



HAL
open science

Two Bayesian methods for multipath propagation parameters estimation

Fabienne Porée, Olivier Rosec, Thierry Chonavel, Jean-Marc Boucher

► **To cite this version:**

Fabienne Porée, Olivier Rosec, Thierry Chonavel, Jean-Marc Boucher. Two Bayesian methods for multipath propagation parameters estimation. ICASSP 2000: IEEE International Conference on Acoustics, Speech and Signal Processing, Jun 2000, Istanbul, Turkey. pp.69 - 72, 10.1109/ICASSP.2000.861866 . hal-02141163

HAL Id: hal-02141163

<https://hal.science/hal-02141163>

Submitted on 27 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TWO BAYESIAN METHODS FOR MULTIPATH PROPAGATION PARAMETERS ESTIMATION

F. Porée, O. Rosec, T. Chonavel and J.M. Boucher

ENST de Bretagne
Technopôle de Brest Iroise, BP 832
29285 Brest cedex, France

ABSTRACT

In this paper we propose a bayesian approach for multipath channel time delay estimation when the signal is received by a square-law detector. Two different methods are presented. The first one is applied when the signal is supplied at the output of the detector. On certain approximations, the derivation of the MAP (Maximum A Posteriori) estimator leads to optimize a l_1 -norm criterion. In the second method, the signal is observed after demodulation and matched-filtering. We introduce a Bernoulli-Gaussian model to account for the sparse properties of the channel impulse response and solve the deconvolution problem using Monte-Carlo simulations. Comparisons between both algorithms are presented on simulations. They show that the second method leads to more accurate estimators, while justifying the validity of the approximations used in the first one.

1. INTRODUCTION

Many problems can be modeled as a linear system where the observed output signal $r(t)$ is the convolution of the input signal $e(t)$ with a sparse spike time series $h(t)$. Depending on the domain of application, the function $h(t)$ can be for instance the reflectivity sequence of a seismic trace [8], or a propagative channel multipath impulse response as in oceanic acoustic tomography [15]. In such cases, $h(t)$ can be written as

$$h(t) = \sum_{p=1}^P \alpha_p \delta(t - \tau_p), \quad (1)$$

where α_p and τ_p are respectively the attenuation and the delay associated to the p^{th} path, and P is the number of paths. In applications such as ocean acoustic tomography [15], $e(t)$ is known by the receiver. Then the problem of estimating the $(\alpha_p, \tau_p)_{p=1, P}$, and possibly the number P of paths, can be addressed via maximum likelihood estimation (see e.g. [2] [9]). Unfortunately, the log-likelihood criterion shows local maxima, and it requires the knowledge of the number of paths.

To overcome these limitations, we propose a bayesian approach [13] which allows to include information upon the parameters by means of a priori distributions. Here we take into account the fact that the paths amplitudes $(\alpha_p)_{p=1, P}$

and the noise have gaussian distributions. In this work, we present two different deconvolution methods. The first one addresses the case where the only available signal is obtained at the output of a square-law detector. On certain approximations, the derivation of the MAP estimator leads to a very simple l_1 -norm criterion. In the second method the signal is complex-valued, and is obtained after demodulation and matched-filtering. We introduce a Bernoulli-Gaussian model to account for the sparse properties of $h(t)$ and solve the deconvolution problem using Monte-Carlo simulations.

The reason for considering both approaches stems from the fact that the first method involves linear approximations of the non-linear transforms in the square-law detector. The simulation results demonstrate the superiority of the Bernoulli-Gaussian model. However we show that processing the signal at the output of the quadratic receiver is quite robust to the above approximations, and thus can be used when only the output of the square-law detector is available.

The problem is presented in section 2. In section 3 et 4, the two different bayesian methods are developed. Results and comparison are discussed in section 4.

2. FORMULATION OF THE PROBLEM

Let $s(t)$ denote the transmitted signal, $r(t)$ the received signal, and $\nu(t)$ an additive gaussian white noise with variance σ_ν^2 , independent from $h(t)$. At the transmitter side $s(t)$ is modulated by a carrier with pulsation ω . Then, the received signal $r(t)$ is in the form:

$$\begin{cases} r(t) = (h * e)(t) + \nu(t) = \sum_{p=1}^P \alpha_p e(t - \tau_p) + \nu(t), \\ \text{with } e(t) = s(t) \cos(\omega t - \phi). \end{cases} \quad (2)$$

At the receiver side, $r(t)$ goes through the square-law detector according to the scheme of the figure 1, where $\tilde{s}(t) = s(-t)$ is the impulse response of the matched-filter.

Let $\underline{x}(t)$ denote the complex signal obtained at the outputs of the matched-filters. It can also be written in the form:

$$\underline{x}(t) = (g * \underline{h})(t) + \underline{b}(t), \quad (3)$$

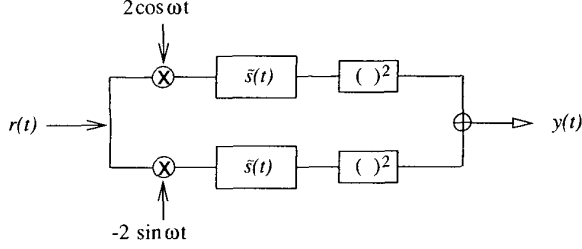


Figure 1: Structure of the receiver.

$$\text{where } \begin{cases} g(t) = (s * \bar{s})(t) \\ \underline{h} = \sum_{p=1}^P \alpha_p \delta(t - \tau_p) \\ \alpha_p = \alpha_p \cos(\omega\tau_p + \phi) + i \alpha_p \sin(\omega\tau_p + \phi). \end{cases}$$

Finally, the sampled observation in the observation interval, say $[1, N]$, $\underline{x} = (\underline{x}(n))_{n=1, N}$, can be written

$$\underline{x}(n) = \sum_{p=1}^P \alpha_p g(n - \tau_p) + \underline{b}(n), \quad (4)$$

where $\underline{b}(n)$ is a complex circular gaussian random variable, with variance $\sigma_{\underline{b}}^2 = 4\sigma_{\nu}^2$ (g is normalized with $g(0) = 1$).

3. L_1 NORM DECONVOLUTION AT THE OUTPUT OF THE RECEIVER

3.1. The received signal $y(t)$

In this first approach we consider, as already in [11], the case where the only available signal is obtained at the output of the quadratic receiver, and denoted $y(t)$. With a view to getting a simple expression of this signal, we consider two approximations. First, the product terms between the signal of interest and the noise are neglected, which is generally justified because of the high SNR at the output of the matched filter. Moreover, the product terms associated with two distinct paths are neglected, which is generally a satisfactory approximation when the autocorrelation function of $s(t)$ is sharp. Then, $y(t)$ can be written as:

$$\begin{aligned} y(t) &= (h * z)(t) + \epsilon(t) \\ &= \sum_{p=1}^P a_p z(t - \tau_p) + \epsilon(t), \end{aligned} \quad (5)$$

with $z(t) = g^2(t)$ and $a_p = |\alpha_p|^2$.

3.2. The l_1 -norm criterion

Assuming that the noise $\underline{b}(t)$ and the attenuations α_p [12] have gaussian distributions:

$$\alpha_p \sim \mathcal{N}(0, \sigma_{\alpha}^2), \quad (6)$$

$$\underline{b}(t) \sim \mathcal{N}(0, \sigma_{\underline{b}}^2), \quad (7)$$

it comes that a_p and $\epsilon(t)$ have exponential distributions:

$$a_p \sim \mathcal{E}(1/2\sigma_{\alpha}^2), \quad (8)$$

$$\epsilon(t) \sim \mathcal{E}(1/2\sigma_{\underline{b}}^2). \quad (9)$$

Without prior information upon the time delays, we assume that they are uniformly distributed in the observation interval: $\tau_p \sim U_{[1, N]}$.

Then, the posterior likelihood minimization problem is:

$$\min_{(a_p, \tau_p)} \frac{\left\| y(t) - \sum_{p=1}^P a_p z(t - \tau_p) \right\|_1}{2\sigma_{\underline{b}}^2} + \frac{\| \underline{a} \|_1}{2\sigma_{\alpha}^2}. \quad (10)$$

In order to optimize the criterion, the time scale is discretized in the same way as in [3] [6]. Finally, we get the problem in the form

$$\min_{\underline{h}} \| \underline{y} - \mathbf{S}_z \underline{h} \|_1 + \lambda \| \underline{h} \|_1, \quad (11)$$

In this criterion, $\lambda = \sigma_{\underline{b}}^2 / \sigma_{\alpha}^2$, \mathbf{S}_z is the matrix of convolution with the sampled $z(t)$, \underline{y} is the data vector, and \underline{h} is the vector of the amplitudes of the paths on the sampled time scale. Moreover let us remark that the spike train can be solved with higher resolution than the received data sampling interval [3] [14].

Also, let us note that the criterion (11) as already been considered in the litterature. In [14], the l_1 -norm term $\| \underline{y} - \mathbf{S}_z \underline{h} \|_1$ is justified by the spike preservation properties, and the penalty term $\lambda \| \underline{h} \|_1$ by the fact that $h(t)$ is sparse. Moreover, λ appears as the weighting factor selected to balance the conflicting priorities of data accountability and addressing the a priori assumption that $h(t)$ is sparse [10]. But in such methods the choice of the value of the parameter λ is often a problem. Here, the bayesian formulation of the problem leads to a simple interpretation of λ , as the inverse of a signal to noise ratio. Thus, it is generally possible to estimate λ in a simple way.

This criterion can be also rewritten in the form:

$$\min_{\underline{h}} \| \underline{c} - A \underline{h} \|_1, \quad (12)$$

with

$$A = \begin{pmatrix} \mathbf{S}_z \\ \lambda I \end{pmatrix}, \text{ and } \underline{c} = \begin{pmatrix} \underline{y} \\ \mathbf{0} \end{pmatrix}. \quad (13)$$

The minimization of the criterion (12) is implemented according to the algorithm presented in [1], that uses the simplex method applied to a linear programming formulation of the problem. Also, with a view to adaptative tracking of time-varying channels parameters, the convex criterion $\| \underline{c} - A \underline{h} \|_1$ can be minimized by means of a simple gradient algorithm.

4. BERNOULLI-GAUSSIAN DECONVOLUTION

4.1. Model presentation

Instead of deconvolving the signal after quadratic detection, we consider here the complex signal \underline{x} at the output of the matched filters. In this case, we are faced to a deconvolution problem in the presence of gaussian noise. The convolutional model (3) can be rewritten as:

$$\underline{x}(n) = \sum_{k=-K}^K g(k) \underline{h}(n - k) + \underline{b}(n). \quad (14)$$

When no prior is available upon \underline{h} the deconvolution is equivalent to the maximization of the likelihood function given by:

$$p(\underline{x}|\tau, \underline{\alpha}) = \frac{1}{(\pi\sigma_b^2)^N} \exp \frac{-(\underline{x} - \mathbf{S}_g \underline{h})^H (\underline{x} - \mathbf{S}_g \underline{h})}{\sigma_b^2}, \quad (15)$$

where \mathbf{S}_g denotes the convolution matrix deduced from g . The maximum likelihood estimate is not an appropriate solution, although it can be shown that it is the minimum variance unbiased estimator.

A major drawback of maximum likelihood deconvolution is that it does not take into account the fact that $\underline{h}(n)$ is a sparse sequence of arrival time. A possible way to regularize the problem is to introduce a Bernoulli variable q to indicate the presence of a path. Then $\underline{h}(n)$ can be modeled by a Bernoulli-Gaussian process defined by:

$$\underline{h}(n) \sim \eta \mathcal{N}(0, \sigma_1^2) + (1 - \eta) \mathcal{N}(0, \sigma_0^2), \quad (16)$$

where

$$\eta = P(q(n) = 1) = 1 - P(q(n) = 0) \quad (17)$$

is the prior probability of finding a path at each sample n . Strictly speaking a pure Bernoulli-Gaussian is obtained when $\sigma_0^2 = 0$. However, as the likelihood of a Dirac distribution is not defined we introduce a non-zero variance $\sigma_0^2 \ll \sigma_1^2$. In geophysics this model is commonly used to characterize the reflectivity of the subsurface, and σ_0^2 describes small heterogeneity in the sedimentary layers [4]. In nuclear science σ_0^2 models some background disturbance noise at the receiver.

4.2. The MCMC approach

In that framework, the problem is to study the posterior likelihood $p(\underline{h}, \mathbf{q}|\underline{x})$. Using Bayes rule and omitting the constant term $p(\underline{x})$ this density can be written as:

$$p(\underline{h}, \mathbf{q}|\underline{x}) \propto p(\underline{x}|\underline{h})p(\underline{h}|\mathbf{q})p(\mathbf{q}). \quad (18)$$

After some easy calculations the posterior log-likelihood becomes:

$$\mathcal{L}(\underline{h}, \mathbf{q}|\underline{x}) = -\frac{(\underline{x} - \mathbf{S}_g \underline{h})^H (\underline{x} - \mathbf{S}_g \underline{h})}{\sigma_b^2} - \frac{\underline{h}^H \mathbf{D} \underline{h}}{\sigma_1^2} - \frac{\underline{h}^H (\mathbf{I} - \mathbf{D}) \underline{h}}{\sigma_0^2} + n_1(\mathbf{q}) \left[\frac{1 - \eta}{\eta} \frac{\sigma_1}{\sigma_0} \right], \quad (19)$$

with $\mathbf{D} = \text{diag}(\mathbf{q})$. Let us notice that given a vector \mathbf{q} , \mathcal{L} is a quadratic function of \underline{h} and so its maximum can be obtained in closed form. Unfortunately the study of \mathcal{L} for the 2^N possible \mathbf{q} sequences is computationally untractable as soon as N becomes large.

To overcome this difficulty it is possible to use simulation methods. For the application we are interested in, suppose we are able to generate a sequence of independent samples $\{(\underline{h}^{(m)}, \mathbf{q}^{(m)}), m = 1, \dots, M\}$ according to $p(\underline{h}, \mathbf{q}|\underline{x})$. Then we can use these samples to make inference on the missing variables such as calculating conditionnal expectations $\mathbb{E}[\mathbf{q}|\underline{x}]$ and $\mathbb{E}[\underline{h}|\underline{x}]$.

Therefore our problem comes down to the simulation of $(\underline{h}, \mathbf{q})$ according to their posterior distribution. However

generating high dimension variables remains a difficult task as soon as the distributions are non-standard. Here the convolution introduces time dependency of the missing variables conditionally on the observation, and no direct simulation scheme is possible. One way of solving this problem is to use MCMC methods. The basic principle is to generate a Markov Chain whose equilibrium distribution is the target distribution. The most popular MCMC algorithms as well as details concerning their convergence properties can be found in [13].

4.3. The Gibbs sampler

One simulation scheme is to simulate the missing variables one sample at a time. Indeed it can be shown (see [5]) that conditionally on $\{\underline{h}(j), q(j), j \neq n\}$, $(\underline{h}(n), q(n))$ can be written as a mixture of gaussian distributions. This motivates the choice of a Gibbs sampler. The algorithm proceeds as follows:

1. Initialization : random choice of $(\underline{h}^{(0)}, \mathbf{q}^{(0)})$;
2. At iteration $m \leq M$, for $n = 1, \dots, N$:
 - choose site i ;
 - simulate $(\underline{h}(i), q(i))$ according to $p(\underline{h}(i), q(i)|\underline{x}, \underline{h}(j), q(j), j \neq i)$.

At each iteration, the sites are visited using a random permutation of $\{1, \dots, N\}$. The behavior of the algorithm can be divided into two periods: first a burning period of M_0 iterations and then a steady state period where it can be assumed that the drawn sampled are distributed according to their posterior distribution.

Then, we calculate for each sample $1 \leq n \leq N$

$$Q(n) = \frac{1}{M - M_0} \sum_{m=M_0+1}^M q^{(m)}(n), \quad (20)$$

$$\underline{H}(n) = \frac{1}{M - M_0} \sum_{m=M_0+1}^M \underline{h}^{(m)}(n). \quad (21)$$

Then we use the following decision rule:

- if $Q(n) > 0.5$ set $\hat{q}(n) = 1$ and $\hat{\underline{h}}(n) = Q(n)\underline{H}(n)$
- else set $\hat{q}(n) = 0$ and $\hat{\underline{h}}(n) = 0$.

5. RESULTS

We now apply these two methods to numerical data. Let $s(t)$ be a M.L.S. (Maximum Length Sequence) [12] of length $n = 2^9 - 1 = 511$ symbols with values $\pm 1/\sqrt{n}$. The auto-correlation of the M.L.S. is a triangular shape, carried by the interval $[-K, K]$, where K is the symbol duration. We assume that $y(t)$ is sampled at a rate $\delta = K/6$.

We simulate two arrival times τ_1 and τ_2 with amplitudes 1, for a SNR equal to 30 dB, ie $\sigma_v^2 = 10^{-3}$. The results are presented for 2 values of the difference $|\tau_1 - \tau_2|$. Figure 2 shows an example of the signals obtained at different locations of the detector.

For each scenario, the performances are tested on 100 traces, corresponding to 100 independent realizations of the

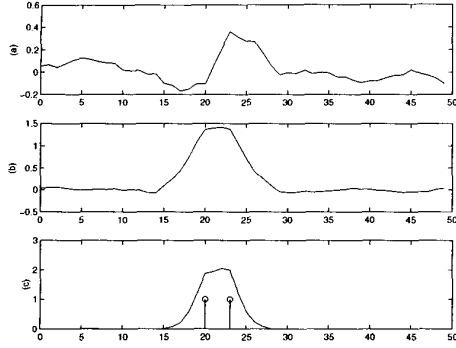


Figure 2: (a) Real part of $\underline{x}(t)$; (b) Imaginary part of $\underline{x}(t)$; (c) Output real signal $y(t)$.

noise. Figure 3 represents the image of the results obtained with each method.

The l_1 -norm method is implemented using standard Nag fortran routines [7]. The parameter λ is chosen equal to $1/2\sigma_0^2$. For the Bernoulli-Gaussian deconvolution algorithm, the model parameters are respectively $\sigma_b^2 = 4.10^{-3}$, $\sigma_1^2 = 1$, $\sigma_0^2 = 10^{-3}$, and $\eta = 0.04$. We run the Gibbs sampler for $M = 500$ iterations and discard the first 100.

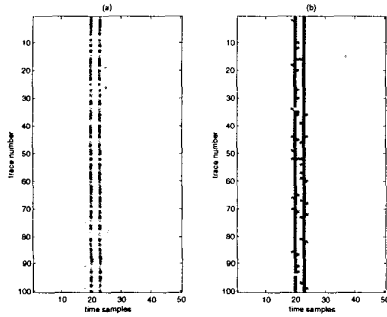


Figure 3: Results obtained on the 100 traces with $|\tau_1 - \tau_2| = K/2$: (a) l_1 -norm method; (b) MCMC method.

We focus here only on the estimation of the delays. Indeed, an accurate estimation of these parameters is of the greatest importance in oceanic acoustic tomography. For that purpose, we define P_{D_1} as the percentage of pulses detected not further than one sample from their true location. Besides P_{D_2} represents the percentage of exact estimation of time delays when the two largest maxima are selected. The results are presented in table 1.

$ \tau_1 - \tau_2 $	l_1		MCMC	
	K	$K/2$	K	$K/2$
P_{D_1}	100	96	100	100
P_{D_2}	91	88.5	98.5	97.5

Table 1: Detection results.

6. CONCLUSION

In this paper we have presented two methods for time delay estimation. We have seen that at the output of the square-law detector a very simple approximate MAP criterion can be considered, that yields good results in terms of path detection and delay estimation. Moreover when the complex data at the output of the matched-filter are available, better accuracy can be achieved via MCMC simulations of the parameters posterior distribution.

7. REFERENCES

- [1] I. Barrodale and F.D.K. Roberts. An improved algorithm for discrete II linear approximation. *SIAM Numer. Anal.*, 10(5):839–848, Oct. 1973.
- [2] M. Feder and E. Weinstein. Parameter estimation of superimposed signals using the em algorithm. *IEEE Trans. on A.S.S.P.*, 36(4):477–489, 1988.
- [3] J.-J. Fuchs. Multipath time-delay detection and estimation. *IEEE Trans. S.P.*, 47(1):237–243, Janv. 1999.
- [4] M. Lavielle. Bayesian deconvolution of bernoulli-gaussian processes. *Signal Processing*, 33:67–79, 1993.
- [5] M. Lavielle. A stochastic algorithm for parametric and non-parametric estimation in the case of incomplete data. *Signal Processing*, 42:3–17, 1995.
- [6] Tze Fen Li. Multipath time delay estimation using regression stepwise procedure. *IEEE Trans. S.P.*, 46(1):191–195, January 1998.
- [7] The Numerical Algorithms Group Limited. *The Nag Fortran Library Manual, Mark 14*, 1st edition, April 1990.
- [8] J.M. Mendel. *Maximum-likelihood deconvolution: a journey into model-based signal processing*. Springer-Verlag, 1990.
- [9] V. Nimier and G. Jourdain. Active estimation of a multipath propagation channel with a bayesian strategy. *Traitement du signal*, 10(3):201–213, 1993.
- [10] M. O'Brien, A. N. Sinclair, and S. M. Kramer. Recovery of a sparse spike time series by l_1 norm deconvolution. *IEEE Trans. S.P.*, 42(12):3353–3365, Dec. 1994.
- [11] F. Poree, T. Chonavel, and T. Terre. Multipath time-delay detection and estimation for ocean acoustic tomography: a bayesian approach. In *Oceans'99 MTS/IEEE Proceedings*, volume 3, pages 1587–1590, 1999.
- [12] J. G. Proakis. *Digital Communications*. Mac Graw Hill, second edition, 1989.
- [13] C. P. Robert. *The bayesian choice*. Springer-Verlag, New-York, 1994.
- [14] H. Taylor, S. Banks, and F. McCoy. Deconvolution with the l_1 norm. *Geophys.*, 44(1):39–52, 1979.
- [15] C. Wunsch W. Munk, P. Worcester. *Tomography Acoustic Oceanic*. Cambridge University Press, 1995.