BAYESIAN SPARSE IMAGE RECONSTRUCTION FOR MRFM

Nicolas Dobigeon, Alfred O. Hero[*]

University of Michigan
Department of EECS
Ann Arbor, MI 48109-2122, USA
`{dobigeon,hero}@umich.edu`

Jean-Yves Tourneret

University of Toulouse
IRIT/INP-ENSEEIHT
31071 Toulouse cedex 7, France
`jean-yves.tourneret@enseeiht.fr`

ABSTRACT

In this paper, we propose a Bayesian model and a Monte Carlo Markov chain (MCMC) algorithm for reconstructing images that consist of only few non-zero pixels. An appropriate distribution that promotes sparsity is proposed as prior distribution for the pixel values. The hyperparameters involved in the modeling are also assigned prior distributions, resulting in a hierarchical model. A Gibbs sampler allows us to draw samples distributed according the full posterior of interest. These samples are then used to approximate standard *maximum a posteriori* (MAP) estimator. By conducting some simulations, we show that the proposed estimator clearly outperforms previous estimators proposed in the literature.

Index Terms— Bayesian inference, sparse representation, MCMC methods, MRFM.

1. INTRODUCTION

A few years ago, a new technology for molecular imagery was proposed [1]. This technique, called Magnetic Resonance Force Microscopy (MRFM), is able to detect molecules with atomic-level resolution. Thanks to the nanometer scale of the sensed scenes, most part of the image is empty: only a few pixels have non-zero values. This paper proposes to reconstruct such sparse images in a full Bayesian framework when the observations are obtained from noisy linear transformations.

Reconstruction of sparse images provided by MRFM has been recently addressed in [2, 3]. In these works, sparsity penalized reconstruction methods based on a LAZE (Laplacian distribution and an atom at zero) prior have been proposed. Then, the reconstruction problem is decomposed into a deconvolution step and a denoising step, yielding an iterative thresholding algorithm. However, in [2], the hyperparameters inherent to the prior model are estimated in an empirical manner. This *ad hoc* hyperparameter choice may lead to incorrect results, particularly at high SNRs for which these methods display increasingly biased estimation of the hyperparameters. We propose here to estimate the unknown parameters and hyperparameters in a fully Bayesian framework.

In this paper, a new prior composed of a mass at zero and a single-sided exponential distribution is introduced. Coupling a standard probability density function (pdf) with an atom at zero is a standard choice to ensure sparsity. This strategy has for instance been used for the detection of located events required in several applications such as spike train deconvolution [4, 5]. In addition, this

prior allows one to take into account the positivity of the pixels, contrary to [2]. When the noise variance is unknown, the full Bayesian posterior can be derived from samples generated by Markov chain Monte Carlo (MCMC) methods. As suggested above, the results of the sparse reconstruction critically depend on the hyperparameters chosen to define the mixture for the image prior. In the Bayesian estimation framework, two different approaches can be considered to estimate these hyperparameters. One approach couples MCMC methods to an expectation-maximization (EM) algorithm [6]. The other approach introduces a second level of hierarchy in the Bayesian formulation by choosing non-informative prior distributions for the hyperparameters. This fully Bayesian approach, adopted in this paper, has been successfully applied to signal segmentation [7] and semi-supervised unmixing of hyperspectral imagery [8].

The paper is organized as follows. The problem and the different notations are introduced in Section 2. Section 3 presents the hierarchical Bayesian model used for sparse image reconstruction. The Gibbs sampling strategy used to sample according to the posterior of interest is detailed in Section 4. Some simulation results are presented in Section 5. Conclusions are reported in Section 6.

2. PROBLEM STATEMENT

Let $\mathbf{X}$ denote the $n_r \times n_c$ unknown 2-dimensional image to be estimated. The observed $n_r \times n_c$ matrix $\mathbf{Y}$ is the following noisy convolution:

$$\mathbf{Y} = \boldsymbol{\kappa} \otimes \mathbf{X} + \mathbf{N}, \tag{1}$$

where $\otimes$ stands for the 2-dimensional convolution operator, $\mathbf{N}$ is an $n_r \times n_c$ observation noise and $\boldsymbol{\kappa}$ is the kernel modeling the response of the imaging device. Typical responses of MRFM tip $\boldsymbol{\kappa}$ can be found in [9] for horizontal and vertical configurations.

The images $\mathbf{X}$ and $\mathbf{Y}$ can be lexicographically vectorized yielding the unknown vectorized image $\mathbf{x} \in \mathbb{R}^M$ and the observed measurement vector $\mathbf{y} \in \mathbb{R}^M$, respectively, with $M = n_r n_c$. With these notations, Eq. (1) can be rewritten as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{H}$ is a $M \times M$ Toeplitz matrix that describes the convolution by the point spread function (psf) $\boldsymbol{\kappa}$. In (2), $\mathbf{n}$ is an additive Gaussian noise sequence distributed according to $\mathbf{n} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}_M\right)$. Note that the noise variance σ^2 is unknown. The problem addressed in the following sections consists of estimating $\mathbf{x}$ and the noise variance σ^2 by ensuring sparsity and positivity constraints for the vector $\mathbf{x}$ given the observations $\mathbf{y}$ and the transformation matrix $\mathbf{H}$.

3. HIERARCHICAL BAYESIAN MODEL

3.1. Likelihood function

The observation model defined in (2) and the Gaussian properties of the noise sequence $\mathbf{n}$ yield:

$$f\left(\mathbf{y}|\mathbf{x},\sigma^2\right) = \left(\frac{1}{2\pi\sigma^2}\right)^M \exp\left(-\frac{\|\mathbf{y}-\mathbf{Hx}\|^2}{2\sigma^2}\right), \qquad (3)$$

where $\|\cdot\|^2$ denotes the standard ℓ_2 norm: $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$.

3.2. Parameter priors

3.2.1. Image prior

As prior distribution for x_i, we propose here to use a mixture of a mass at zero and a single-sided exponential distribution:

$$f\left(x_i|w,a\right) = (1-w)\delta\left(x_i\right) + \frac{w}{a}\exp\left(-\frac{x_i}{a}\right)\mathbf{1}_{\mathbb{R}_+^*}\left(x_i\right), \quad (4)$$

where $\delta\left(\cdot\right)$ is the Dirac function and $\mathbf{1}_{\mathbb{E}}\left(x\right)$ is the indicator function defined on $\mathbb{E}$:

$$\mathbf{1}_{\mathbb{E}}\left(x\right) = \left\{\begin{array}{ll} 1, & \text{if } x \in \mathbb{E}, \\ 0, & \text{otherwise.} \end{array}\right. \qquad (5)$$

By assuming the components x_i to be a priori independent ($i = 1,\ldots,M$), the following prior distribution is obtained for $\mathbf{x}$:

$$f\left(\mathbf{x}|w,a\right) = \prod_{i=1}^{M}\left[(1-w)\delta\left(x_i\right) + \frac{w}{a}\exp\left(-\frac{x_i}{a}\right)\mathbf{1}_{\mathbb{R}_+^*}\left(x_i\right)\right]. \tag{6}$$

This prior is similar to the LAZE distribution (weighted average of a Laplacian pdf and an atom at zero) used, for example, in [2]. However, the proposed prior in (4) also ensures the positivity of the pixel values to be estimated.

3.2.2. Noise variance prior

Classically, a conjugate inverse-Gamma distribution with parameters $\frac{\nu}{2}$ and $\frac{\gamma}{2}$ is chosen as prior distribution for the noise variance:

$$\sigma^2|\nu,\gamma \sim \mathcal{IG}\left(\frac{\nu}{2},\frac{\gamma}{2}\right). \qquad (7)$$

In the following, ν will be fixed to $\nu = 2$ and γ will be an hyperparameter that will be estimated. This choice significantly reduces the complexity of the Bayesian algorithm. It has been suggested in many works including [7, 8].

3.3. Hyperparameter priors

The hyperparameter vector associated with the previous prior distributions is $\boldsymbol{\Phi} = \{a,\gamma,w\}$. Obviously, the accurateness of the proposed Bayesian model for image reconstruction depends on the values of these hyperparameters. We propose to estimate these hyperparameters from the data using a hierarchical Bayesian model. This model requires to choose prior distributions, detailed below, for the hyperparameters.

3.3.1. Hyperparameter a

A conjugate inverse-Gamma distribution is assumed for hyperparameter a:

$$a|\boldsymbol{\alpha} \sim \mathcal{IG}\left(\alpha_0,\alpha_1\right), \qquad (8)$$

with $\boldsymbol{\alpha} = [\alpha_0,\alpha_1]^T$. The fixed hyperparameters α_0 and α_1 have been chosen to obtain a vague prior, i.e. $\alpha_0 = \alpha_1 = 10^{-10}$.

3.3.2. Hyperparameter γ

Non informative Jeffrey's prior is chosen as prior distribution for hyperparameter γ:

$$f\left(\gamma\right) \propto \frac{1}{\gamma}\mathbf{1}_{\mathbb{R}_+}\left(\gamma\right). \qquad (9)$$

3.3.3. Hyperparameter w

A conjugate beta distribution with fixed hyperparameters β_1 and β_0 is chosen as prior distribution for w:

$$w|\boldsymbol{\beta} \sim \mathcal{B}\left(\beta_1,\beta_0\right), \qquad (10)$$

with $\boldsymbol{\beta} = [\beta_0,\beta_1]^T$ and where $\mathcal{B}\left(a,b\right)$ denotes the Beta distribution with parameters (a,b). By choosing $\beta_0 = \beta_1 = 1$, the Beta distribution reduces to the uniform distribution on $[0,1]$.

Assuming that the individual hyperparameters are independent the full hyperparameter prior distribution for $\boldsymbol{\Phi}$ can be expressed as:

$$f\left(\boldsymbol{\Phi}|\boldsymbol{\alpha},\boldsymbol{\beta}\right) = f\left(a\right)f\left(\gamma\right)f\left(w\right). \qquad (11)$$

3.4. Posterior distribution

The posterior distribution of $\{\boldsymbol{\theta},\boldsymbol{\Phi}\}$ can be computed as follows:

$$f\left(\boldsymbol{\theta},\boldsymbol{\Phi}|\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right) \propto f\left(\mathbf{y}|\boldsymbol{\theta}\right)f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right)f\left(\boldsymbol{\Phi}|\boldsymbol{\alpha},\boldsymbol{\beta}\right), \qquad (12)$$

with

$$f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right) = f\left(\mathbf{x}|a,w\right)f\left(\sigma^2|\gamma\right), \qquad (13)$$

where $f\left(\mathbf{y}|\boldsymbol{\theta}\right)$ and $f\left(\boldsymbol{\Phi}|\boldsymbol{\alpha},\boldsymbol{\beta}\right)$ have been defined in (3) and (11). This hierarchical structure allows one to integrate out the parameter σ^2 and the hyperparameter vector $\boldsymbol{\Phi}$ in the full posterior distribution (12), yielding:

$$f\left(\mathbf{x}|\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right) \propto \frac{B\left(\beta_1+n_1,\beta_0+n_0\right)}{\|\mathbf{y}-\mathbf{Hx}\|^M}\frac{\Gamma\left(n_1+\alpha_0\right)}{\left(\|\mathbf{x}\|_1+\alpha_1\right)^{n_1+\alpha_0}}. \tag{14}$$

with $B\left(\beta_1,\beta_0\right) = \Gamma\left(\beta_1\right)\Gamma\left(\beta_0\right)/\Gamma\left(\beta_1+\beta_0\right)$, where $\Gamma\left(\cdot\right)$ denotes the Gamma function, and $n_1 = \|\mathbf{x}\|_0$ and $n_0 = M - \|\mathbf{x}\|_0$.

The next section presents an appropriate Gibbs sampling strategy that allows ones to generate samples distributed according to the posterior distribution $f\left(\mathbf{x}|\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right)$.

4. GIBBS SAMPLER

The principles of the Gibbs sampler can be found in many textbooks, including [10].

4.1. Generation of samples according to $f\left(\mathbf{x}\,|\sigma^2,\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right)$

To generate samples distributed according to $f\left(\mathbf{x}\,|\sigma^2,\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right)$, it is very convenient to sample according to $f\left(\mathbf{x},w,a\,|\sigma^2,\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right)$. Thus, generating samples according to $f\left(\mathbf{x}\,|\sigma^2,\mathbf{y},\boldsymbol{\alpha},\boldsymbol{\beta}\right)$ can be achieved in the 3-step procedure detailed in paragraphs 4.1.1, 4.1.2 and 4.1.3.

4.1.1. Generation of samples according to $f\left(w\,|\mathbf{x},\boldsymbol{\beta}\right)$

Straightforward computations lead to:

$$f\left(w\,|\mathbf{x},\boldsymbol{\beta}\right) \propto (1-w)^{n_0+\beta_0-1}w^{n_1+\beta_1-1}, \qquad (15)$$

where n_0 and n_1 have been defined in the previous section. Therefore, generation of samples according to $f\left(w\,|\mathbf{x},\boldsymbol{\beta}\right)$ is achieved as follows:

$$w\,|\mathbf{x},\boldsymbol{\beta} \sim \mathcal{B}\left(\beta_1+n_1,\beta_0+n_0\right). \qquad (16)$$

4.1.2. Generation of samples according to $f\left(a\,|\mathbf{x},\boldsymbol{\alpha}\right)$

Looking at the joint posterior distribution (12), we obtain:

$$a\,|\mathbf{x},\boldsymbol{\alpha} \sim \mathcal{IG}\left(\|\mathbf{x}\|_0+\alpha_0,\|\mathbf{x}\|_1+\alpha_1\right). \qquad (17)$$

Table 1. Parameters used to compute the MRFM psf.

Parameter		Value
Description	Name	
Amplitude of external magnetic field	B_{ext}	9.6×10^3 G
Value of B_{mag} in the resonant slice	B_{res}	1.0×10^4 G
Radius of tip	R_0	4.0 nm
Distance from tip to sample	d	7.6 nm
Cantilever tip moment	m	4.6×10^5 emu
Peak cantilever oscillation oscillation	x_{pk}	0.8 nm
Maximum magnetic field gradient	G_{max}	125

Fig. 1. Left: Psf of the MRFM tip. Right: unknown sparse image to be estimated.

4.1.3. Generation of samples according to $f\left(\mathbf{x}\,|w,a,\sigma^2,\mathbf{y}\right)$

The posterior distribution of each component x_i ($i = 1, \ldots, M$) conditionally upon the others can be easily derived:

$$f\left(x_i|w,a,\sigma^2,\mathbf{x}_{-i},\mathbf{y}\right) \propto (1-w_i)\delta\left(x_i\right) + w_i\phi_+\left(x_i|\mu_i,\eta_i^2\right),\tag{18}$$

with

$$\begin{cases} \mathbf{e}_i = \mathbf{y} - \sum_{k\neq i} x_k\mathbf{h}_i, \\[2mm] \eta_i^2 = \dfrac{\sigma^2}{\|\mathbf{h}_i\|^2}, \\[2mm] \mu_i = \eta_i^2\left(\dfrac{\mathbf{h}_i^T\mathbf{e}_i}{\sigma^2} - \dfrac{1}{a}\right), \\[2mm] u_i = \dfrac{w}{a}C\left(\mu_i,\eta_i^2\right)\exp\left(\dfrac{\mu_i^2}{2\eta_i^2}\right), \\[2mm] w_i = \dfrac{u_i}{u_i + (1-w)}, \end{cases}\tag{19}$$

where $\mathbf{x}_{-i}$ stands for the vector $\mathbf{x}$ whose ith component has been removed and $\mathbf{h}_i$ denotes the ith column of $\mathbf{H}$. In (18), $\phi_+\left(\cdot, m, s^2\right)$ stands for the pdf of the truncated Gaussian distribution defined on $\mathbb{R}_+^*$ with hidden parameters equal to the mean m and the variance s^2. Therefore, from (18), $x_i|w,a,\sigma^2,\mathbf{x}_{-i},\mathbf{y}$ is a Bernoulli-truncated Gaussian variable with parameter $\left(w_i,\mu_i,\eta_i^2\right)$. Generation of samples distributed according to the Bernoulli-truncated Gaussian can be easily performed by using a Bernoulli auxiliary variable and the simulation scheme studied in [11].

4.2. Generation of samples according to $f\left(\sigma^2\,|\mathbf{x},\mathbf{y}\right)$

The posterior distribution of $\sigma^2\,|\mathbf{x},\mathbf{y}$ is

$$\sigma^2\,|\mathbf{x},\mathbf{y} \sim \mathcal{IG}\left(\frac{M}{2}, \frac{\|\mathbf{y} - \mathbf{Hx}\|^2}{2}\right).\tag{20}$$

according to which it is easy to sample.

5. SIMULATIONS

This section considers a 32×32 synthetic image simulated using the prior in (4) with parameter $a = 1$ and $w = 0.02$ (see Fig. 1 (right)). In this figure and in the following ones, white pixels stand for zero values. Following the MRFM tip model in [9], a 10×10 2-dimensional convolution kernel, represented in Fig. 1 (left), has been generated using the physical parameters given in Table 1 (see also [3] for more details). The corresponding matrix $\mathbf{H}$ is of size 1024×1024. The observed measurements $\mathbf{y}$, depicted
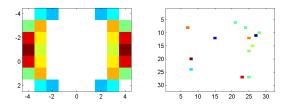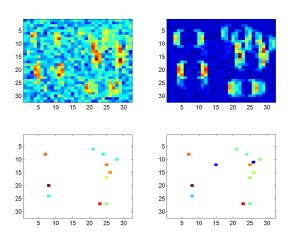
in Fig. 1 (right) are of size $M = 1024$. These observations are corrupted by an additive Gaussian noise with two different variances $\sigma^2 = 2.0 \times 10^{-1}$ and $\sigma^2 = 3.1 \times 10^{-3}$, corresponding to signal-to-noise ratios SNR = 2dB and SNR = 20dB respectively. The observations are processed by the proposed algorithm that consists of $N_{\text{MC}} = 2000$ iterations of the Gibbs sampler with $N_{\text{bi}} = 300$ burn-in iterations. Then the MAP estimator of the unknown image $\mathbf{x}$ is computed by keeping the generated sample that maximizes the posterior distribution in (14):

$$\hat{\mathbf{x}}_{\text{MAP}} \approx \underset{\mathbf{x}\in\mathcal{X}}{\arg\max}\, f\left(\mathbf{x}|\mathbf{y}\right),\tag{21}$$

with $\mathcal{X} = \left\{\mathbf{x}^{(t)}, t = 1, \ldots, N_{\text{MC}}\right\}$. These estimates are depicted in Fig. 2 for the two levels of noise considered. The estimated images are very similar to the actual image, even with low SNR.
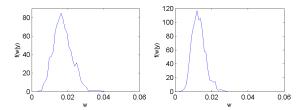

Fig. 2. Top, left (resp. right): noisy observations for SNR = 2dB (resp. 20dB). Bottom, left (resp. right): reconstructed image for SNR = 2dB (resp. 20dB).

As the proposed algorithm generates samples distributed according to the posterior distribution in (14), these samples can be used to compute the posterior distributions of each parameter. As examples, the posterior distributions of the hyperparameter w and the noise variance σ^2, are shown in figures 3 and 4. These estimated distributions are in good agreement with the actual values of these parameters[1] for the two SNR levels. The posterior distributions of two

---

[1]Similar results have been obtained for the hyperparameter a. They are omitted for brevity.

different pixels are finally depicted in Fig. 5. These posteriors are also in agreement with the actual values of these pixels that are represented in dotted red line in these figures.

Fig. 3. Posterior distribution of hyperparameter w (left: $\mathrm{SNR} = 2\mathrm{dB}$, right: $\mathrm{SNR} = 20\mathrm{dB}$).
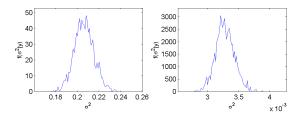
The results provided by the proposed method have been compared with those provided by methods that also estimate the hyperparameters. Firstly, the techniques proposed in [2, 3] are based on EM algorithms that perform empirical estimation of the unknown hyperparameters. Therein, two empirical Bayesian estimators, denoted Emp-MAP-Lap and Emp-MAP-LAZE, based on a Laplacian or a LAZE prior respectively, are studied. Here we compare the estimators of [2,3] to the proposed MAP estimator under the model and the algorithm presented in Sections 3 and 4. As in [2,3], the comparison is conducted with respect to different measures of performance. First let $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ denote the reconstruction error when $\hat{\mathbf{x}}$ is the estimator of the image $\mathbf{x}$ to be recovered. To measure the performance of the sparse reconstruction, four criteria have been used: the ℓ_0, ℓ_1 and ℓ_2-norms of $\mathbf{e}$ to measure the accuracy of the reconstruction and the ℓ_0-norm of the estimator $\hat{\mathbf{x}}$ to measure its sparsity. Table 2 gathers the four performance measures for the four different studied algorithms. It appears that the proposed Bayesian MAP estimator yiels better performance criteria than the others estimators.

Table 2. Reconstruction performances for different algorithms.

Method	Error criterion			
	$\|e\|_0$	$\|e\|_1$	$\|e\|_2$	$\|\hat{\mathbf{x}}\|_0$
SNR = 2dB				
Landweber	1024	372.96	14.50	1024
Emp-MAP-Lap	13	15.70	4.87	0
Emp-MAP-LAZE	48	10.63	1.80	46
Proposed MAP	13	1.65	0.57	11
SNR = 20dB				
Landweber	1024	178.51	7.01	1024
Emp-MAP-Lap	29	1.62	0.42	27
Emp-MAP-LAZE	54	2.04	0.35	53
Proposed MAP	14	0.34	0.13	13

6. CONCLUSIONS

We presented a novel Bayesian estimation algorithm to reconstruct sparse images provided by MRFM. A mixture of a one-sided exponential distribution and a mass at zero was chosen as prior distribution for the pixel values to ensure their non-negativity and the sparsity of the estimated image. The hyperparameters involved in this mixture were directly estimated from the data via a hierarchical

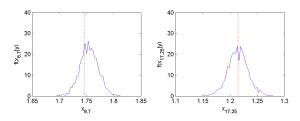**Fig. 4**. Posterior distribution of hyperparameter σ^2 (left: $\mathrm{SNR} = 2\mathrm{dB}$, right: $\mathrm{SNR} = 20\mathrm{dB}$).

Fig. 5. Posteriors distributions of the non-zero values of $\mathbf{x}$ for $\mathrm{SNR} = 20\mathrm{dB}$, (actual values are depicted with dotted red lines).

model, yielding a fully Bayesian approach. A Gibbs sampling strategy was proposed approximate the MAP estimate of the unknown image. The derived Bayesian estimate provided very encouraging results when compared to other previously proposed estimators.

7. REFERENCES

[1] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui, "Single spin detection by magnetic resonance force microscopy," *Nature*, vol. 430, pp. 329–332, July 2004.

[2] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction using sparse priors," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Oct. 2006, pp. 1261–1264.

[3] M. Ting, "Signal processing for magnetic resonance force microscopy," Ph.D. dissertation, Univ. of Michigan, Ann Arbor, MI, May 2006.

[4] J. J. Kormylo and J. M. Mendel, "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 482–488, May 1982.

[5] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 2988–2998, Dec. 1996.

[6] M. Lavielle and E. Lebarbier, "An application of MCMC methods for the multiple change-points problem," *Signal Processing*, vol. 81, no. 1, pp. 39–53, Jan. 2004.

[7] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Joint segmentation of multivariate astronomical time series: Bayesian sampling with a hierarchical model," *IEEE Trans. Signal Processing*, vol. 55, no. 2, pp. 414–423, Feb. 2007.

[8] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2684–2695, July 2008.

[9] J. Mamin, R. Budakian, and D. Rugar, "Point response function of an MRFM tip," IBM Research Division, Tech. Rep., Oct. 2003.

[10] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[11] C. P. Robert, "Simulation of truncated normal variables," *Statistics and Computing*, vol. 5, no. 2, pp. 121–125, June 1995.