BAYESIAN DECONVOLUTION OF POISSONIAN POINT SOURCES

†*Guillaume Stawinski* - †‡*Arnaud Doucet* - ♦*Patrick Duvaut*
†LETI CEA-Technologies Avancées CEN Saclay
DEIN/SPE Bat. 451 91191 Gif-sur-Yvette cedex, France.
‡Signal Processing Group, Departement of Engineering
University of Cambridge, Trumpington St., CB2 1PZ Cambridge, UK.
♦ETIS-ENSEA URA CNRS 2235 95014 Cergy Pontoise cedex, France.
Tel: +33 1 69.08.88.20 - Fax: +33 1 69.08.88.20
e-mail: stawinski@serin.cea.fr - ad2@eng.cam.ac.uk - duvaut@ensea.fr

ABSTRACT

In this article, we address the problem of Bayesian deconvolution of point sources with Poisson statistics. A high level Bayesian approach is proposed to solve this problem. The original image is modeled as a list of an unknown number of points sources with unknow parameters. A prior distribution reflecting our degree of belief is introduced on all these unknown parameters, including the number of sources. All Bayesian inference relies on the posterior distribution. This latter admitting no analytical expression, we estimate it using an original Reversible Jump Markov Chain Monte Carlo method. The algorithm developed is tested over real data. It displays satisfactory results compared to traditional low level Bayesian approaches.

1 INTRODUCTION

In this article, we consider the problem of deconvolution of point sources images with Poissonian statistics. This problem is motivated by an industrial application where one wants to locate radioactive sources with a portative Gamma camera [5]. The available image is the result of the convolution of the point sources by the spread impulse response of the low resolution Gamma camera. The aim is to develop an algorithm to detect and locate precisely the point sources in space. Related problems occur in astronomy when one is interested in locating stars.

A traditional statistical approach to solve this kind of problems is a *low level* Bayesian approach. The original image, i.e. the image before convolution by the impulse response of the sensor, is modeled using a Markov Random Field (MRF) prior distribution. The likelihood is defined by the assumption of Poisson statistics. Then Bayesian inference is performed to obtain the Maximum A Posteriori (MAP) or the conditional expectation using the Expectation Maximisation (EM) algorithm or Markov chain Monte Carlo (MCMC) methods [8], [3]. MRF prior models are often used because of their ability to model global properties using local constraints. However, this kind of prior distributions is not well-adapted to model non-homogeneous images such as point sources images.

We adopt here a different *high level* Bayesian approach where the original picture is modeled as a list of geometrical objects rather than as a list of pixels. This kind of approach was suggested recently by Baddeley and Van Lieshout [1] who use a marked point process model as an object prior, with the points representing the locations of objects and the marks being the variables needed to describe the objects themselves. We follow here a similar approach. More precisely, we assume that our picture is a list of an unknown number of radioactive sources, each radioactive source being assumed to have a circular Gaussian shape with unknown amplitude, variance and location. This kind of parametric models involves typically fewer parameters than MRF models and is able in our case to describe more accurately the original image. A prior distribution for all these unknown parameters including the number of sources is defined, the likelihood function being not modified. All Bayesian inference on the unknown image is then based on the posterior distribution allowing to perform joint detection and estimation of the sources parameters. However, similarly to the case where MRF prior models are used, it is impossible to evaluate analytically the posterior distribution of interest. To estimate this distribution, we propose to use a MCMC method, i.e. we build a ergodic Markov chain which admits as limiting distribution the posterior distribution of interest. The posterior distribution being defined on a union of subspaces of different dimensions, it is not possible to use classical MCMC methods. Recently, Green [4] has developed a general methodology to solve such problems: the reversible jump MCMC methods. We propose here an original MCMC sampler relying on this theory to estimate the posterior distribution.

The paper is organized as follows. Section 2 describes the Bayesian model and our estimation objectives. In section 3, the main steps of Bayesian computation are given. The performance of this algorithm is illustrated by computer simulations on real data in section 4. Finally, some conclusions are drawn in section 5.

2 STATISTICAL MODEL AND AIM

We first define the statistical model and then state our estimation objectives.

2.1 Statistical model

The image $x(i,j)$ to estimate is defined on a compact set $T \subseteq \mathbb{N}^2$. This image is convolved by the known impulse response of the sensor $h(m,n)$ and we obtain the intensity image $\lambda(i,j)$:

$$\lambda(i,j) = \sum_n \sum_m h(i-m, j-n)\, x(m,n) \qquad (1)$$

The Gamma camera introduces statistical fluctuations and we only observe

$$y(i,j) \sim \mathcal{P}(\lambda(i,j)) \qquad (2)$$

where $\mathcal{P}(\lambda(i,j))$ is the Poisson distribution of parameter $\lambda(i,j)$, i.e.

$$\Pr(y(i,j) = m) = \exp(-\lambda(i,j)) \frac{[\lambda(i,j)]^m}{m!} \qquad (3)$$

Equations (1) and (2) define the likelihood of the model.

We now define a *high level* prior model for the unobserved image $x(m,n)$. This image is modeled as a sum of an unknown number k of sources of circular Gaussian shape with unknown parameters, i.e.

$$x(m,n) = \sum_{i=1}^{k} A_{i,k} \exp\left(- \frac{\left(m - c_{i,k}^1\right)^2 + \left(n - c_{i,k}^2\right)^2}{2\sigma_{i,k}^2} \right) \qquad (4)$$

when the number of sources k sources is given, where $A_{i,k} \in \mathbb{R}^+$, $\mathbf{c}_{i,k} = \left(c_{i,k}^1, c_{i,k}^2\right) \in \mathbb{N}^2$, $\sigma_{i,k}^2 \in \mathbb{R}^+$ are respectively the amplitude, the center and the variance of the i^{th} source for the model with k sources.

We follow here a full Bayesian approach, that is the unknown parameters k and $\boldsymbol{\theta}_k \triangleq \left(A_{i,k}, \mathbf{c}_{i,k}, \sigma_{i,k}^2\right)$ are regarded as random. The space of parameters $\boldsymbol{\Theta}$ is a countable union of subspaces $\boldsymbol{\Theta} = \cup_{k=0}^{\infty} \boldsymbol{\Theta}_k$ where $\boldsymbol{\Theta}_k$ represents the space of parameters when the number of sources k is given, it is a subspace of $(\mathbb{R}^+)^{2k} \times (\mathbb{N})^{2k}$. We assume the following prior structure for $\boldsymbol{\theta}_k$:

$$p(\boldsymbol{\theta}_k | k) = p\left(\{\mathbf{c}_{i,k}\}_{i=1,\ldots,k} | k \right) \prod_{i=1}^{k} p(A_{i,k}|k)\, p(\sigma_{i,k}^2 | k) \qquad (5)$$

In the next subsections, we define more precisely these prior distributions.

Prior distribution for the centres For the centers, a simple uniformative uniform distribution on T. That is for $k \geq 1$

$$p\left(\{\mathbf{c}_{i,k}\}_{i=1,\ldots,k} | k \right) = [area(T)]^{-k} \prod_{i=1}^{k} \mathbb{I}_T(\mathbf{c}_{i,k}) \qquad (6)$$

If additional prior information is available, it is possible to select for example a discrete equivalent to the Strauss process [7]. It would allow us to introduce interaction parameters which roughly speaking control the degree of repulsion between the centres.

Prior distribution for the amplitudes The unknown positive amplitudes are assumed distributed according to a Gamma law:

$$A_{i,k} \sim \mathcal{G}(\alpha_{i,k}, \beta) \qquad (7)$$

One would like to select improper prior distribution as it is often done in Bayesian estimation. However one must be very careful as such an approach is not valid in a model selection framework. It would in all cases lead to the selection of the model with the smaller number of parameters: this is the Lindleys' paradox [2]. We consider here a hierarchical model whose aim is to provide a flexible and weakly informative prior distribution on the parameters. That is we allow the priors α and β to depend on hyperparameters. It seems natural to take the conjugate law:

$$\alpha_{i,k} \sim \mathcal{G}(\alpha_{k_1}, \beta_{k_1}) \qquad (8)$$

$$\beta \sim \mathcal{G}(\alpha_{k_2}, \beta_{k_2}) \qquad (9)$$

where α_{k_1}, β_{k_1}, α_{k_2} and β_{k_2} are selected to obtain a vague prior distribution.

Prior distribution for the variance The unknown variances $\sigma_{i,k}^2$ are assumed distributed according to an inverse Gamma distribution:

$$\sigma_{i,k}^2 \sim \mathcal{IG}(\upsilon_0, \gamma_0) \qquad (10)$$

We adopt for a vague proper prior distribution with an infinite variance.

Prior distribution for the number of sources We assume that the number of sources k is distributed uniformly in $[0, \ldots, k_{\max}]$, i.e.

$$p(k) = (k_{\max} + 1)^{-1}\, \mathbb{I}_{[0,\ldots,k_{\max}]}(k) \qquad (11)$$

2.2 Estimation objectives

Given the observed image $\mathbf{y} = \{y(i,j)\}_{(i,j) \in [1,\ldots,N_x] \times [1,\ldots,N_y]}$, all Bayesian inference is based of the posterior distribution

$$p(\boldsymbol{\theta}_k, k | \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta}_k, k)\, p(\boldsymbol{\theta}_k | k)\, p(k)}{p(\mathbf{y})} \qquad (12)$$

From this posterior distribution, it is for example possible to compute Bayes factor $p(\mathbf{y}|k)$:

$$p(\mathbf{y}|k) = \int_{\boldsymbol{\Theta}_k} p(\mathbf{y} | \boldsymbol{\theta}_k, k)\, p(\boldsymbol{\theta}_k | k)\, d\boldsymbol{\theta}_k \qquad (13)$$

and the conditional expectation of the parameters $\mathbb{E}[\boldsymbol{\theta}_k|\mathbf{y},k]$ for a given model order

$$\mathbb{E}[\boldsymbol{\theta}_k|\mathbf{y},k] = \int_{\boldsymbol{\Theta}_k} \boldsymbol{\theta}_k p(\boldsymbol{\theta}_k|\mathbf{y},k)\, d\boldsymbol{\theta}_k \tag{14}$$

Bayesian inference is based on the posterior distribution $p(\boldsymbol{\theta}_k,k|\mathbf{y})$. Unfortunately, it is impossible to express analytically this distribution and its features of interest as it involves integrating high-dimensional complex functions.

3 BAYESIAN COMPUTATION USING REVERSIBLE JUMP MCMC

We develop here a Markov Chain Monte Carlo methods (MCMC) to estimate the posterior distribution $p(\boldsymbol{\theta}_k,k|\mathbf{y})$, see [8] for a review on MCMC methods. The posterior distribution $p(\boldsymbol{\theta}_k,k|\mathbf{y})$ being defined on a union of subspaces of different dimensions, it is not possible to use classical MCMC methods. A simple solution would consist of upper bounding k by say $k_{\max}$ and running $k_{\max}+1$ independent MCMC samplers, each being associated to a fixed model order $k=0,...,k_{\max}$. However, this approach suffers from severe drawbacks. Firstly, it is computationally very expensive since $k_{\max}$ can be large. Secondly, the same computational effort is attributed to each value of k. In fact, some of these values are of no interest in practice because they have a very weak posterior Bayes factor $p(k|\mathbf{y})$. One alternative solution would be to construct an MCMC sampler that would be able to sample directly from the joint distribution on $\boldsymbol{\Theta} = \cup_{k=0}^{k_{\max}} \boldsymbol{\Theta}_k$. Standard MCMC methods are not able to "jump" between subspaces $\boldsymbol{\Theta}_k$ of different dimensions. Recently, Green has introduced a new efficient class of MCMC samplers, the so-called reversible jump MCMC that are able to solve such problems [4]. His method is based on a general state-space MH algorithm. One proposes candidates according to a set of proposal distributions. These candidates are randomly accepted according to an acceptance ratio which ensures reversibility and thus invariance of the Markov chain with respect to the posterior distribution. Here, the chain must move across subspaces of different dimensions, and therefore the proposal distributions are more complex. Specifically, for general moves between subspaces, a naive evaluation of the acceptance ratio would be impossible as it would require the evaluation of the ratio of probability measures between subspaces of different dimensions. To avoid this problem, Green has proposed to perform reversible jumps between different subspaces via proper dimension matching, see [4] for details. The following reversible jumps have been selected in our application:

1. birth of a new source,
2. death of an existing source,
3. merge of two neighbour sources,
4. split of a source into two neighbour sources,

Moves (1) to (4) perform dimension changes from k to $k-1$ or $k+1$. These moves are defined from purely heuristic considerations, the only condition to be fulfilled being to maintain the correct equilibrium distribution. A particular choice will only have influence on the rate of convergence of the algorithm. Other moves may be proposed, but we have found the one proposed satisfactory. The algorithm proceeds as follows.

Reversible Jump MCMC Sampler

1. Initialization. Start with a random $k^{(0)}$ and associated random parameters $\boldsymbol{\theta}_{k^{(0)}} = \left\{\mathbf{c}_{i,k^{(0)}}, A_{i,k^{(0)}}, \sigma^2_{i,k^{(0)}}\right\}$ for $i=1,\ldots,k^{(0)}$. Set $p=1$.
2. Iteration p.
   (a) Select one of the following moves uniformly at random.
   - Birth of a new source whose parameters are proposed randomly.
   - Death of a source selected randomly.
   - Split of a source selected randomly into two close sources.
   - Merge of two close sources into one.

   (b) Update the values of the current parameters $\boldsymbol{\theta}_{k^{(p)}}$.
3. Set $p \leftarrow p+1$ and go to step 2.

The update move is based on classical Metropolis Hastings steps. For further details, the reader is invited to consult [6].

4 SIMULATION RESULTS

We applied our algorithm to a real image of the Gamma camera developed by the CEA. This picture represents two radioactive cobalt sources which are very close. In Fig. 1, a zoom on this image is displayed. We run our algorithm for 1000 iterations, the first 300 iterations are assumed to correspond to the so-called "burn-in" period of the Markov chain and are discarded. In Fig. 2, we present the simulated values of the model order $\left\{k^{(p)}; p=1,\ldots,1000\right\}$.

It clearly appears that, according to a posterior Bayes factor criterion, the model with $k=2$ sources is the selected. In Tab. 1, we give the conditional expectation of the parameters for $k=2$, that is $\mathbb{E}[\boldsymbol{\theta}_2|\mathbf{y},2]$.

i	$c^1_{i,2}$	$c^2_{i,2}$	$A_{i,2}$	$\sigma^2_{i,2}$
1	20.1	16.2	15.2	3.6
2	27.1	19.6	18.8	3.4

Tab. 1: conditional expectation of the parameters

The estimated source image $x(i,j)$ with parameters $k = 2$ and $\mathbb{E}[\boldsymbol{\theta}_2|\mathbf{y},k{=}2]$ is displayed in Fig. 3.

5 CONCLUSION

We have addressed the problem of deconvolution of point sources in nuclear imaging using a high level Bayesian approach that allows to address jointly the problems of detection and estimation. To perform Bayesian computation, a reversible jump Markov chain Monte Carlo sampler has been developed. This algorithm displays satisfactory results on both synthetic and real data.

6 ACKNOWLEDGMENTS

The authors would like to acknowledge E. Barat, O. Gal and J.C. Trama.

References

[1] A. J. Baddeley and M.N.M Van Lieshout, "Stochastic Geometry Models in High Level Vision", in K.V Mardia and G.K Kanji (eds), *Statistics and Image,* vol. 20, Abingdon: Carfax Publishing, chap. 11, 1995, pp. 235-256.

[2] J.O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag 1985.

[3] J. Besag, P.J. Green, D.Higdon et K. Mengersen, "Bayesian Computation and Stochastic Systems", *Stat. Sc.*, vol. 10, 1995, pp. 3-66.

[4] P. J Green, "Reversible Jump MCMC Computation and Bayesian Model Determination", *Biometrika*, vol. 82, 1995, pp. 711-732.

[5] G. Stawinski, A. Doucet and P. Duvaut, "A Cascaded Point Process Model for Bayesian Image Restoration", in *Proc. ISCP'98 IASTED*, Las Palmas, Spain, Feb. 1998.

[6] G. Stawinski, A. Doucet and P. Duvaut, "Reversible Jump Markov Chain Monte Carlo for Bayesian Deconvolution of Point Sources with Poisson Statistics", in *Proc. SPIE Conf. Bayesian Inference and Inverse Problems*, San Diego, USA, July 1998.

[7] D.J. Strauss, "A Model for Clustering", *Biometrika*, vol. 62, 1975, pp. 467-475.

[8] L. Tierney, "Markov Chains for Exploring Posterior Distribution", *Ann. Stat.,* 1994, pp. 1701-1728.

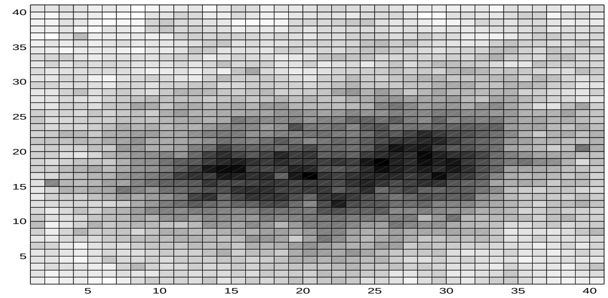

Figure 1: Observed image

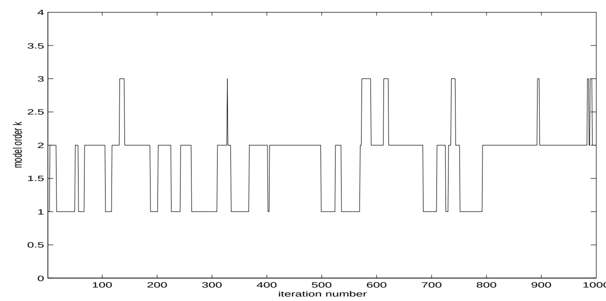

Figure 2: Simulated Markov chain

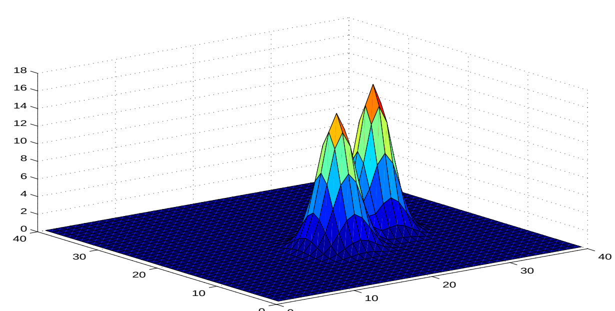

Figure 3: 3D representation: estimated point sources