta}$ becomes pathological if λ is strictly greater than one. This is not only a potential source of degeneracy, but also a source of arbitrariness regarding the choice of angular unit that is not mentioned in [5], though it may partly explain the degeneracies reported by Kwakernaak. The intrinsic problem actually lies in the behaviour of the MMAP estimator itself, which depends on the value of a dimensioned quantity relatively to a constant. A drastic solution would be to replace the MMAP estimator by another Bayesian estimator. However, we found no substitute candidate that would provide the same trade-off between numerical complexity and performance. Instead, a more pragmatic approach is to ensure that λ is lower than one.

For this purpose, we choose the sampling rate δ/L used to generate the grid $\mathcal{D}$ as the angular unit, since the value of L must be already be chosen high enough with regard to the expected number of sources and to the number N of sensors. The result is $\widehat{\lambda} = \widehat{p}/L$, which stays lower than one in practice. Overflow can be easily prevented anyway in the SMLR step, where $\widehat{p}$ is updated.

5.3. Estimation of parameters σ^2 and $\sigma^2_{\mathbf{n}}$

We now discuss the problem of minimizing (9) w.r.t. the variances σ^2 and σ^2_{n}. As shown in [3], this may be reduced to a one-dimensional optimization problem using an appropriate change of variable, for instance $(\sigma^2, \sigma^2_{\mathrm{n}}) \rightarrow (\alpha, \sigma^2_{\mathrm{n}})$, with $\alpha \stackrel{\Delta}{=} \sigma^2/\sigma^2_{\mathrm{n}}$. The whole procedure converges in a small number of iterations (no more than ten in our practical experiments) and it is easy to implement.

6. CONCLUSION

Joint detection and localization are performed within a unified Bayesian framework, so that no preliminary estimate of the number of sources is required. Compared to preexisting methods for estimating DOAs, the PG approach leads to a penalized version of the SML method. If it were constrained to a given number of sources, it would only boil down to a simplified version of the standard SML estimator, for which the signal amplitude correlation matrix $\mathbf{S}$ is taken proportional to identity. The PG method should rather be viewed as a generalized version of SML that allows to tackle an unconstrained number of sources. Whereas usual SML would degenerate towards a meaningless maximal number of sources, it is shown that the new PG version does theoretically provide a finite set of sources as the most likely one. Finally, all unknown quantities are estimated from a unique regularized "stochastic" likelihood function, including the number of sources, their locations and statistical parameters $\mathcal{H}$. Further works are currently studied to valid extensions of the proposed estimation and optimization techniques to SML-like frameworks.

7. SIMULATION RESULTS

Compared to standard information-based approaches, simulations results show that the PG method brings substantial improvement in terms of detection performance, while it behaves as expected for sufficiently large data samples, quite similarly to SML estimation in terms of localization performance . Nevertheless, the PG method behaves particularly better for small data sets, even for a single experiment, as depicted below.

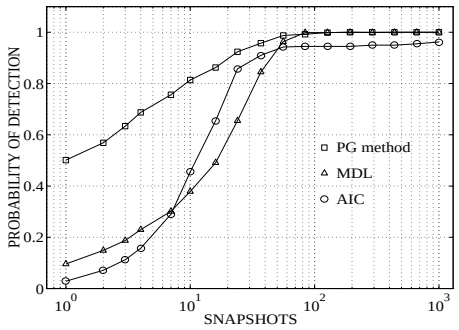

Figure 1: two sources located at angles -5 and 5 deg. impinge on a 4-sensors uniform linear array with spacing between two consecutive sensors equal to half a wavelength. Each source has SNR = 10 dB. Probability of correct detection, examined in terms of detecting two sources, is evaluated as a function of the number of snapshots. For each point, 800 independent trials were generated with initial parameters: $\lambda^{(0)} = 0.833$, $\sigma^{2(0)} = 1.4$, $\sigma^{2(0)}_{\mathrm{n}} = 0.05$ and $L = 24$.

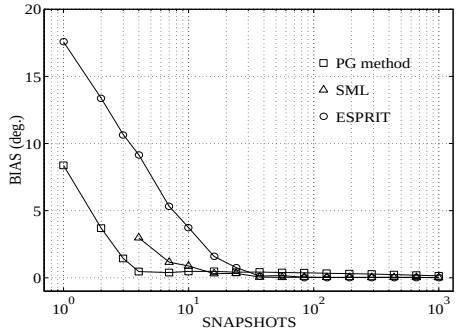

Figure 2: conditionally to a correct detection for the PG method, corresponding estimation bias for one of the two sources is evaluated; these results correspond to the situation described on Fig. 1. The first three points for the SML method are not plotted since obtained results were meaningless, which is due to a lack of data.

8. REFERENCES

[1] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields", *Signal Processing*, vol. 11, pp. 329–337, 1986.

[2] J. J. Kormylo and J. M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes", *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 482–488, 1982.

[3] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation", *IEEE Trans. Signal Processing*, vol. 44, no. 12, December 1996.

[4] F. Dublanchet, P. Duvaut, and J. Idier, *Complex sinusoid analysis by Bayesian deconvolution of the discrete Fourier transform*, pp. 323–328, Maximum entropy and Bayesian methods. Kluwer Academic Publ., Santa Fe, U.S.A., K. Hanson edition, 1995.

[5] H. Kwakernaak, "Estimation of pulse heights and arrival times", *Automatica*, vol. 16, pp. 367–377, 1980.

[6] F. Dublanchet, J. Idier, and P. Duvaut, "Direction-of-arrival and frequency estimation using Poisson-Gaussian deconvolution", *submitted to IEEE Trans. Signal Processing*, 1996.