ROBUST BAYESIAN SPECTRAL ANALYSIS VIA MCMC SAMPLING

*Arnaud Doucet - Christophe Andrieu
Signal Processing Group, Department of Engineering
University of Cambridge, Trumpington St., CB2 1PZ Cambridge, UK.
e-mail: ad2@eng.cam.ac.uk - ca226@eng.cam.ac.uk

ABSTRACT

In this paper, the harmonic retrieval problems in white Gaussian noise, non-Gaussian impulsive noise and in presence of threshold observations are addressed using a Bayesian approach. Bayesian models are proposed that allow us to define posterior distributions on the parameter space. All Bayesian inference is then based on these distributions. Unfortunately, a direct evaluation of these latters and of their features requires evaluation of some complicated high-dimensional integrals. Efficient stochastic algorithms based on Markov chain Monte Carlo methods are presented to perform Bayesian computation. In simulation, these algorithms are able to estimate the unknown parameters in highly degraded conditions.

1 INTRODUCTION

The harmonic retrieval problem is a fundamental problem in signal processing. Numerous methods have been developed to solve this problem. In this paper we follow a Bayesian approach. We address the problem of Bayesian harmonic retrieval in Gaussian noise and two important practical extensions of this problem. The first extension considers the case where the observation noise is not Gaussian but impulsive. Such heavy-tailed noise arises in many applications in signal processing. The second extension is the harmonic retrieval problem with hard-threshold observations, this last problem arises in radar science.

Bayesian models are proposed that allow us to define posterior probability distributions over the space of possible structures of the signal. Bayesian estimation of harmonic signals in white Gaussian noise has already been the subject of many recent works [1, 2, 4, 5]. The main problem of this attractive approach is that the posterior distribution appears highly nonlinear in its parameters, thus precluding analytical calculations. Except in some simple cases, where it is possible to perform analytical approximations [2], estimation of the posterior distribution requires numerical methods. Deterministic numerical methods and classical Monte Carlo methods have already been proposed [4] but they are not flexible and require the careful tuning of many parameters to be efficient. That is why recently Markov chain Monte Carlo (MCMC) methods [6] have been proposed to solve this problem, see [1] and the references therein. We present here efficient MCMC algorithms to perform Bayesian computation for the three problems adressed and demonstrate their performance using computer simulations.

* reversed alphabetical order

2 HARMONIC RETRIEVAL IN GAUSSIAN NOISE

2.1 Bayesian model

Let $\mathbf{y} \triangleq (y_0, y_1, \ldots, y_{T-1})^{\mathrm{T}}$ be an observed vector of T real data samples. $\mathbf{y}$ is the superimposition of k, $1 \leq k < \lfloor T/2 \rfloor$, sinusoids corrupted by noise, where the noise sequence $\mathbf{n} \triangleq (n_0, \ldots, n_{T-1})^{\mathrm{T}}$ is zero-mean white Gaussian noise of variance σ^2. In a vector-matrix form, we have

$$\mathbf{y} = \mathbf{D}(\boldsymbol{\omega})\,\mathbf{a} + \mathbf{n} \tag{1}$$

where we denote $\mathbf{D}(\boldsymbol{\omega})$ as the $T \times 2k$ matrix defined by

$$\begin{aligned}
\mathbf{D}(\boldsymbol{\omega}) &\triangleq \begin{pmatrix} \mathbf{d}_{c_1} & \mathbf{d}_{s_1} & \mathbf{d}_{c_2} & \cdots & \mathbf{d}_{c_k} & \mathbf{d}_{s_k} \end{pmatrix} \\
\mathbf{d}_{c_j} &\triangleq \begin{pmatrix} 1 & \cos[\omega_j] & \cdots & \cos[\omega_j (T-1)] \end{pmatrix}^{\mathrm{T}} \\
\mathbf{d}_{s_j} &\triangleq \begin{pmatrix} 0 & \sin[\omega_j] & \cdots & \sin[\omega_j (T-1)] \end{pmatrix}^{\mathrm{T}} \\
\mathbf{a} &\triangleq \begin{pmatrix} a_{c_1} & a_{s_1} & \cdots & \cdots & a_{c_k} & a_{s_k} \end{pmatrix}^{\mathrm{T}} \\
\boldsymbol{\omega} &\triangleq (\omega_1, \ldots, \omega_k)^{\mathrm{T}}
\end{aligned}$$

with $a_{c_j} = a_j \cos[\phi_j]$ and $a_{s_j} = -a_j \sin[\phi_j]$, a_j, ω_j and ϕ_j being respectively the amplitude, the radial frequency and the phase of the j^{th} sinusoid. The parameters $\boldsymbol{\theta} \triangleq (\boldsymbol{\omega}, \mathbf{a}, \sigma^2)$ are unknown. To complete the Bayesian model, we set a prior distribution on $\boldsymbol{\theta}$. We assume a vague improper prior distribution for $\boldsymbol{\theta}$ [1, 2, 5]:

$$p(\mathbf{a}, \boldsymbol{\omega}, \sigma^2) \propto \mathbb{I}_{\Omega}(\boldsymbol{\omega}) / \sigma^2 \tag{2}$$

where $\Omega \triangleq \left\{ \boldsymbol{\omega} \in (0, \pi)^k ; 0 \leq \omega_1 < \omega_2 < \ldots < \omega_k < \pi \right\}$. $\propto$ means "proportional to" and $\mathbb{I}_A(\cdot)$ is the indicator function of the set A.

2.2 Estimation objectives

Bayesian inference about $\boldsymbol{\theta}$ is based on the posterior distribution $p(\boldsymbol{\theta}|\,\mathbf{y})$ obtained from Bayes' theorem. Our aim is to estimate this joint distribution from which, by standard probability marginalisation and transformation techniques, one can obtain all posterior features of interest including the marginal distributions, posterior modes or conditional expectations. To develop an efficient algorithm, it is worth noticing that $p(\boldsymbol{\omega}|\,\mathbf{y})$ can be evaluated analytically up to a normalizing constant. Indeed, by straightforward calculations [1], we obtain:

$$p(\boldsymbol{\omega}|\,\mathbf{y}) \propto |\mathbf{M}(\boldsymbol{\omega})|^{\frac{1}{2}} \left[\mathbf{y}^{\mathrm{T}}\mathbf{P}(\boldsymbol{\omega})\,\mathbf{y}\right]^{-\frac{T-2k}{2}} \mathbb{I}_{\Omega}(\boldsymbol{\omega}) \qquad (3)$$

with

$$\begin{aligned} \mathbf{M}^{-1}(\boldsymbol{\omega}) &= \mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega})\,\mathbf{D}(\boldsymbol{\omega}) \\ \mathbf{m}(\boldsymbol{\omega}) &= \mathbf{M}(\boldsymbol{\omega})\,\mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega})\,\mathbf{y} \\ \mathbf{P}(\boldsymbol{\omega}) &= \mathbf{I}_T - \mathbf{D}(\boldsymbol{\omega})\,\mathbf{M}(\boldsymbol{\omega})\,\mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega}) \end{aligned} \qquad (4)$$

2.3 Bayesian computation

Denoting $\boldsymbol{\omega}_{-j}^{(i)} \triangleq \left(\omega_1^{(i)}, ..., \omega_{j-1}^{(i)}, \omega_{j+1}^{(i-1)}, ..., \omega_k^{(i-1)}\right)$, the iterative MCMC algorithm proceeds as follows. The parameters are randomly initialized $\left\{\mathbf{a}^{(0)}, \boldsymbol{\omega}^{(0)}, \sigma^{2(0)}\right\}$ at iteration $i = 0$, then the algorithm proceeds as follows at iteration i, $i \geq 1$.

For $j = 1, ..., k$

- Sample ω_j according to a mixture of two Metropolis-Hastings (MH) kernels of invariant distribution $p\left(\omega_j|\,\mathbf{y}, \boldsymbol{\omega}_{-j}^{(i)}\right)$, see 2.3.1.

End For.

- Optional step. Sample $\left(\mathbf{a}^{(i)}, \sigma^{2(i)}\right) \sim p\left(\mathbf{a}, \sigma^2\middle|\,\mathbf{y}, \boldsymbol{\omega}^{(i)}\right)$.

These steps are detailed in the following subsections. In what follows, in order to simplify notation, we drop the superscript $\cdot^{(i)}$ from all variables at iteration i.

2.3.1 Sampling the frequencies

We sample the frequencies one-at-a-time using MH steps [6]. The target distribution is the so-called full conditional distribution of a frequency,

$$p\left(\omega_j|\,\mathbf{y}, \boldsymbol{\omega}_{-j}\right) \propto |\mathbf{M}(\boldsymbol{\omega})|^{\frac{1}{2}} \left[\mathbf{y}^{\mathrm{T}}\mathbf{P}(\boldsymbol{\omega})\,\mathbf{y}\right]^{-\frac{T-2k}{2}} \mathbb{I}_{\Omega}(\boldsymbol{\omega}) \qquad (5)$$

With probability $0 < \rho < 1$, we perform a MH step with proposal distribution $q_1\left(\omega_j'|\,\omega_j\right)$ independent of the current state ω_j:

$$q_1\left(\omega_j'|\,\omega_j\right) \propto \sum_{l=0}^{T_p-1} p_l \mathbb{I}_{[l\pi/T_p,(l+1)\pi/T_p)}\left(\omega_j'\right) \qquad (6)$$

where p_l is the value of the squared modulus of the FFT of the observations $\mathbf{y}$ at frequency $l\pi/T_p$. This proposal distribution allows the Markov chain to reach quickly the regions of interest of the posterior distribution. With probability $1-\rho$, we perform a MH step with proposal distribution $q_2\left(\omega_j'|\,\omega_j\right)$ satisfying

$$\omega_j'|\,\omega_j \sim \mathcal{N}\left(\omega_j, \sigma_{RW}^2\right) \qquad (7)$$

$\mathcal{N}(\mathbf{m}, \boldsymbol{\Sigma})$ being a normal distribution of mean $\mathbf{m}$ and covariance $\boldsymbol{\Sigma}$. This random walk is introduced to perform a local exploration of the posterior distribution. In both cases ω_j' is accepted with the probability

$$\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}(\boldsymbol{\omega})\,\mathbf{y}}{\mathbf{y}^{\mathrm{T}}\mathbf{P}(\boldsymbol{\omega}')\,\mathbf{y}}\right)^{\frac{T-2k}{2}} \frac{|\mathbf{M}(\boldsymbol{\omega}')|^{1/2}\, q_i(\omega|\,\omega')}{|\mathbf{M}(\boldsymbol{\omega})|^{1/2}\, q_i(\omega'|\,\omega)} \qquad (8)$$

for $i = 1, 2$ where $\boldsymbol{\omega}' \triangleq \left(\omega_1, ..., \omega_{j-1}, \omega_j', \omega_{j+1}, ..., \omega_k\right)$.

2.3.2 Sampling the nuisance parameters

If one is not interested in estimating $\left(\mathbf{a}, \sigma^2\right)$, then it is not necessary to sample these parameters contrary to [5]. Otherwise, we obtain by straightforward calculations:

$$\sigma^2|\left(\mathbf{y}, \boldsymbol{\omega}\right) \sim \mathcal{IG}\left(\frac{T-2k}{2}, \frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}(\boldsymbol{\omega})\,\mathbf{y}}{2}\right) \qquad (9)$$

$$\mathbf{a}|\left(\mathbf{y}, \boldsymbol{\omega}, \sigma^2\right) \sim \mathcal{N}\left(\mathbf{m}(\boldsymbol{\omega}), \mathbf{M}(\boldsymbol{\omega})\right) \qquad (10)$$

$\mathcal{IG}$ denoting the inverse Gamma distribution.

This algorithm is much more efficient practically and theoretically than the one proposed in [5]. Theoretically, it converges uniformally geometrically towards the posterior distribution [1] and in practice we observe a quick convergence of the simulated Markov chain.

3 HARMONIC RETRIEVAL IN NON-GAUSSIAN NOISE

We extend here this algorithm to non-Gaussian impulsive noise modeled by a mixture of Gaussians.

3.1 Bayesian model and Estimation objectives

The model is similar to the one developed in Section 2 except that the noise sequence $\mathbf{n} \triangleq (n_0, \ldots, n_{T-1})^{\mathrm{T}}$ is now i.i.d. modeled as a mixture of Gaussians, *i.e.* $n_t|\,\sigma_t^2 \sim \mathcal{N}\left(0, \sigma_t^2\right)$ where σ_t^2 is itself a random variable.

Finite mixture: The noise is here modeled as a two-component Gaussian mixture, *i.e.*

$$\sigma_t^2|\left(\alpha^2, \sigma^2, \lambda\right) \sim \lambda\delta_{\sigma^2}(d\cdot) + (1-\lambda)\,\delta_{\alpha^2\sigma^2}(d\cdot) \qquad (11)$$

where $\alpha^2 < 1$ and $\delta_u(d\cdot)$ denotes the delta-Dirac measure located in u. With probability $1-\lambda$, $n_t|\,\sigma_t^2 \sim \mathcal{N}\left(0, \alpha^2\sigma^2\right)$ and otherwise $n_t|\,\sigma_t^2 \sim \mathcal{N}\left(0, \sigma^2\right)$. This last component allows to model impulsive noise. It is of interest to introduce the unobserved missing data set $\boldsymbol{\xi} \triangleq (\xi_0, \ldots, \xi_{T-1})^{\mathrm{T}}$ which take values in $\{l_1, l_2\}^T$ and such that $\Pr\left(\xi_t = l_1|\,\lambda, \alpha^2\right) = \lambda$ and

$$\begin{aligned} n_t|\left(\alpha^2, \sigma^2, \xi_t = l_1\right) &\sim \mathcal{N}\left(0, \sigma^2\right) \\ n_t|\left(\alpha^2, \sigma^2, \xi_t = l_2\right) &\sim \mathcal{N}\left(0, \alpha^2\sigma^2\right) \end{aligned} \qquad (12)$$

We assume that $p\left(\boldsymbol{\theta}, \alpha^2, \lambda,\right) = p(\boldsymbol{\theta})\,p\left(\alpha^2, \lambda\right)$ where α^2 and λ are assumed distributed according to uniform prior distributions $\alpha^2 \sim \mathcal{U}_{(0,1)}$, $\lambda \sim \mathcal{U}_{(0,1)}$.

Continuous mixture: The modeling by a finite mixture can be restrictive. In particular, it does not allow to model Cauchy and Laplacian noises. One can use a continuous mixture of Gaussians to solve this problem [1]. Because of space limitations, we only present here the algorithm to compute the posterior distribution in the case of finite mixture distributions. More precisely, we derive an algorithm to estimate $p(\boldsymbol{\theta},\boldsymbol{\xi}|\,\mathbf{y})$ as, similarly to (3), one can evaluate analytically up to a normalizing constant $p(\boldsymbol{\omega}|\,\mathbf{y},\boldsymbol{\xi})$.

3.2 Bayesian computation

At iteration $i \geq 1$, the MCMC algorithm proceeds as follows.

- For $j = 1,...,k$
- Sample ω_j according to a mixture of MH kernels of invariant distribution $p\left(\omega_j|\,\mathbf{y},\boldsymbol{\xi}^{(i-1)},\boldsymbol{\omega}_{-j}^{(i)}\right)$.

End For.
For $t = 0,...,T-1$

- Sample $\xi_t^{(i)} \sim p\left(\xi_t|\,\mathbf{y},\boldsymbol{\xi}_{-t}^{(i)},\boldsymbol{\omega}^{(i)},\alpha^{2(i-1)},\lambda^{(i-1)}\right)$.

End For.

- Sample $\left(\mathbf{a}^{(i)},\sigma^{2(i)}\right) \sim p\left(\mathbf{a},\sigma^2\middle|\,\mathbf{y},\boldsymbol{\xi}^{(i)},\boldsymbol{\omega}^{(i)},\alpha^{2(i-1)}\right)$.
- Sample $\left(\alpha^{2(i)},\lambda^{(i)}\right) \sim p\left(\alpha^2,\lambda\middle|\,\mathbf{y},\boldsymbol{\xi}^{(i)},\boldsymbol{\omega}^{(i)},\mathbf{a}^{(i)},\sigma^{2(i)}\right)$.

3.2.1 *Sampling the frequencies and the missing data*

To sample the frequencies, we adopt a similar strategy to the one presented in 2.3.1 except that the target distribution is now $p\left(\omega_j|\,\mathbf{y},\boldsymbol{\xi}^{(i-1)},\boldsymbol{\omega}_{-j}^{(i)}\right)$ and that $q_1\left(\omega_j'\middle|\,\omega_j\right)$ is based at iteration i on the FFT of the observations $\mathbf{y}$ weighted with the current inverse covariance matrix $\overline{\boldsymbol{\Delta}}^{-1}$ of the noise determined by $\boldsymbol{\xi}^{(i-1)}$, $\overline{\boldsymbol{\Delta}}$ being defined in 3.2.2. Sampling the missing data is easy as $p\left(\xi_t|\,\mathbf{y},\boldsymbol{\xi}_{-t},\boldsymbol{\omega},\alpha^2,\lambda\right)$ is a discrete probability distribution that can be evaluated.

3.2.2 *Sampling the nuisance parameters*

$$\begin{aligned}
\sigma^2\middle|\left(\mathbf{y},\boldsymbol{\xi},\boldsymbol{\omega},\alpha^2\right) &\sim \mathcal{IG}\left(\frac{T-2k}{2},\frac{\mathbf{y}^{\mathrm{T}}\overline{\mathbf{P}}(\boldsymbol{\omega})\,\mathbf{y}}{2}\right) && (13)\\
\mathbf{a}|\left(\mathbf{y},\boldsymbol{\xi},\boldsymbol{\omega},\alpha^2,\sigma^2\right) &\sim \mathcal{N}\left(\overline{\mathbf{m}}(\boldsymbol{\omega}),\overline{\mathbf{M}}(\boldsymbol{\omega})\right) && (14)
\end{aligned}$$

with

$$\begin{aligned}
\overline{\mathbf{M}}^{-1}(\boldsymbol{\omega}) &= \mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega})\,\overline{\boldsymbol{\Delta}}^{-1}\mathbf{D}(\boldsymbol{\omega})\\
\overline{\mathbf{m}}(\boldsymbol{\omega}) &= \overline{\mathbf{M}}(\boldsymbol{\omega})\,\mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega})\,\overline{\boldsymbol{\Delta}}^{-1}\mathbf{y} \qquad (15)\\
\overline{\mathbf{P}}(\boldsymbol{\omega}) &= \mathbf{I}_T-\mathbf{D}(\boldsymbol{\omega})\,\overline{\mathbf{M}}(\boldsymbol{\omega})\,\mathbf{D}^{\mathrm{T}}(\boldsymbol{\omega})
\end{aligned}$$

$\overline{\boldsymbol{\Delta}}$ being a $T \times T$ diagonal matrix with $\overline{\boldsymbol{\Delta}}_{i,i} = \mathbb{I}_{\{\xi_{i-1}=l_1\}}+\alpha^2\mathbb{I}_{\{\xi_{i-1}=l_2\}}$.

3.2.3 *Sampling the mixture parameters*

Denoting $n_1 = \sum_{t=0}^{T-1}\mathbb{I}_{\left\{\xi_t^{(i)}=l_1\right\}}$ and $n_2 = T-n_1$:

$$\lambda|\,\boldsymbol{\xi}^{(i)} \sim \mathcal{B}e\left(n_1+1,n_2+1\right) \qquad (16)$$

$\mathcal{B}e$ being the Beta distribution. For $n_2 \geq 2$

$$\alpha^2\middle|\left(\mathbf{y},\boldsymbol{\xi},\boldsymbol{\omega},\mathbf{a},\sigma^2\right) \sim \mathcal{IG}\left(\frac{n_2-2}{2},\frac{\gamma}{2}\right)\mathbb{I}_{(0,1)}(\cdot) \qquad (17)$$

For $0 \leq n_2 \leq 1$, efficient rejection methods can be used.

4 HARMONIC RETRIEVAL WITH THRESHOLD DATA

4.1 Signal model and Estimation objectives

The statistical model for $\mathbf{y} = (y_0,y_1,\ldots,y_{T-1})^{\mathrm{T}}$ is similar to the one in Section 2, *i.e.* k sinusoids in white Gaussian noise. But in this case, we do not observe $\mathbf{y}$ but $\mathbf{z} \triangleq (z_0,z_1,\ldots,z_{T-1})^{\mathrm{T}}$, where $\mathbf{z}$ is a threshold version of $\mathbf{y}$, *i.e.*

$$\begin{aligned}
z_j &= y_{\max} && \text{if } y_j \geq y_{\max}\\
z_j &= y_j && \text{if } y_{\min} \leq y_j \leq y_{\max} \qquad (18)\\
z_j &= y_{\min} && \text{if } y_j \leq y_{\min}
\end{aligned}$$

$y_{\min}$ and $y_{\max}$ ($y_{\min} < y_{\max}$) being known thresholds. The parameters of the sinusoids and the variance of the noise are unknown. Given the data set $\mathbf{z}$, our objective is to estimate these parameters, that is $\boldsymbol{\theta} \triangleq \left(\mathbf{a}^{\mathrm{T}},\boldsymbol{\omega}^{\mathrm{T}},\sigma^2\right)^{\mathrm{T}}$ where $p(\boldsymbol{\theta}) \propto \mathbb{I}_{\Omega}(\boldsymbol{\omega})/\,\sigma^2$.

Our aim is to estimate the posterior distribution $p(\boldsymbol{\theta}|\,\mathbf{z})$ and its features. More precisely, we derive an algorithm to estimate $p(\boldsymbol{\theta},\mathbf{y}|\,\mathbf{z})$ as, similarly to (3), one can evaluate analytically up to a normalizing constant $p(\boldsymbol{\omega}|\,\mathbf{y},\mathbf{z}) = p(\boldsymbol{\omega}|\,\mathbf{y})$.

4.2 Bayesian computation

At iteration $i \geq 1$, the MCMC algorithm proceeds as follows.

- Sample $\mathbf{y}^{(i)} \sim p\left(\mathbf{y}|\,\mathbf{z},\mathbf{a}^{(i-1)},\boldsymbol{\omega}^{(i-1)},\sigma^{2(i-1)}\right)$.

For $j = 1,...,k$

- Sample ω_j according to a mixture of two MH kernels of invariant distribution $p\left(\omega_j|\,\mathbf{y}^{(i)},\boldsymbol{\omega}_{-j}^{(i)}\right)$.

End For.

- Sample $\left(\mathbf{a}^{(i)},\sigma^{2(i)}\right)$ according to $p\left(\mathbf{a},\sigma^2\middle|\,\mathbf{y}^{(i)},\boldsymbol{\omega}^{(i)}\right)$.

4.2.1 *Sampling the missing data* $\mathbf{y}$

The missing data $\mathbf{y}$ are conditionally statistically independent and we easily obtain:

$$\begin{aligned}
y_t|\left(z_t,\mathbf{a},\boldsymbol{\omega},\sigma^2\right) &\sim \delta_{z_t}(dy_t) \text{ if } (y_{\min} \leq z_t \leq y_{\max}) && (19)\\
y_t|\left(z_t,\mathbf{a},\boldsymbol{\omega},\sigma^2\right) &\sim \mathcal{N}\left(\overline{y}_t,\sigma^2\right)\mathbb{I}_{(-\infty,y_{\min})}(\cdot) \text{ if } z_t < y_{\min}\\
y_t|\left(z_t,\mathbf{a},\boldsymbol{\omega},\sigma^2\right) &\sim \mathcal{N}\left(\overline{y}_t,\sigma^2\right)\mathbb{I}_{(y_{\max},+\infty)}(\cdot) \text{ if } z_t > y_{\max}
\end{aligned}$$

where $\overline{y}_t = \mathbf{e}_{t+1}^{\mathrm{T}}\mathbf{D}(\boldsymbol{\omega})\,\mathbf{a}$, $\mathbf{e}_{t+1}$ being a T-dimensional vector with all components equal to 0 except the component t which is equal to 1.

4.2.2 *Sampling the frequencies and the nuisance parameters*

To sample the frequencies we adopt the same strategy as in 2.3.1 except that $q_1\left(\omega_j'\middle|\,\omega_j\right)$ is based at iteration i on the FFT of the simulated observations $\mathbf{y}^{(i)}$. To sample the nuisance parameters, the method is similar to the one described in 2.3.2. The main difference is that, even if one is not interested in estimating $\left(\mathbf{a},\sigma^2\right)$, this simulation step is necessary to sample the missing data $\mathbf{y}$.

5 SIMULATIONS

These algorithms require the specification of parameters that have no influence on the posterior distribution but only on the speed of convergence of the algorithm. We set $\rho = 0.2$, $T_p = T$ and $\sigma_{RW} = 1/(5T)$. These parameters have been determined in a rather heuristic way. They are the first values we tried and they provide the Markov chain with very satisfactory properties. The following parameters have been selected for the sinusoids: $T = 128$, $k = 2$. We define $E_i \triangleq a_{c_i}^2 + a_{s_i}^2$. $E_1 = 20$, $E_2 = 18$, $-\arctan(a_{s_1}/a_{c_1}) = 0$, $-\arctan(a_{s_2}/a_{c_2}) = \pi/4$, $f_1 \triangleq \omega_1/2\pi = 0.2$ and $f_2 \triangleq \omega_2/2\pi = 0.3$. First a signal in non-Gaussian impulsive noise has been simulated with $\sigma = 5.6$, $\lambda = 0.02$ and $\alpha^2 = 0.1$. In Fig. 1, this signal is displayed as well as the posterior distribution of the missing data $\boldsymbol{\xi}$. In Fig. 2, $p(f_1|\mathbf{y})$ and $p(f_2|\mathbf{y})$ are given. Then a threshold signal has been simulated $y_{\max} = -y_{\min} = 3$ and $\sigma = 5.6$. In Fig. 3, $\mathbf{y}$ and $\mathbf{z}$ are displayed. In Fig. 4, $p(f_1|\mathbf{z})$ and $p(f_2|\mathbf{z})$ are presented. Other results are presented in [1].

6 CONCLUSION

In this paper, the harmonic retrieval problems in Gaussian noise, non-Gaussian impulsive noise and with threshold observations have been treated. These complex statistical problems have been addressed in a Bayesian framework. Bayesian models have been proposed and efficient Markov chain Monte Carlo methods have been proposed to perform Bayesian computation. In simulations, these algorithms allow us to estimate in a satisfactory way the unknown parameters in difficult conditions. Extensions to continuous mixtures and colored noise cases are presented in [1].

References

[1] C. Andrieu and A. Doucet, "Robust Bayesian Harmonic Retrieval using MCMC," technical report, Dept. Eng., Univ. Cambridge, 1998.

[2] G.L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Note in Statistics, vol. 48, Springer-Verlag, New-York, 1988.

[3] L. Devroye, *Non-Uniform Random Variate Generation*, Springer, New-York, 1986.

[4] P.M. Djurić and H. Li, "Bayesian spectrum estimation of harmonic signals," *IEEE Sig. Proc. Letters*, vol. 2, no. 11, pp. 213-215, 1995.

[5] L. Dou and R.J.W. Hodgson, "Bayesian inference and Gibbs sampling in spectral analysis and parameter estimation: I," *Inverse Problems*, pp. 1069-1085, 1995.

[6] L. Tierney, "Markov chains for exploring posterior distributions," *Ann. Stat.*, pp. 1701-1728, 1994.

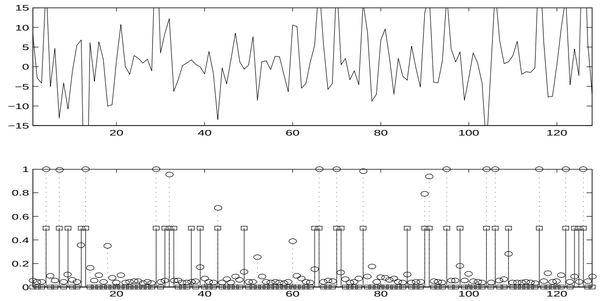

Figure 1: Top: observations $\mathbf{y}$ (threshold for presentation). Bottom: Missing data $\xi_t = l_1$ set to an arbitrary height 0.5 and $\Pr(\xi_t = l_1|\mathbf{y})$.

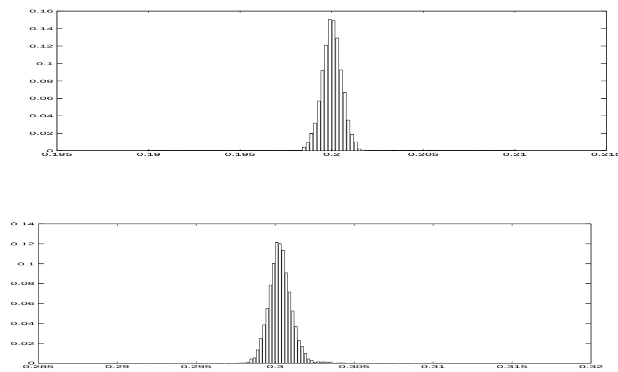

Figure 2: Marginal posterior distributions of the frequencies $p(f_1|\mathbf{y})$ and $p(f_2|\mathbf{y})$.

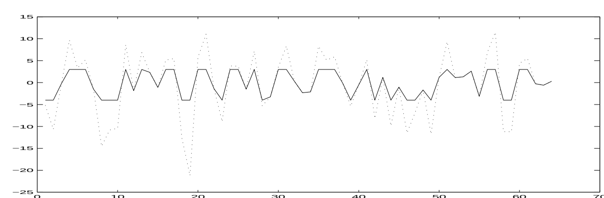

Figure 3: $\mathbf{y}$ and threshold obervations $\mathbf{z}$.

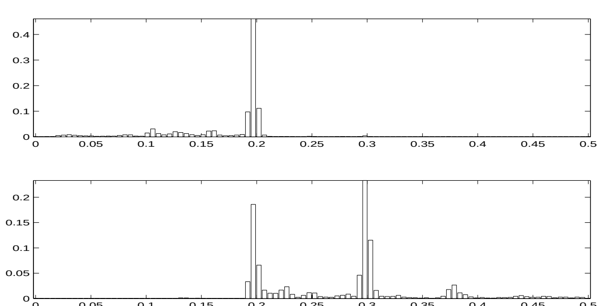

Figure 4: Marginal posterior distributions of the frequencies $p(f_1|\mathbf{z})$ and $p(f_2|\mathbf{z})$.