# BAYESIAN DECOMPOSITION TREES WITH APPLICATION TO SIGNAL DENOISING

*D. Leporini*
Laboratoire des Signaux et Systèmes,
CNRS/UPS and GDR ISIS, ESE, Gif-sur-Yvette Cedex, France
e-mail: leporini@lss.supelec.fr

## ABSTRACT

Tree-structured dictionaries of orthonormal bases (wavelet packet/Malvar's wavelets) provide a natural framework to answer the problem of finding a "best representation" of both deterministic and stochastic signals. In this paper, we reformulate the "best basis" search as a model selection problem and present a Bayesian approach where the decomposition operators themselves are considered as model parameters. Denoising applications are subsequently presented to substantiate the proposed methodology.

## 1 Introduction

Basing multiscale representations of signals embedded in noise on statistical approaches has recently been of great research interest [4, 2, 6]. The optimization of representation typically takes place over a tree-structured dictionary of orthonormal bases (wavelet packet/Malvar's wavelets), and aims at statistically distinguishing the signal components from those of the noise. These classical dictionaries, in turn, may be extended to more general decomposition trees by considering node-varying decomposition operators.

In this paper, we define the "best basis" search in these generalized dictionaries in a fully Bayesian perspective by considering the dictionary itself, and subsequently the optimal representation, as model stochastic parameters. Since a complete statistical description of the "best basis" is provided in this context by its posterior distribution, the objective is not necessarily to derive an estimate based on the decomposition onto the "best basis". As a consequence, this framework is particularly useful to obtain *a posteriori* mean estimates obtained by averaging signal estimates on distinctive bases. In the sequel, we introduce nonstationary decomposition trees in reference to the so-called nonstationary wavelet packets and present a Bayesian model for signal representation in denoising applications. We then propose a reversible jump Markov chain Monte Carlo (MCMC) sampler [3] to deal with the variable-dimension problem induced by the chosen non-homogeneous priors on the signal transformations.

## 2 Nonstationary decomposition trees

Let $\mathbb{R}^K$, with $K = 2^p$, $p \in \mathbb{N}^*$, denote the space of real discrete-time signals of length $K$. Discrete decompositions on the interval are used throughout the paper. We define the dictionary $\mathcal{D}$ of possible representations with the help of a finite set of decomposition operator pairs $\mathcal{S} = \{(\mathcal{F},\mathcal{G})_1, \ldots, (\mathcal{F},\mathcal{G})_N\}$ through the relations

$$\begin{aligned}\mathcal{B}_{0,0} &= \cup_{k=1}^{K}\{\delta[n-k]\}_{n=1,\ldots,K},\\ \mathcal{B}_{j+1,2m} &= \mathcal{F}^*_{j,m}\mathcal{F}_{j,m}\mathcal{B}_{j,m},\\ \mathcal{B}_{j+1,2m+1} &= \mathcal{G}^*_{j,m}\mathcal{G}_{j,m}\mathcal{B}_{j,m},\end{aligned}$$

subject to

$$\mathrm{Span}\{\mathcal{B}_{j,m}\} = \mathrm{Span}\{\mathcal{B}_{j+1,2m}\} \overset{\perp}{\oplus} \mathrm{Span}\{\mathcal{B}_{j+1,2m+1}\},$$

for all $(\mathcal{F}_{j,m},\mathcal{G}_{j,m}) \in \mathcal{S}$. We recall that $(\mathcal{F}^*_{j,m},\mathcal{G}^*_{j,m})$ corresponds to the recomposition adjoint operator pair while $\mathcal{B}_{j,m}$ denotes the orthonormal basis corresponding to the node $(j,m)$ (with $j \in \{0,\ldots,J\}$ and $m \in \{0,\ldots,2^j-1\}$) of the dictionary. In other words, these decomposition operators realize the following partition of identity

$$\mathcal{F}^*_{j,m}\mathcal{F}_{j,m} + \mathcal{G}^*_{j,m}\mathcal{G}_{j,m} = I.$$

An orthonormal basis of $\mathbb{R}^K$ is subsequently obtained according to $\mathcal{B}_{\mathcal{I}} = \cup_{(j,m)/I_{j,m}\subset\mathcal{I}}\mathcal{B}_{j,m}$ where $\mathcal{I}$ is a partition of $[0,1[$ in intervals $I_{j,m} = [2^{-j}m, 2^{-j}(m+1)[$, similarly to the wavelet packet case [10]. We point out that this general framework encompasses Malvar's wavelets, possibly nonstationary wavelet packets [1], and (by straightforward extensions) $M$-band wavelet packets [8]. We now state the problem under study. We assume the following model for the observation of a random process realization

$$y(t) = x(t) + w(t), \quad t \in \{1,\ldots,K\},$$

where $w(t)$ is i.i.d. normal, with zero mean and finite variance $\sigma^2$, although more general noise models may be adopted in some cases [5]. Recovery of the underlying *unknown* signal $x(t)$ is of interest. Given a risk to be minimized (typically the mean square error), the ideal estimation would require an oracle for the optimal decomposition basis $\mathcal{B}^*$ which is completely characterized by the corresponding sequence of optimal decomposition operators. This oracle is of particular interest for practical applications such as underwater acoustic signal processing where a wide variety of phenomena are encountered, implying that $\mathcal{B}^*$ is not *a priori* known. Then, in a Bayesian framework, the oracle for $\mathcal{B}^*$ is provided by the posterior distribution $p(\mathcal{B}^* \mid \boldsymbol{y})$ which, in turn, is used to nonlinearly estimate the underlying signal of interest. In the sequel, we take advantage of the binary tree structure of the decompositions to propose a Bayesian approach to the involved integration problem based on stochastic algorithms.

## 3 Bayesian framework

Let $\boldsymbol{x}^{\mathcal{B}^*}$ denote the $K$-dimensional vector of time samples of the underlying process $x(t)$ in $\mathcal{B}^*$. Following [6, 5], we choose non-homogeneous Bernoulli-Gaussian priors to reflect the desired parsimonious representation of the process $\boldsymbol{x}^{\mathcal{B}^*}$. Using the orthonormality property of the decompositions, we subsequently obtain

$$p(\boldsymbol{y} \mid \mathcal{B}^*_{j,m}, \boldsymbol{\theta}^*_{j,m}, \sigma^2) = \prod_{k=1}^{K2^{-j}} \Big[(1-\varepsilon_{j,m}) g(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \sigma^2) + \varepsilon_{j,m} g(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \tilde{\sigma}^2_{j,m})\Big], \quad (1)$$

where $g(\cdot \mid s^2)$ denotes the Gaussian $\mathcal{N}(0, s^2)$ PDF and $\boldsymbol{\theta}^*_{j,m} = [\varepsilon_{j,m}, \tilde{\sigma}^2_{j,m}]$, with $\tilde{\sigma}^2_{j,m} \geq \sigma^2$. In other words, the noise statistical properties remain unchanged in any basis of the dictionary. This mixture model is used in tandem with an allocation hidden vector $\boldsymbol{q}^*$ of independent random variables defining the following conditional densities

$$\begin{aligned} p(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \boldsymbol{q}^*_{j,m}[k] = 0) &= g(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \sigma^2) \\ p(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \boldsymbol{q}^*_{j,m}[k] = 1) &= g(\boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k] \mid \tilde{\sigma}^2_{j,m}), \end{aligned}$$

with $P(\boldsymbol{q}^*_{j,m}[k] = 1) = \varepsilon_{j,m} \in [0,1]$. We recall that the set of model parameters (including here $\mathcal{B}^*$) are distributed in a Bayesian framework according to prior probabilities providing the posterior distribution of interest

$$p(\mathcal{B}^*, \boldsymbol{\theta}^*, \boldsymbol{q}^* \mid \boldsymbol{y}) \propto p\big(\boldsymbol{y}^{\mathcal{B}^*} \mid \mathcal{B}^*, \boldsymbol{\theta}^*, \boldsymbol{q}^*\big) \, p\big(\boldsymbol{\theta}^*, \boldsymbol{q}^* \mid \mathcal{B}^*\big) p(\mathcal{B}^*),$$

where $\boldsymbol{\theta}^* = [\sigma^2, \cup_{(j,m)/\mathcal{B}_{j,m} \subset \mathcal{B}^*} \boldsymbol{\theta}^*_{j,m}]$. Note that our state of knowledge concerning the functional/statistical nature of the signal under study is now expressed by the likelihood function (1) and priors $p\big(\boldsymbol{\theta}^*, \boldsymbol{q}^* \mid \mathcal{B}^*\big)$ and $p(\mathcal{B}^*)$. In particular, this latter distribution expresses our degree of belief (or ignorance) concerning the optimal dictionary and the associated decomposition bases.

We further assume that the parameter vector prior reads

$$p\big(\boldsymbol{\theta}^*, \boldsymbol{q}^* \mid \mathcal{B}^*\big) = p(\sigma^2 \mid \mathcal{B}^*) \prod_{(j,m)/\mathcal{B}_{j,m} \subset \mathcal{B}^*} P(\boldsymbol{q}^*_{j,m} \mid \varepsilon_{j,m}) p(\tilde{\sigma}^2_{j,m} \mid \sigma^2),$$

to provide an independent local modeling of $\boldsymbol{y}^{\mathcal{B}^*}$. In order to minimize the mean square error, we propose to estimate the signal $\boldsymbol{x}$ using the posterior expectation $\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{y}]$ which is expressed as

$$\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{y}] = \mathbb{E}_{\mathcal{B}^*, \boldsymbol{\theta}^*, \boldsymbol{q}^*}\Big[\mathbb{E}[\boldsymbol{x} \mid \mathcal{B}^*, \boldsymbol{\theta}^*, \boldsymbol{q}^*, \boldsymbol{y}]\Big], \quad (2)$$

with

$$\mathbb{E}[\boldsymbol{x}^{\mathcal{B}^*_{j,m}}[k] \mid \boldsymbol{\theta}^*_{j,m}, \sigma^2, \boldsymbol{q}^*_{j,m}, \boldsymbol{y}^{\mathcal{B}^*_{j,m}}] = \frac{\tilde{\sigma}^2_{j,m} - \sigma^2}{\tilde{\sigma}^2_{j,m}} \boldsymbol{q}^*_{j,m}[k] \boldsymbol{y}^{\mathcal{B}^*_{j,m}}[k],$$

for the involved model (1). The evaluation of (2) being analytically intractable, we resort to reversible jump MCMC methods to answer the problem of dimension changing of the parameter space induced by the non-homogeneous likelihood.

## 4 Reversible jump MCMC sampler

We recall that MCMC algorithms allow the construction of ergodic Markov chains whose equilibrium distribution corresponds to a target posterior density (given by $p(\mathcal{B}^*, \boldsymbol{\theta}^*, \boldsymbol{q}^* \mid \boldsymbol{y})$ in our problem) upon which any Bayesian inference is based. In particular, the posterior expectation (2) may be approximated for $N \gg 1$ by

$$\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{y}] \approx \frac{1}{N - n_0} \sum_{n=n_0}^{N-1} \mathbb{E}[\boldsymbol{x} \mid \mathcal{B}^{*(n)}, \boldsymbol{\theta}^{*(n)}, \boldsymbol{q}^{*(n)}, \boldsymbol{y}],$$

where $n_0 < N$ denotes the burn-in period of the chain, under mild conditions (namely aperiodicity and irreducibility) on the generated Markov chain $\Big\{\big(\mathcal{B}^{*(n)}, \boldsymbol{\theta}^{*(n)}, \boldsymbol{q}^{*(n)}\big); n = 0, \ldots, N-1\Big\}$. Possible moves of the chain are defined through the following randomly scanned tansition kernels

$(a)$ a change in the decomposition operators $(\mathcal{F}_{j,m}, \mathcal{G}_{j,m})$ at a randomly chosen node $(j,m)$,

$(b)$ a change in the parameter vector $[\boldsymbol{\theta}^*_{j,m}, \sigma^2]$ and allocation variables $\boldsymbol{q}^*_{j,m}$ where $(j,m)$ is a terminal node,

$(c)$ a new decomposition, *i.e.* the addition of two terminal nodes,

$(d)$ a recomposition, *i.e.* the deletion of two terminal nodes.

Those transition kernels satisfy the desired (weak) convergence conditions. Note that the last two transitions induce a change in the parameter subspace dimensionality and make it compelling to resort to reversible jump samplers. The principle consists of defining here a move from $\boldsymbol{\theta}_{j,m}$ to $(\boldsymbol{\theta}_{j+1,2m}, \boldsymbol{\theta}_{j+1,2m+1})$ with the help of a reversible deterministic function $f_{j,j+1}(\cdot)$ along with a random vector $\boldsymbol{u} \in \mathbb{R}^2$ ($\boldsymbol{\theta}_{j,m} \in [0,1] \times \mathbb{R}_+$) verifying

$$(\boldsymbol{\theta}_{j+1,2m}, \boldsymbol{\theta}_{j+1,2m+1}) = f_{j,j+1}(\boldsymbol{\theta}_{j,m}, \boldsymbol{u}).$$

The random vector $\boldsymbol{u}$ completes the parameter space at resolution level $j$ in order to define a common dominating measure. We choose to generate a two-dimensional random vector of independent beta $\mathcal{B}e(3,3)$ variables to obtain the new parameters according to

$$\begin{aligned} \varepsilon_{j+1,2m} &= 2u_1\varepsilon_{j,m}, \\ \varepsilon_{j+1,2m+1} &= 2(1-u_1)\varepsilon_{j,m}, \\ \varepsilon_{j+1,2m}\tilde{\sigma}^2_{j+1,2m} &= 2u_2\varepsilon_{j,m}\tilde{\sigma}^2_{j,m}, \\ \varepsilon_{j+1,2m+1}\tilde{\sigma}^2_{j+1,2m+1} &= 2(1-u_2)\varepsilon_{j,m}\tilde{\sigma}^2_{j,m}, \end{aligned}$$

subject to the model constraints $(\varepsilon_{j+1,2m}, \varepsilon_{j+1,2m+1}) \in [0,1]^2$, $\tilde{\sigma}^2_{j+1,2m} \geq \sigma^2$ and $\tilde{\sigma}^2_{j+1,2m+1} \geq \sigma^2$. Note that this setting results in the following reconstruction equations (*i.e.* the inverse transformation $f^{-1}_{j,j+1}(\cdot)$ associated with the recomposition move $(d)$)

$$\begin{aligned} \varepsilon_{j,m} &= \frac{\varepsilon_{j+1,2m} + \varepsilon_{j+1,2m+1}}{2}, \\ \varepsilon_{j,m}\tilde{\sigma}^2_{j,m} &= \frac{\varepsilon_{j+1,2m}\tilde{\sigma}^2_{j+1,2m} + \varepsilon_{j+1,2m+1}\tilde{\sigma}^2_{j+1,2m+1}}{2}, \end{aligned}$$

which amounts to the conservation of energy for the signal component in the mixture. The Jacobian of the transformation is then given by

$$\mathcal{J} = \frac{8\varepsilon_{j,m}\tilde{\sigma}^2_{j,m}}{u_1(1-u_1)}.$$

Similarly to the classical (*i.e.* without dimension changing) Metropolis-Hastings (M-H) algorithms, the parameter vector proposal $(\boldsymbol{\theta}_{j+1,2m}, \boldsymbol{\theta}_{j+1,2m+1})$ in the decomposition move is then only accepted with probability $\alpha_{j,j+1}$ given (using simplified expressions) by

$$\alpha_{j,j+1} = \min\left\{1, \frac{p(\boldsymbol{s}' \mid \boldsymbol{y})P(\boldsymbol{q}_{j,m})}{p(\boldsymbol{s} \mid \boldsymbol{y})P\Big((\boldsymbol{q}_{j+1,2m}, \boldsymbol{q}_{j+1,2m+1})\Big)} \frac{p_{j+1,j}\mathcal{J}}{p_{\boldsymbol{U}}(\boldsymbol{u})p_{j,j+1}}\right\}.$$

In the previous expression, the symbol $\boldsymbol{s} = (\mathcal{B}_{j,m}, \boldsymbol{\theta}_{j,m}, \boldsymbol{q}_{j,m})$ stands for the current state, while $\boldsymbol{s}' = \Big((\mathcal{B}_{j+1,2m}, \mathcal{B}_{j+1,2m+1}),\ (\boldsymbol{\theta}_{j+1,2m}, \boldsymbol{\theta}_{j+1,2m+1}),\ (\boldsymbol{q}_{j+1,2m}, \boldsymbol{q}_{j+1,2m+1}))\Big)$ corresponds to the proposal, and $p_{j,j+1}$ denotes the prior probability of the decomposition move which depends on the current representation basis. The proposal for $(\boldsymbol{q}_{j+1,2m}, \boldsymbol{q}_{j+1,2m+1})$ is obtained using the full conditional distribution given the proposed parameter vector $(\boldsymbol{\theta}_{j+1,2m}, \boldsymbol{\theta}_{j+1,2m+1})$. Consequently, the chain remains in its previous state with probability $1-\alpha_{j,j+1}$, which in particular guarantees the desired aperiodicity condition. Note that our choice for moves $(c)$ and $(d)$, involving two consecutive decomposition levels, corresponds to the simplest basis update, and more general transitions may be alternatively/additionally proposed to move more rapidly across the tree. The parameter vector update (move $(b)$) is implemented using conjugate beta and inverse gamma priors for $\varepsilon_{j,m}$ and $(\tilde{\sigma}^2_{j,m}, \sigma^2)$ respectively, through the classical Data Augmentation algorithm [9, 7]

$$\begin{aligned} \sigma^{2(n+1)} &\sim p(\sigma^2 \mid \boldsymbol{q}^{*(n)}, \boldsymbol{y}^{\mathcal{B}^*}), \\ \boldsymbol{\theta}^{*(n+1)}_{j,m} &\sim p(\boldsymbol{\theta}^*_{j,m} \mid \sigma^{2(n+1)}, \boldsymbol{q}^{*(n)}_{j,m}, \boldsymbol{y}^{\mathcal{B}^*_{j,m}}), \\ \boldsymbol{q}^{*(n+1)}_{j,m} &\sim P(\boldsymbol{q}^*_{j,m} \mid \boldsymbol{\theta}^{*(n+1)}_{j,m}, \sigma^{2(n+1)}, \boldsymbol{y}^{\mathcal{B}^*_{j,m}}). \end{aligned}$$

We finally choose to implement an M-H step for move $(a)$ by first selecting a node $(j,m)$ at random in the current decomposition tree, and then drawing a proposal $(\mathcal{F}_{j,m}, \mathcal{G}_{j,m})$ from the prior distribution on the operator set $\mathcal{S}$. Without additional information, this distribution is considered as uniform. This operator pair provides in turn a new decomposition basis and is likely to modify the representation at children nodes $(j', m')$, with $j' > j$ such that $\mathcal{B}_{j',m'}$ belongs to the current basis. We therefore simultaneously update the associated parameter vector $\boldsymbol{\theta}_{j',m'}$ with proposals drawn from their prior distribution, while $\boldsymbol{q}_{j',m'}$ is again obtained using the full conditional distribution given the proposed parameters.

## 5 Simulations

To show the interest of the proposed method, we present the results obtained with two examples of underwater acoustic signals given in Fig. 1. The first process corresponds to a biological signal involving transient phenomena while the second one is artificial and corresponds to a modulated waveform. In both cases, the noise level results in a Signal-to-Noise Ratio of 0 dB. The set of decomposition operators is composed of three unitary transforms given by a single level wavelet packet decomposition, a time segmentation and a discrete cosine transform, providing enough structure to reproduce both wavelet packet (WP) and Malvar's wavelet (MW) decompositions. The maximum level of decomposition was fixed to $J = 5$ and the parameter informative priors are given by $\varepsilon_{j,m} \sim \mathcal{B}e(1,3)$, $\tilde{\sigma}^2_{j,m} \sim \mathcal{IG}(1, 1/2\widehat{\sigma}^2_{\boldsymbol{y}^{\mathcal{B}}})$ and $\sigma^2 \sim \mathcal{IG}(3, 1/4\widehat{\sigma}^2_{\boldsymbol{y}^{\mathcal{B}}})$ subject to $\tilde{\sigma}^2_{j,m} \geq \sigma^2$. The prior distribution on $\varepsilon_{j,m}$ expresses the desired parsimonious character of signal representations, while the mode of the variance priors corresponds to a robust estimate of the variance of $\boldsymbol{y}$ in the current decomposition basis denoted by $\widehat{\sigma}^2_{\boldsymbol{y}^{\mathcal{B}}}$. Indeed this estimate is close to $\sigma^2$ when the decomposition onto the basis $\mathcal{B}$ leads to a parsimonious representation of the signal of interest. Our approach is then compared in terms of normalized mean square error (NMSE $= ||\widehat{\boldsymbol{x}} - \boldsymbol{x}||_2^2/||\boldsymbol{x}||_2^2$) to the two corresponding best basis selection algorithms introduced in [6] using the same model (1). Note that, in this latter work, a maximum likelihood/generalized likelihood approach was used to determine the "best basis". This approach was shown to provide improved performances with respect to classical thresholding policies.

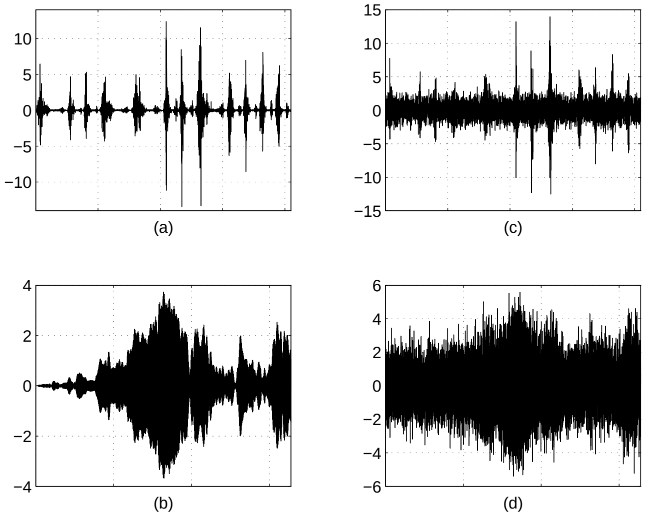


Figure 1: (a) biological signal, (b) artificial signal, (c) and (d) noisy versions of (a) and (b) respectively.

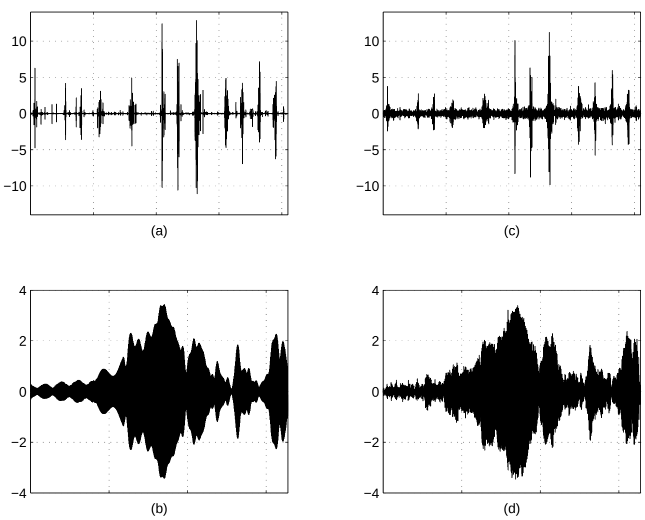


Figure 2: (a) and (c) biological signal estimates using WP (NMSE = 0.35) and Posterior Expectation (NMSE = 0.29) respectively, (b) and (d) artificial signal estimates using MW (NMSE = $4.6\ \ 10^{-2}$) and Posterior Expectation (NMSE = $3.5\ \ 10^{-2}$) respectively.

For illustration, the results obtained with the proposed algorithm are presented in Fig. 2 along with the estimates derived in the best (fixed) dictionary using the approach developed in [6]. As expected, our approach demonstrates its adaptation properties to the *unknown* signal of interest.

## 6 Conclusion

In this paper, a fully Bayesian approach to "best basis" representation of noisy signals over tree-structured dictionaries of bases has been presented. This approach makes use of non-homogeneous statistical models, and hinges upon the construction of a Markov chain whose stationary distribution corresponds to the posterior distribution of interest. This Markov chain, in turn, is used to nonlinearly estimate the underlying signal via posterior expectation.


## Acknowledgement

The author wishes to thank Mr. Di Martino (DCN/CERDSM) for providing the underwater acoustic data.


## References


[1] A. Cohen. Non-stationary multiscale analysis. In C. K. Chui, L. Montefusco, and L. Puccio, editors, *Wavelets. Theory, algorithms and applications.* Academic Press, San Diego, CA, 1994.

[2] D. L. Donoho and I. M. Johnstone. Ideal denoising in an orthogonal basis chosen from a library of bases. *C. R. Acad. Sci. Paris*, 319:1317–1322, 1994.

[3] P. J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82:711–732, 1995.

[4] H. Krim and J.-C. Pesquet. On the statistics of best bases criteria. In A. Antoniadis, editor, *Wavelets and statistics.* Lecture Notes in Statistics, Springer Verlag, 1995.

[5] D. Leporini. Méthodes MCMC pour la Décomposition en Paquets d'Ondelettes de Signaux Transitoires. In *Seizième Colloque GRETSI (Grenoble, FRANCE)*, pages 1455–1458, 1997.

[6] J.-C. Pesquet, H. Krim, D. Leporini, and E. Hamman. Bayesian Approach to Best Basis Selection. In *Proc. IEEE Conf. Acoust., Speech, Signal Processing*, pages 2634–2638, Atlanta, USA, May 7-9 1996.

[7] C. P. Robert. *Mixtures of distributions: inference and estimation.* Chapman and Hall, in *Practical Markov chain Monte Carlo*, pp. 441-464, London, 1996.

[8] P. Stephen, P. N. Heller, R. A. Gospinath, and G. B. Burrus. Theory of regular $M$-band wavelet bases. *IEEE Trans. Signal Processing*, SP-41:3497–3511, Dec. 1993.

[9] M. Tanner and W. Wong. The calculation of posterior distributions. *J.Amer. Statist. Assoc.*, 82:528–550, 1987.

[10] M. V. Wickerhauser. INRIA lectures on wavelet packet algorithms. In *Ondelettes et paquets d'ondelettes*, pages 31–99, Roquencourt, France, Jun. 17-21 1991.