Hierarchical Bayesian sparse image reconstruction with application to MRFM

Nicolas Dobigeon, Jean-Yves Tourneret, Alfred O. Hero

Hierarchical Bayesian Sparse Image Reconstruction With Application to MRFM

Nicolas Dobigeon, *Member, IEEE*, Alfred O. Hero, *Fellow, IEEE*, and Jean-Yves Tourneret, *Senior Member, IEEE*

***Abstract*—This paper presents a hierarchical Bayesian model to reconstruct sparse images when the observations are obtained from linear transformations and corrupted by an additive white Gaussian noise. Our hierarchical Bayes model is well suited to such naturally sparse image applications as it seamlessly accounts for properties such as sparsity and positivity of the image via appropriate Bayes priors. We propose a prior that is based on a weighted mixture of a positive exponential distribution and a mass at zero. The prior has hyperparameters that are tuned automatically by marginalization over the hierarchical Bayesian model. To overcome the complexity of the posterior distribution, a Gibbs sampling strategy is proposed. The Gibbs samples can be used to estimate the image to be recovered, e.g., by maximizing the estimated posterior distribution. In our fully Bayesian approach, the posteriors of all the parameters are available. Thus, our algorithm provides more information than other previously proposed sparse reconstruction methods that only give a point estimate. The performance of the proposed hierarchical Bayesian sparse reconstruction method is illustrated on synthetic data and real data collected from a tobacco virus sample using a prototype MRFM instrument.**

***Index Terms*—Bayesian inference, deconvolution, Markov chain Monte Carlo (MCMC) methods, magnetic resonance force microscopy (MRFM) imaging, sparse representation.**

I. Introduction

FOR several decades, image deconvolution has been of increasing interest [2], [47]. Image deconvolution is a method for reconstructing images from observations provided by optical or other devices and may include denoising, deblurring or restoration. The applications are numerous including astronomy [49], medical imagery [48], remote sensing [41] and photography [55]. More recently, a new imaging technology, called magnetic resonance force microscopy (MRFM), has been developed (see [38] and [29] for reviews). This nondestructive method allows one to improve the detection sensitivity of standard magnetic resonance imaging (MRI) [46]. Three-dimensional MRI at 4 nm spatial resolution has recently been achieved by the IBM MRFM prototype for imaging the proton density of a tobacco virus [8]. Because of its potential atomic-level resolution,[1] the 2-D or 3-D images resulting from this technology are naturally sparse in the standard pixel basis. Indeed, as the observed objects are molecules, most of the image is empty space. In this paper, a hierarchical Bayesian model is proposed to perform reconstruction of such images.

Sparse signal and image deconvolution has motivated research in many scientific applications including: spectral analysis in astronomy [4]; seismic signal analysis in geophysics [7], [45]; and deconvolution of ultrasonic B-scans [39]. We propose here a hierarchical Bayesian model that is based on selecting an appropriate prior distribution for the unknown image and other unknown parameters. The image prior is composed of a weighted mixture of a standard exponential distribution and a mass at zero. When the nonzero part of this prior is chosen to be a centered normal distribution, this prior reduces to a Bernoulli-Gaussian process. This distribution has been widely used in the literature to build Bayesian estimators for sparse deconvolution problems (see [5], [16], [24], [28], [33], or more recently [3] and [17]). However, choosing a distribution with heavier tail may improve the sparsity inducement of the prior. Combining a Laplacian distribution with an atom at zero results in the so-called LAZE prior. This distribution has been used in [27] to solve a general denoising problem in a non-Bayesian quasi-maximum likelihood estimation framework. In [52] and [54], this prior has also been used for sparse reconstruction of noisy images, including MRFM. The principal weakness of these previous approaches is the sensitivity to hyperparameters that determine the prior distribution, e.g., the LAZE mixture coefficient and the weighting of the prior versus the likelihood function. The hierarchical Bayesian approach proposed in this paper circumvents these difficulties. Specifically, a new prior composed of a mass at zero and a single-sided exponential distribution is introduced, which accounts for positivity and sparsity of the pixels in the image. Conjugate priors on the hyperparameters of the image prior are introduced. It is this step that makes our approach hierarchical Bayesian. The full Bayesian posterior can then be derived from samples generated by Markov chain Monte Carlo (MCMC) methods [44].

The estimation of hyperparameters involved in the prior distribution described above is the most difficult task and poor estimation leads to instability. Empirical Bayes (EB) and Stein unbiased risk (SURE) solutions were proposed in [52], [54] to

N. Dobigeon was with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA. He is now with the University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse cedex 7, France (e-mail: nicolas.dobigeon@enseeiht.fr).

A. O. Hero is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: hero@umich.edu).

J.-Y. Tourneret is with the University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse cedex 7, France (e-mail: jean-yves.tourneret@enseeiht.fr).

[1]Note that the current state of art of the MRFM technology allows one to acquire images with nanoscale resolution. However, atomic-level resolution might be obtained in the future.

deal with this issue. However, instability was observed especially at higher signal-to-noise ratios (SNR). In the Bayesian estimation framework, two approaches are available to estimate these hyperparameters. One approach couples MCMC methods to an expectation-maximization (EM) algorithm or to a stochastic EM algorithm [30], [32] to maximize a penalized likelihood function. The second approach defines noninformative prior distributions for the hyperparameters; introducing a second level of hierarchy into the Bayesian formulation. This latter fully Bayesian approach, adopted in this paper, has been successfully applied to signal segmentation [11], [14], [15] and semi-supervised unmixing of hyperspectral imagery [13].

Only a few papers have been published on reconstruction of images from MRFM data [6], [8], [56], [58]. In [21], several techniques based on linear filtering and maximum-likelihood principles were proposed that do not exploit image sparsity. More recently, Ting *et al.* has introduced sparsity penalized reconstruction methods for MRFM (see [54] or [53]). The reconstruction problem has been formulated as a decomposition into a deconvolution step and a denoising step, yielding an iterative thresholding framework. In [54] the hyperparameters are estimated using penalized log-likelihood criteria including the SURE approach [50]. Despite promising results, especially at low SNR, penalized likelihood approaches require iterative maximization algorithms that are often slow to converge and can get stuck on local maxima [10]. In contrast to [54], the fully Bayesian approach presented in this paper converges quickly and produces estimates of the entire posterior and not just the local maxima. Indeed, the hierarchical Bayesian formulation proposed here generates Bayes-optimal estimates of all image parameters, including the hyperparameters.

In this paper, the response of the MRFM imaging device is assumed to be known. While it may be possible to extend our methods to unknown point spread functions, e.g., along the lines of [22] and [23], the case of sparse blind deconvolution is outside of the scope of this paper.

This paper is organized as follows. The deconvolution problem is formulated in Section II. The hierarchical Bayesian model is described in Section III. Section IV presents a Gibbs sampler that allows one to generate samples distributed according to the posterior of interest. Simulation results, including extensive performance comparison, are presented in Section V. In Section VI, we apply our hierarchical Bayesian method to reconstruction of a tobacco virus from real MRFM data. Our main conclusions are reported in Section VII.

II. Problem Formulation

Let $\mathbf{X}$ denote a $l_1 \times \cdots \times l_n$ unknown n-dimensional pixelated image to be recovered (e.g., $n = 2$ or $n = 3$). Observed are a collection of P projections $\mathbf{y} = [y_1, \ldots, y_P]^T$ which are assumed to follow the model

$$\mathbf{y} = T(\boldsymbol{\kappa}, \mathbf{X}) + \mathbf{n} \tag{1}$$

where $T(\cdot,\cdot)$ stands for a bilinear function, $\mathbf{n}$ is a $P \times 1$ dimension noise vector and $\boldsymbol{\kappa}$ is the kernel that characterizes the response of the imaging device. On the right-hand side of (1), $\mathbf{n}$ is an additive Gaussian noise sequence distributed according to $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_P)$, where the variance σ^2 is assumed to be unknown.

Note that in standard deblurring problems, the function $T(\cdot,\cdot)$ represents the standard n-dimensional convolution operator $\otimes$. In this case, the image $\mathbf{X}$ can be vectorized yielding the unknown image $\mathbf{x} \in \mathbb{R}^M$ with $M = P = l_1 l_2 \ldots l_n$. With this notation, (1) can be rewritten

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad \text{or} \quad \mathbf{Y} = \boldsymbol{\kappa} \otimes \mathbf{X} + \mathbf{N} \tag{2}$$

where $\mathbf{y}$ (resp. $\mathbf{n}$) stands for the vectorized version of $\mathbf{Y}$ (resp. $\mathbf{N}$) and $\mathbf{H}$ is an $P \times M$ matrix that describes convolution by the point spread function (psf) $\boldsymbol{\kappa}$.

The problem addressed in the following sections consists of estimating $\mathbf{x}$ and σ^2 under sparsity and positivity constraints on $\mathbf{x}$ given the observations $\mathbf{y}$, the psf $\boldsymbol{\kappa}$ and the bilinear function[2] $T(\cdot,\cdot)$.

III. Hierarchical Bayesian Model

A. Likelihood Function

The observation model defined in (1) and the Gaussian properties of the noise sequence $\mathbf{n}$ yield

$$f(\mathbf{y}|\mathbf{x}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{P/2} \exp\left(-\frac{\|\mathbf{y} - T(\boldsymbol{\kappa}, \mathbf{x})\|^2}{2\sigma^2}\right) \tag{3}$$

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

B. Parameter Prior Distributions

The unknown parameter vector associated with the observation model defined in (1) is $\boldsymbol{\theta} = \{\mathbf{x}, \sigma^2\}$. In this section, we introduce prior distributions for these two parameters; which are assumed to be independent.

1) Image Prior: First consider the exponential distribution with shape parameter $a > 0$

$$g_a(x_i) = \frac{1}{a} \exp\left(-\frac{x_i}{a}\right) \mathbf{1}_{\mathbb{R}_+^*}(x_i) \tag{4}$$

where $\mathbf{1}_{\mathbb{E}}(x)$ is the indicator function defined on $\mathbb{E}$

$$\mathbf{1}_{\mathbb{E}}(x) = \begin{cases} 1, & \text{if } x \in \mathbb{E} \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Choosing $g_a(\cdot)$ as prior distributions for $x_i (i = 1, \ldots, M)$ leads to a maximum *a posteriori* (MAP) estimator of $\mathbf{x}$ that corresponds to a maximum ℓ_1-penalized likelihood estimate with a positivity constraint.[3] Indeed, assuming the component $x_i (i = 1, \ldots, P)$ *a priori* independent allows one to write the full prior distribution for $\mathbf{x} = [x_1, \ldots, x_M]^T$

$$g_a(\mathbf{x}) = \left(\frac{1}{a}\right)^M \exp\left(-\frac{\|\mathbf{x}\|_1}{a}\right) \mathbf{1}_{\{\mathbf{x} \succ 0\}}(\mathbf{x}) \tag{6}$$

[2]In the following, for sake of conciseness, the same notation $T(\cdot,\cdot)$ will be adopted for the bilinear operations used on n-dimensional images $\mathbf{X}$ and used on $M \times 1$ vectorized images $\mathbf{x}$.

[3]Note that a similar estimator using a Laplacian prior for $x_i (i = 1, \ldots, M)$ was proposed in [51] for regression problems, referred to as the LASSO estimator, but without positivity constraint.

where $\{\mathbf{x} \succ 0\} = \{\mathbf{x} \in \mathbb{R}^m;\ x_i > 0, \forall i = 1, \ldots, m\}$ and $\|\cdot\|_1$ is the standard ℓ_1 norm $\|\mathbf{x}\|_1 = \sum_i |x_i|$. This estimator has interesting sparseness properties for Bayesian estimation [1] and signal representation [20].

Coupling a standard probability density function (pdf) with an atom at zero is another alternative to encourage sparsity. This strategy has for instance been used for located event detection [28] such as spike train deconvolution [5], [7]. In order to increase the sparsity of the prior, we propose to use the following distribution derived from $g_a(\cdot)$ as prior distribution for x_i

$$f(x_i|w, a) = (1 - w)\delta(x_i) + w g_a(x_i) \tag{7}$$

where $\delta(\cdot)$ is the Dirac function. This prior is similar to the LAZE distribution (Laplacian pdf and an atom at zero) introduced in [27] and used, for example, in [52], [54] for MRFM. However, since $g_a(x_i)$ is zero for $x_i \leq 0$, the proposed prior in (7) accounts for the positivity of the nonzero pixel values, a constraint that exists in many imaging modalities such as MRFM. By assuming the components x_i to be *a priori* independent ($i = 1, \ldots, M$), the following prior distribution for $\mathbf{x}$ is obtained:

$$f(\mathbf{x}|w, a) = \prod_{i=1}^{M} \left[(1 - w)\delta(x_i) + w g_a(x_i)\right]. \tag{8}$$

Introducing the index subsets $\mathcal{I}_0 = \{i; x_i = 0\}$ and $\mathcal{I}_1 = \overline{\mathcal{I}}_0 = \{i; x_i \neq 0\}$ allows one to rewrite the previous equation as follows:

$$f(\mathbf{x}|w, a) = \left[(1 - w)^{n_0} \prod_{i \in \mathcal{I}_0} \delta(x_i)\right] \left[w^{n_1} \prod_{i \in \mathcal{I}_1} g_a(x_i)\right] \tag{9}$$

with $n_\epsilon = \operatorname{card}\{\mathcal{I}_\epsilon\}$, $\epsilon \in \{0, 1\}$. Note that $n_0 = M - n_1$ and $n_1 = \|\mathbf{x}\|_0$ where $\|\cdot\|_0$ is the standard ℓ_0 norm $\|\mathbf{x}\|_0 = \#\{i; x_i \neq 0\}$.

2) Noise Variance Prior: A conjugate inverse-Gamma distribution with parameters $\nu/2$ and $\gamma/2$ is chosen as prior distribution for the noise variance [43, Appendix A]

$$\sigma^2|\nu, \gamma \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\gamma}{2}\right). \tag{10}$$

In the following, the shape parameter ν will be fixed to $\nu = 2$ and the scale parameter γ will be estimated as an hyperparameter (see [13], [14], [40]). Note that choosing the inverse-Gamma distribution $\mathcal{IG}(\nu/2, \gamma/2)$ as a prior for σ^2 is equivalent to choosing a Gamma distribution $\mathcal{G}(\nu/2, \gamma/2)$ as a prior for $1/\sigma^2$.

C. Hyperparameter Priors

The hyperparameter vector associated with the aforementioned prior distributions is $\boldsymbol{\Phi} = \{a, \gamma, w\}$. Obviously, the accuracy of the proposed Bayesian model depends on the values of these hyperparameters. Sometimes prior knowledge may be available, e.g., the mean number of nonzero pixels in the image. In this case these parameters can be tuned manually to their true values. However, in many practical situations such prior information is not available and these hyperparameters must be estimated directly from the data. Priors for these hyperparameters, sometimes referred to as "hyperpriors" are given below.

1) Hyperparameter a: A conjugate inverse-Gamma distribution is assumed for the scale parameter α of the distribution $g_a(\cdot)$ of nonzero pixel intensities

$$a|\boldsymbol{\alpha} \sim \mathcal{IG}(\alpha_0, \alpha_1) \tag{11}$$

with $\boldsymbol{\alpha} = [\alpha_0, \alpha_1]^T$. Similarly to [19], the fixed hyperparameters α_0 and α_1 have been chosen to obtain a vague prior: $\alpha_0 = \alpha_1 = 10^{-10}$.

2) Hyperparameter γ : A noninformative Jeffreys' prior [25], [26] is assumed for the scale parameter of the inverse Gamma prior density on the noise variance σ^2

$$f(\gamma) \propto \frac{1}{\gamma} \mathbf{1}_{\mathbb{R}_+}(\gamma). \tag{12}$$

The combination of the priors (10) and (12) yields the noninformative Jeffreys' prior on σ^2. Note that there is no difference between choosing a noninformative Jeffrey's prior for σ^2 and the proper hierarchical prior defined by (10) and (12). Indeed, integrating over the hyperparameter γ in the joint $f(\sigma^2, \gamma)$ distribution yields

$$\begin{aligned} f(\sigma^2) &= \int f(\sigma^2|\gamma) f(\gamma) d\gamma \\ &\propto \left(\frac{1}{\sigma^2}\right)^2 \int \exp\left(-\frac{\gamma}{2\sigma^2}\right) d\gamma \\ &\propto \frac{1}{\sigma^2}. \end{aligned} \tag{13}$$

However, in more complex noise models, the hierarchical priors $f(\sigma^2|\gamma)$ and $f(\gamma)$ are not equivalent to such a simple prior on σ^2. For example, as in [12], this pair of hierarchical priors is easily generalizable to conditionally Gaussian noise with spatial correlation and spatially varying signal-to-noise ratio.

3) Hyperparameter ω: A uniform distribution on the simplex $[0, 1]$ has been chosen as prior distribution for the mean proportion of nonzero pixels

$$\omega \sim \mathcal{U}([0, 1]). \tag{14}$$

This is the least informative prior on the image sparsity factor. Assuming that the individual hyperparameters are statistically independent the full hyperparameter prior distribution for $\boldsymbol{\Phi}$ can be expressed as

$$\begin{aligned} f(\boldsymbol{\Phi}|\boldsymbol{\alpha}) &= f(w) f(\gamma) f(a) \\ &\propto \frac{1}{\gamma a^{\alpha_0 + 1}} \exp\left(-\frac{\alpha_1}{a}\right) \\ &\quad \times \mathbf{1}_{\{0,1\}}(w) \mathbf{1}_{\mathbb{R}^+}(a) \mathbf{1}_{\mathbb{R}^+}(\gamma). \end{aligned} \tag{15}$$

D. Posterior Distribution

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

$$f(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{y}, \boldsymbol{\alpha}) \propto f(\mathbf{y}|\boldsymbol{\theta}) f(\boldsymbol{\theta}|\boldsymbol{\Phi}) f(\boldsymbol{\Phi}|\boldsymbol{\alpha}) \tag{16}$$

with

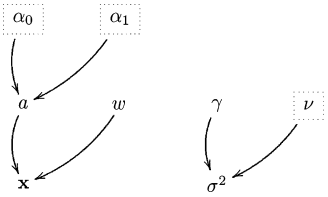

Fig. 1. DAG for the parameter priors and hyperpriors (the fixed nonrandom hyperparameters appear in dashed boxes).

$$f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right) = f\left(\mathbf{x}|a,w\right) f\left(\sigma^2|\gamma\right) \tag{17}$$

where $f\left(\mathbf{y}|\boldsymbol{\theta}\right)$ and $f\left(\boldsymbol{\Phi}|\boldsymbol{\alpha}\right)$ have been defined in (3) and (15). This hierarchical structure, represented on the directed acyclic graph (DAG) in Fig. 1, allows one to integrate out the parameter σ^2 and the hyperparameter vector $\boldsymbol{\Phi}$ in (16) to obtain the posterior of the image given the measured data and the parameters $\mathbf{x}$

$$f\left(\mathbf{x}|\mathbf{y},\boldsymbol{\alpha}\right) \propto \frac{B\left(1+n_1,1+n_0\right)}{\left\|\mathbf{y}-T\left(\boldsymbol{\kappa},\mathbf{x}\right)\right\|^P} \times \frac{\Gamma\left(n_1+\alpha_0\right)}{\left[\left\|\mathbf{x}\right\|_1+\alpha_1\right]^{n_1+\alpha_0}} \mathbf{1}_{\{\mathbf{x}\succ 0\}}\left(\mathbf{x}\right). \tag{18}$$

In (18), as defined in paragraph III-B1, $n_1 = \left\|\mathbf{x}\right\|_0$, $n_0 = M - \left\|\mathbf{x}\right\|_0$ and $B\left(\cdot,\cdot\right)$ stands for the Beta function $B\left(u,v\right) = \Gamma\left(u\right)\Gamma\left(v\right)/\Gamma\left(u+v\right)$, where $\Gamma(\cdot)$ denotes the Gamma function.

The next section presents an appropriate Gibbs sampling strategy [44, Chap. 10] that allows one to generate an image sample distributed according to the posterior distribution $f\left(\mathbf{x}|\mathbf{y},\boldsymbol{\alpha}\right)$.

IV. Gibbs Sampling Strategy for Sparse Image Reconstruction

In this section, we describe the Gibbs sampling strategy for generating samples $\left\{\mathbf{x}^{(t)}\right\}_{t=1,\ldots}$ distributed according to the posterior distribution in (18). As simulating directly according to (18) is difficult, it is much more convenient to generate samples distributed according to the joint posterior $f\left(\mathbf{x},\sigma^2,a,w|\mathbf{y},\boldsymbol{\alpha}\right)$. This Gibbs sampler produces sequences $\left\{\mathbf{x}^{(t)}\right\}_{t=1,\ldots}$, $\left\{\sigma^{2(t)}\right\}_{t=1,\ldots}$, $\left\{a^{(t)}\right\}_{t=1,\ldots}$, $\left\{w^{(t)}\right\}_{t=1,\ldots}$ which are Markov chains with stationary distributions $f\left(\mathbf{x}|\mathbf{y},\boldsymbol{\alpha}\right)$, $f\left(\sigma^2|\mathbf{y},\boldsymbol{\alpha}\right)$, $f\left(a|\mathbf{y},\boldsymbol{\alpha}\right)$ and $f\left(w|\mathbf{y},\boldsymbol{\alpha}\right)$, respectively, [44, p. 345]. Then, the MAP estimator of the unknown image $\mathbf{x}$ will be computed by retaining among $\mathcal{X} = \left\{\mathbf{x}^{(t)}\right\}_{t=1,\ldots}$ the generated sample that maximizes the posterior distribution in (18) [35, p. 165]

$$\begin{aligned}\hat{\mathbf{x}}_{\mathrm{MAP}} &= \arg\max_{\mathbf{x}\in\mathbb{R}_+^M} f\left(\mathbf{x}|\mathbf{y}\right) \\ &\approx \arg\max_{\mathbf{x}\in\mathcal{X}} f\left(\mathbf{x}|\mathbf{y}\right).\end{aligned} \tag{19}$$

The main steps of this algorithm are given in Sections IV-A and IV-D (see also Algorithm 1 below).

Algorithm 1: Gibbs Sampling Algorithm for Sparse Image Reconstruction:

- Initialization:
  - — Sample parameter $\mathbf{x}^{(0)}$ from the pdf in (9).
  - — Sample parameter $\widetilde{\sigma}^{2(0)}$ from the pdf in (10).
  - — Set $t \leftarrow 1$.
- Iterations: for $t = 1, 2, \ldots,$ do:
  1. Sample hyperparameter $w^{(t)}$ from the pdf in (21).
  2. Sample hyperparameter $a^{(t)}$ from the pdf in (22).
  3. For $i = 1, \ldots, M$, sample parameter $x_i^{(t)}$ from the pdf in (23).
  4. Sample parameter $\widetilde{\sigma}^{2(t)}$ from the pdf in (26).
  5. Set $t \leftarrow t+1$.

A. Generation of Samples According to $f\left(w\left|\mathbf{X}\right.\right)$

Using (9), the following result can be obtained:

$$f\left(w\left|\mathbf{x}\right.\right) \propto \left(1-w\right)^{n_0} w^{n_1} \tag{20}$$

where n_0 and n_1 have been defined in paragraph Section III-B-I. Therefore, samples from $f\left(w\left|\mathbf{x}\right.\right)$ can be generated by simulating from an image dependent Beta distribution

$$w\left|\mathbf{x}\right. \sim \mathcal{B}e\left(1+n_1,1+n_0\right). \tag{21}$$

B. Generation of Samples According to $f\left(a\left|\mathbf{x},\boldsymbol{\alpha}\right.\right)$

The form of the joint posterior distribution (16) implies that samples of a can be generated by simulating from an image dependent inverse-Gamma distribution

$$a\left|\mathbf{x},\boldsymbol{\alpha}\right. \sim \mathcal{IG}\left(\left\|\mathbf{x}\right\|_0+\alpha_0, \left\|\mathbf{x}\right\|_1+\alpha_1\right). \tag{22}$$

C. Generation of Samples According to $f\left(\mathbf{x}\left|w,a,\sigma^2,\mathbf{y}\right.\right)$

The LAZE-type prior (7) chosen for x_i $(i = 1, \ldots, M)$ yields a posterior distribution of $\mathbf{x}$ that is not closed form. However, one can easily derive the posterior distribution of each pixel intensity x_i $(i = 1, \ldots, M)$ conditioned on the intensities of the rest of the image. Indeed straightforward computations (Appendix A) yield

$$f\left(x_i|w,a,\sigma^2,\mathbf{x}_{-i},\mathbf{y}\right) \propto \left(1-w_i\right)\delta\left(x_i\right) + w_i\phi_+\left(x_i|\mu_i,\eta_i^2\right) \tag{23}$$

where $\mathbf{x}_{-i}$ stands for the vector $\mathbf{x}$ whose ith component has been removed and μ_i and η_i^2 are given in Appendix A. In (23), $\phi_+\left(\cdot,m,s^2\right)$ stands for the pdf of the truncated Gaussian distribution defined on $\mathbb{R}_+^*$ with hidden mean and variance parameters equal to m and s^2, respectively

$$\phi_+\left(x,m,s^2\right) = \frac{1}{C\left(m,s^2\right)} \exp\left[-\frac{\left(x-m\right)^2}{2s^2}\right] \mathbf{1}_{\mathbb{R}_+^*}\left(x\right) \tag{24}$$

with

$$C\left(m,s^2\right) = \sqrt{\frac{\pi s^2}{2}}\left[1+\mathrm{erf}\left(\frac{m}{\sqrt{2s^2}}\right)\right]. \tag{25}$$

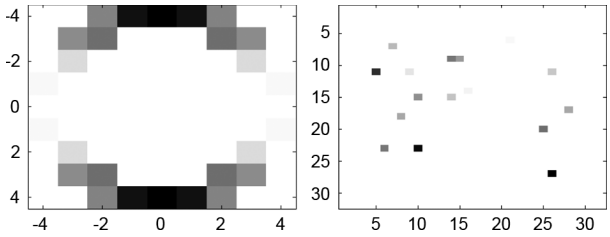

Fig. 2. Left: psf of the MRFM tip. Right: unknown sparse image to be estimated.

The form in (23) specifies $x_i|w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}$ as a Bernoulli-truncated Gaussian variable with parameter $\left(w_i, \mu_i, \eta_i^2\right)$. Appendix C presents an algorithm that can be used to generate samples from this distribution. This algorithm generates samples distributed according to $f\left(\mathbf{x}\middle|w, \sigma^2, a, \mathbf{y}\right)$ by successively updating the coordinates of $\mathbf{x}$ using a sequence of M Gibbs moves (requiring generation of Bernoulli-truncated Gaussian variables).

D. Generation of Samples According to $f\left(\sigma^2|\mathbf{x}, \mathbf{y}\right)$

Samples are generated in the following way:

$$\sigma^2|\mathbf{x}, \mathbf{y} \sim \mathcal{IG}\left(\frac{P}{2}, \frac{\|\mathbf{y} - T(\boldsymbol{\kappa}, \mathbf{x})\|^2}{2}\right). \tag{26}$$

V. Simulation on Synthetic Images

A. Reconstruction of 2-D Image

In this subsection, a 32×32 synthetic image, depicted in Fig. 2 (right panel), is simulated using the prior in (9) with parameters $a = 1$ and $w = 0.02$. In Figs. 2 and 3, white pixels stands for zero intensity values. A general analytical derivation of the psf of the MRFM tip has been given in [34] and with further explanation in [54]. Following this model, we defined a 10×10 2-D convolution kernel, the psf represented in Fig. 2 (left panel), that corresponds to the physical parameters shown in Table I. The associated psf matrix $\mathbf{H}$ introduced in (2) is of size 1024×1024. The observed measurements $\mathbf{y}$, which are of size $P = 1024$ and depicted in Fig. 3 (top panel), are corrupted by an additive Gaussian noise with two different variances $\sigma^2 = 1.2 \times 10^{-1}$ and $\sigma^2 = 1.6 \times 10^{-3}$, corresponding to signal-to-noise ratios $\mathrm{SNR} = 2$ dB and $\mathrm{SNR} = 20$ dB, respectively.

1) Simulation Results: The observations are processed by the proposed algorithm using $N_{\mathrm{MC}} = 2000$ iterations of the Gibbs sampler with $N_{\mathrm{bi}} = 300$ burn-in iterations. The computation time for completing 100 iterations of the proposed algorithm is 80 s for an unoptimized MATLAB 2007b 32-bit implementation on a 2.2-GHz Intel Core 2, while 100 iterations of the Landweber and empirical Bayesian algorithms require 0.15 s and 2 s, respectively. The MAP image reconstruction computed using (19) is depicted in Fig. 3 (bottom panel) for the two levels of noise considered. Observe that the estimated image is very similar to the actual image, Fig. 2 (right panel), even at low SNR.

Moreover, as the proposed algorithm generates samples distributed according to the posterior distribution in (18), these

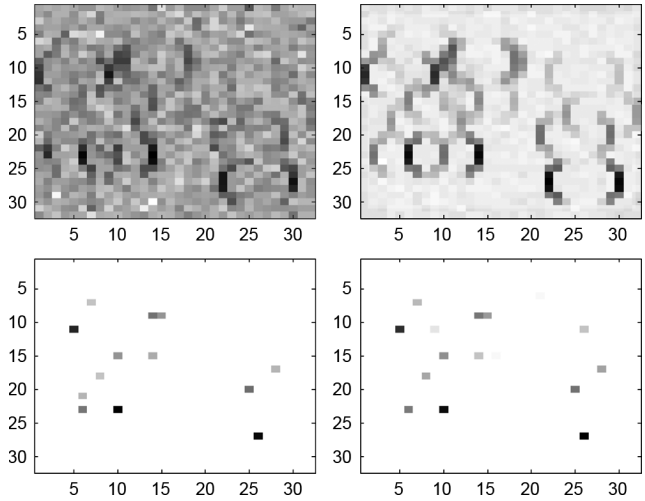

Fig. 3. Top, left (resp. right): noisy observations for $\mathrm{SNR} = 2$ dB (resp. 20 dB). Bottom, left (resp. right): reconstructed image for $\mathrm{SNR} = 2$ dB (resp. 20 dB).

TABLE I
Parameters Used to Compute the MRFM psf

Parameter		Value
Description	Name	
Amplitude of external magnetic field	B_{ext}	9.4×10^3 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	6.0 nm
Cantilever tip moment	m	4.6×10^5 emu
Peak cantilever oscillation	x_{pk}	0.8 nm
Maximum magnetic field gradient	G_{max}	125

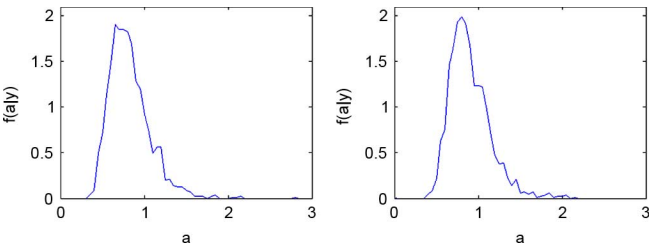

Fig. 4. Posterior distribution of hyperparameter a (left: $\mathrm{SNR} = 2$ dB, right: $\mathrm{SNR} = 20$ dB).

samples can be used to compute the posterior distributions of each parameter. For illustration, the posterior distributions of the hyperparameters a and w, as well as the noise variance σ^2, are shown in Figs. 4–6. These estimated distributions are in good agreement with the ground truth values of these parameters, randomly drawn from the prior distribution.

In many applications, a measure of confidence that a given pixel or pixel region is nonzero is of interest. Our Bayesian approach can easily generate such measures of confidence in the

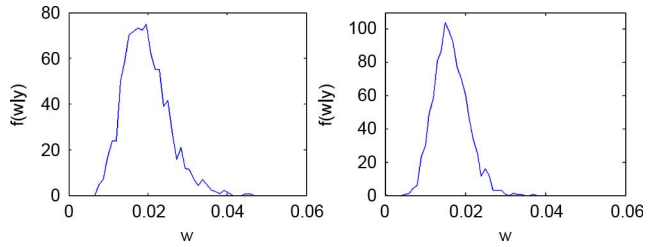

Fig. 5. Posterior distribution of hyperparameter w (left: $\mathrm{SNR} = 2$ dB, right: $\mathrm{SNR} = 20$ dB).

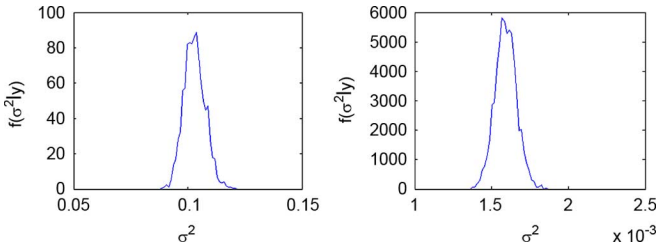

Fig. 6. Posterior distribution of hyperparameter σ^2 (left: $\mathrm{SNR} = 2$ dB, right: $\mathrm{SNR} = 20$ dB).

form of posterior probabilities of the specified event, sometimes known as the Bayesian p-value. Following the strategy detailed in Appendix C, the proposed Gibbs sampler generates a collection of samples $\left\{\mathbf{x}^{(t)}\right\}_{t=1,\ldots,N_{\mathrm{MC}}}$, distributed according the posterior Bernoulli-truncated Gaussian distribution in (23). This sampling requires the generation of indicator variables z_i $(i = 1,\ldots,n)$ that reflect the presence or the absence of nonzero pixel values. It is the indicator variable z_i that $x_i > 0$ that provides information about nonzero pixels in the image. Using the equivalences $\{z_i = 0\} \Leftrightarrow \{x_i = 0\}$ and $\{z_i = 1\} \Leftrightarrow \{x_i > 0\}$, the posterior probability $\mathrm{P}\left[x_i > 0|\mathbf{y}, \boldsymbol{\alpha}\right]$ can be easily obtained by averaging over the Gibbs samples of the binary variables $\left\{z_i^{(t)}\right\}_{t=N_{\mathrm{bi}}+1,\ldots,N_{\mathrm{MC}}}$. To illustrate, these probabilities are depicted in Fig. 7. In addition, these Gibbs samples can be used to compute the probability of having nonzero pixels in a given area of the image. The estimated posterior probability for the event that a nonzero pixel is present inside the small red rectangle in the figure is equal to 45% for the case of $\mathrm{SNR} = 2$ dB. Conversely, the posterior probability of having a nonzero pixel in the green box is 5%. For $\mathrm{SNR} = 20$ dB the MAP algorithm correctly detects up the presence of a pixel in this region. On the other hand, even though at $\mathrm{SNR} = 2$ dB the MAP reconstruction has not detected this pixel, we can be 45% confident of the presence of such a pixel in the red rectangular region on the left panel of Fig. 7.

The posterior distributions of four different pixels are depicted in Fig. 8. These posteriors are consistent with the actual values of these pixels that are represented as dotted red lines in these figures. In particular, in all cases the actual values all lie within the 75% central quantile of the posterior distribution.

2) Comparison of Reconstruction Performances: Here we compare our proposed hierarchical Bayesian method to the sparse reconstruction methods of [52], [54]. The techniques proposed in [52], [54] are based on penalized likelihood EM algorithms that perform empirical estimation of the unknown

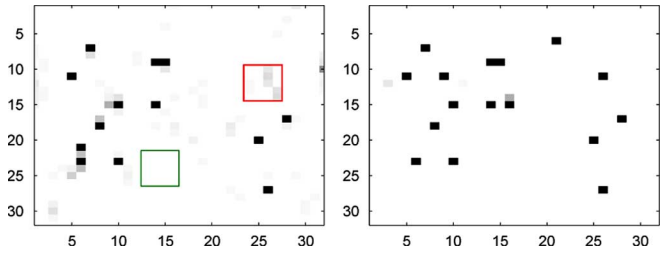

Fig. 7. Posterior probabilities of having nonzero pixels (left: $\mathrm{SNR} = 2$ dB, right: $\mathrm{SNR} = 20$ dB). The probability of having at least one nonzero pixel in the red (resp. green) box-delimited area is 45% (resp. 5%).

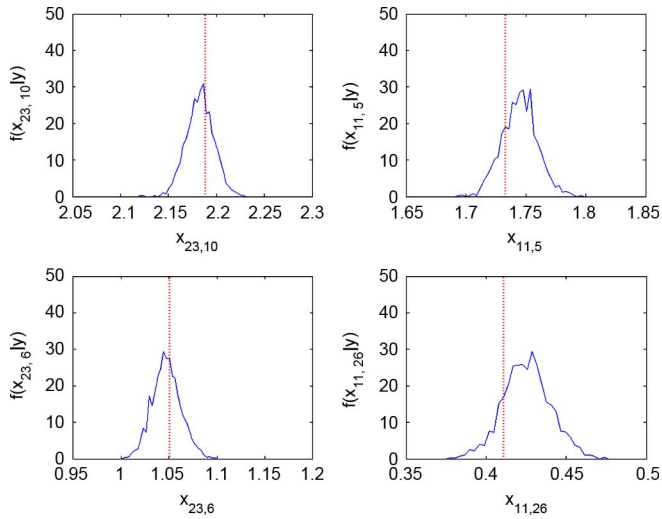

Fig. 8. Posterior distributions of the nonzero values of $\mathbf{x}$ for 4 different pixel locations and for $\mathrm{SNR} = 20$ dB (actual pixel intensity values are depicted with dotted red lines).

hyperparameters. Therein, two empirical Bayesian estimators, denoted Emp-MAP-Lap and Emp-MAP-LAZE, based on a Laplacian or a LAZE prior respectively, were proposed. We also compare to the standard Landweber algorithm [31] that has been previously used to perform MRFM image reconstruction [8], [57]. These are compared to our hierarchical Bayesian MAP reconstruction algorithm, given in (19), and also to a minimum mean square error (MMSE) reconstruction algorithm extracted from the estimated full Bayes posterior (18). The MMSE estimator of the image $\mathbf{x}$ is obtained by empirical averaging over the last $N_r = 1700$ samples of the Gibbs sampler according to

$$\begin{aligned}\hat{\mathbf{x}}_{\mathrm{MMSE}} &= \mathrm{E}\left[\mathbf{x}|\mathbf{y}\right] \\ &\approx \frac{1}{N_r}\sum_{t=1}^{N_r}\mathbf{x}^{(N_{\mathrm{bi}}+t)}. \end{aligned} \tag{27}$$

As in [54], we compare the various reconstruction algorithms with respect to several performance criteria. 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. These criteria are: the ℓ_0, ℓ_1, and ℓ_2-norms of $\mathbf{e}$, which measures the accuracy of the reconstruction, and the ℓ_0-norm of the estimator $\hat{\mathbf{x}}$, which measures its sparsity. As pointed out in [54], a human observer can usually not visually detect the presence of nonzero intensities if they are

TABLE II
RECONSTRUCTION PERFORMANCES FOR DIFFERENT SPARSE IMAGE RECONSTRUCTION ALGORITHMS

Method	Error criterion					
	$\|e\|_0$	$\|e\|_\delta$	$\|e\|_1$	$\|e\|_2$	$\|\hat{\mathbf{x}}\|_0$	$\|\hat{\mathbf{x}}\|_\delta$
SNR = 2dB						
Landweber	1024	990	339.76	13.32	1024	990
Emp-MAP-Lap	18	17	14.13	4.40	**0**	**0**
Emp-MAP-LAZE	60	58	9.49	1.44	55	55
Proposed MMSE	1001	30	3.84	**0.72**	1001	27
Proposed MAP	**19**	**16**	**2.38**	0.81	13	13
SNR = 20dB						
Landweber	1024	931	168.85	6.67	1024	931
Emp-MAP-Lap	33	18	1.27	0.31	28	23
Emp-MAP-LAZE	144	19	1.68	0.22	144	27
Proposed MMSE	541	**5**	**0.36**	**0.11**	541	**16**
Proposed MAP	**19**	7	0.39	0.13	**16**	**16**

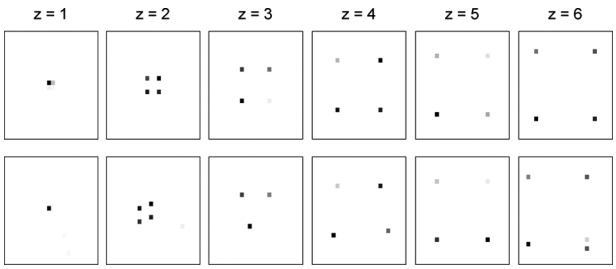

Fig. 9. Top: Slices of the sparse image to be recovered. Bottom: Slices of the estimated sparse image.

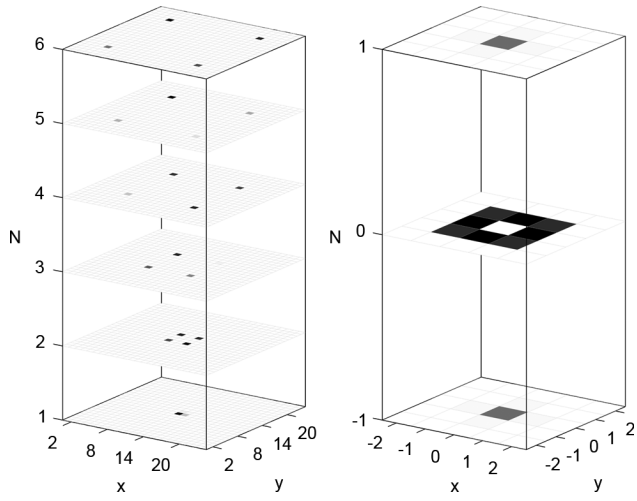

Fig. 10. Left: $24 \times 24 \times 6$ unknown image to be recovered. Right: $5 \times 5 \times 3$ kernel modeling the psf.

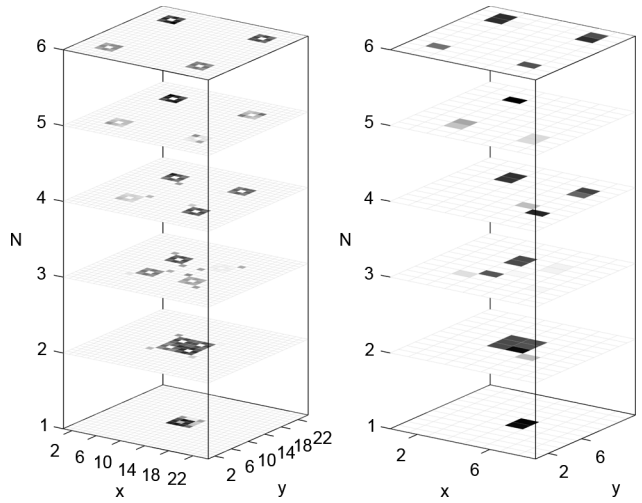

Fig. 11. Left: $24 \times 24 \times 6$ regularly sampled convolved image. Left: $12 \times 8 \times 6$ undersampled observed image.

below a small threshold. Thus, a less strict measure[4] of sparsity than the ℓ_0-norm, which is denoted $\|\cdot\|_\delta$, is the number of reconstructed image pixels that are less than a given threshold δ

$$\|\hat{\mathbf{x}}\|_\delta = \sum_{i=1}^{M} \mathbf{1}_{\hat{x}_i<\delta}(\hat{x}_i)$$
$$\|\mathbf{e}\|_\delta = \sum_{i=1}^{M} \mathbf{1}_{e_i<\delta}(e_i). \quad (28)$$

It what follows, δ has been chosen as $\delta = 10^{-2}\|\mathbf{x}\|_\infty$. To summarize, the following criteria have been computed for the image in paragraph V-A1 for two levels of SNR: $\|e\|_0$, $\|e\|_\delta$, $\|e\|_1$, $\|e\|_2$, $\|\hat{\mathbf{x}}\|_0$ and $\|\hat{\mathbf{x}}\|_\delta$.

Table II shows the six performance measures for the five different algorithms studied. The proposed Bayesian methods (labeled "proposed MMSE" and "proposed MAP" in the table) outperform the other reconstruction algorithms in terms of ℓ_1 or ℓ_2-norms. Note that the MMSE estimation of the unknown image is a nonsparse estimator in the ℓ_0-norm sense. This is due to the very small but nonzero posterior probability of nonzero value at many pixels. The sparsity measure $\|\cdot\|_\delta$ indicates that most of the pixels are in fact very close to zero. The MAP reconstruction method seems to achieve the best balance between the sparsity of the solution and the minimization of the reconstruction error. Of course, by its very construction, the MMSE reconstruction will always have lower mean square error.

B. Reconstruction of Undersampled 3-D Images

As discussed below in Section VI, the prototype IBM MRFM instrument [8] collects data projections as irregularly spaced, or undersampled, spatial samples. In this subsection, we indicate how the image reconstruction algorithm can be adapted to this undersampled scenario in 3-D. We illustrate by a concrete example. First, a $24 \times 24 \times 6$ image is generated such that 4 pixels have nonzero values in each z slice. The resulting data is

[4]The introduced measure of sparsity is denoted $\|\cdot\|_\delta$. This is an abuse of notation since it is not a norm.

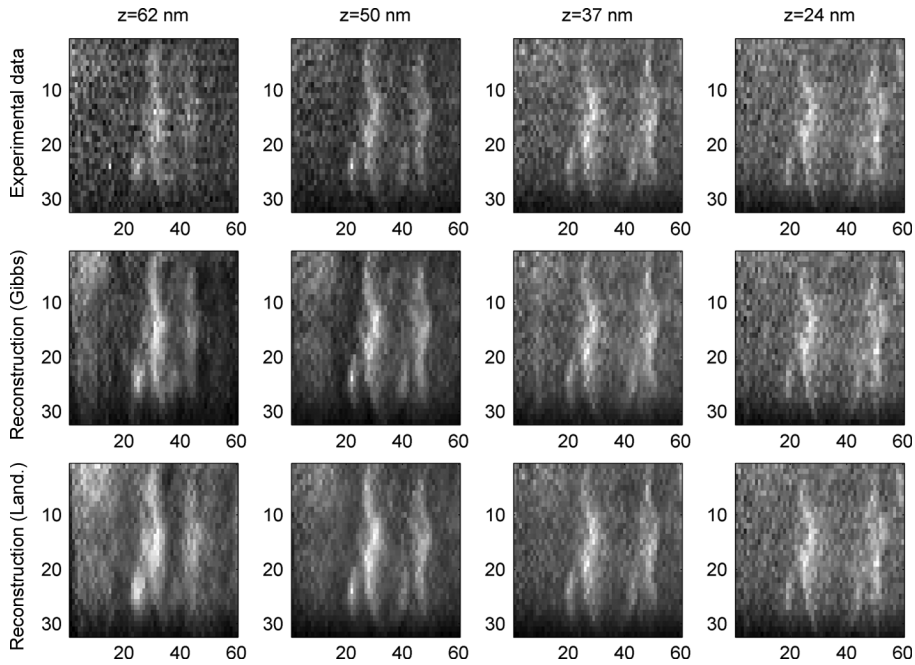

Fig. 12. Top: Experimental scan data where black (resp. white) pixel represents low (resp. high) density of spin (as in [8]). Middle: Scan data reconstructed from the proposed hierarchical Bayesian algorithm. Bottom: Scan data reconstructed from the Landweber algorithm.

depicted in Figs. 9 (top) and 10 (left). This image to be recovered is assumed to be convolved with a $5 \times 5 \times 3$ kernel that is represented in Fig. 10 (right). The resulting convolved image is depicted in Fig. 11 (left). However, the actual observed image is an undersampled version of this image. More precisely, the sampling rates are assumed to be $d_x = 2$, $d_y = 3$, $d_z = 1$, respectively, in the 3 dimensions. Consequently the observed 3-D image, shown in Fig. 11, is of size $12 \times 8 \times 6$. Finally, an i.i.d. Gaussian noise with $\sigma = 0.02$ is added following the model in (1). Note that under these assumptions, the application $T(\cdot,\cdot)$ can be split into two standard operations following the composition:

$$T(\boldsymbol{\kappa}, \mathbf{X}) = g_{d_x,d_y,d_z}(\boldsymbol{\kappa} \otimes \mathbf{X}) \tag{29}$$

where $g_{d_x,d_y,d_z}(\cdot)$ stands for the undersampling function.

The proposed hierarchical Bayesian algorithm is used to perform the sparse reconstruction with undersampled data. The number of Monte Carlo runs was fixed to $N_{\mathrm{MC}} = 2000$ with $N_{\mathrm{bi}} = 300$ burn-in iterations. Fig. 9 shows the result of applying the proposed MAP estimator to the estimated posterior.

VI. Application on Real MRFM Images

Here, we illustrate the hierarchical Bayesian MAP reconstruction algorithm for real 3-D MRFM data. The data is a set of MRFM projections of a sample of tobacco virus. Comprehensive details of both the experiment and the MRFM data acquisition protocol are given in [8] and the supplementary materials [9]. The observed sample consists of a collection of Tobacco mosaic virus particles that constitute a whole viral segment in addition to viral fragments. The projections are computed from the measured proton distribution and the 3-D psf following the protocol described in [8] and [9]. The resulting scan data are depicted in Fig. 12 (top) for four different distances between the MRFM tip and the sample: $d = 24$ nm, $d = 37$ nm, $d = 50$ nm and $d = 62$ nm. Each of these x-y slices is of size 60×32 pixels.

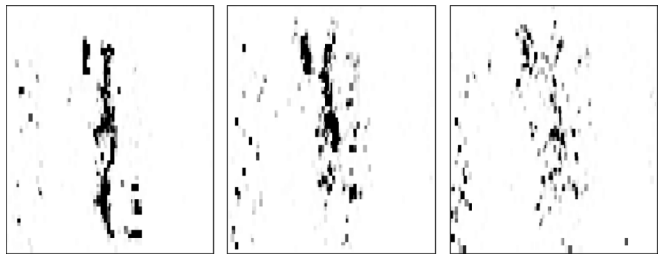

Fig. 13. Three horizontal slices of the estimated image.

These experimental data are undersampled, i.e., the spatial resolution of the MRFM tip, and, therefore, the psf function, is finer than the resolution of the observed slices. Consequently, these data have been deconvolved taking into account the oversampling rates defined by $d_x = 3$, $d_y = 2$ and $d_z = 3$ in the three directions. The MAP estimate of the unknown image is computed from $N_{\mathrm{MC}} = 1000$ Gibbs samples of the proposed Bayesian algorithm initialized with the output of a single Landweber iteration. Several more iterations of the Landweber algorithm would produce the reconstructions reported in [8]. Three horizontal slices of the estimated image[5] are depicted in Fig. 13. A 3-D view of the estimated profile of the virus fragments is shown in Fig. 14. The MMSE estimates of the parameters introduced in Section III are $\hat{\sigma}^2_{\mathrm{MMSE}} = 0.10$, $\hat{a}_{\mathrm{MMSE}} =$

[5]Note that most part of the estimated 3-D image is empty space due to the very localized proton spin centers in the image.

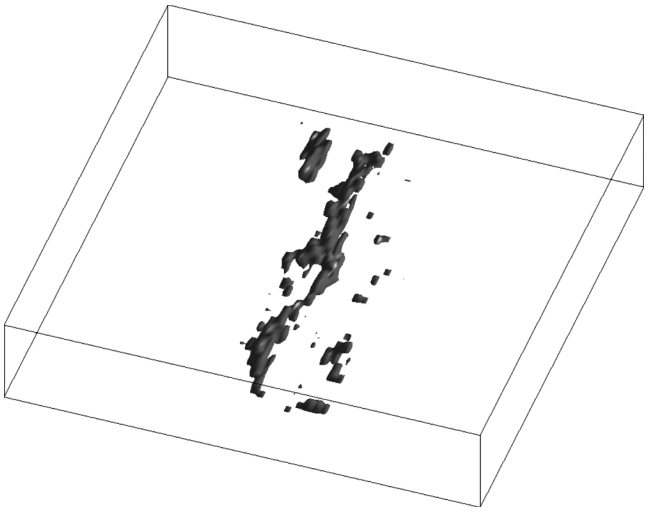

Fig. 14. Three-dimensional view of the estimated profile of the Tobacco virus fragments.

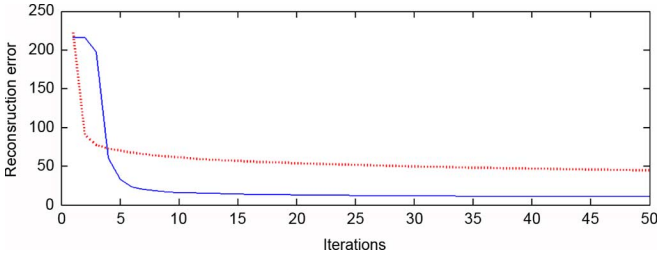

Fig. 15. Error of the reconstructions as functions of the iteration number for the proposed algorithm (continuous blue line) and Landweber algorithm (dotted red line).

1.9×10^{-12} and $\hat{w}_{\mathrm{MMSE}} = 1.4 \times 10^{-2}$. The image reconstructions produced by the Landweber and Bayesian MAP algorithms are shown in Fig. 12.

By forward projecting the estimated virus image through the point spread function one can visually evaluate the goodness of fit of the reconstruction to the raw measured data. This is depicted in Fig. 12. These figures are clearly in good agreement with the observed data (top). To evaluate the convergence speed, the reconstruction error is represented in Fig. 15 as a function of the iterations for the proposed Bayesian and the Landweber algorithms. This shows that the convergence rate of our algorithm is significantly better than the Landweber algorithm.

VII. Conclusion

This paper presented a hierarchical Bayesian algorithm for deconvolving sparse positive images corrupted by additive Gaussian noise. A Bernoulli-truncated exponential distribution was proposed as a prior for the sparse image to be recovered. The unknown hyperparameters of the model were integrated out of the posterior distribution of the image, producing a full posterior distribution that can be used for estimation of the pixel values by extracting the mode (MAP) or the first moment (MMSE). An efficient Gibbs sampler was used to generate approximations to these estimates. The derived Bayesian estimators significantly outperformed several previously proposed sparse reconstruction algorithms. Our approach was implemented on real MRFM data to reconstruct a 3-D image of a tobacco virus. Future work will include extension of the proposed method to other sparse bases, inclusion of uncertain point spread functions, and investigation of molecular priors. Future investigations might also include a comparison between the proposed MCMC approach and variational Bayes approaches.

Appendix A
Derivation of the Conditional Posterior Distribution $f\left(x_i \left| w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right.\right)$

The posterior distribution of each component x_i $(i = 1, \ldots, M)$ conditionally upon the others is linked to the likelihood function (3) and the prior distribution (7) via the Bayes' formula

$$f\left(x_i | w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right) \propto f\left(\mathbf{y} | \mathbf{x}, \sigma^2\right) f\left(x_i | w, a\right). \tag{30}$$

This distribution can be easily derived by decomposing $\mathbf{x}$ on the standard orthonormal basis

$$\mathbb{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_M\} \tag{31}$$

where $\mathbf{u}_i$ is the ith column of the $M \times M$ identity matrix. Indeed, let decompose $\mathbf{x}$ as follows:

$$\mathbf{x} = \tilde{\mathbf{x}}_i + x_i \mathbf{u}_i \tag{32}$$

where $\tilde{\mathbf{x}}_i$ is the vector $\mathbf{x}$ whose ith element has been replaced by 0. Then the linear property of the operator $T\left(\boldsymbol{\kappa}, \cdot\right)$ allows one to state

$$T\left(\boldsymbol{\kappa}, \mathbf{x}\right) = T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i\right) + x_i T\left(\boldsymbol{\kappa}, \mathbf{u}_i\right). \tag{33}$$

Consequently, (30) can be rewritten

$$f\left(x_i | w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right) \propto \exp\left(-\frac{\left\|\mathbf{e}_i - x_i \mathbf{h}_i\right\|^2}{2\sigma^2}\right) \times \left[(1-w)\delta\left(x_i\right) + \frac{w}{a}\exp\left(-\frac{x_i}{a}\right)\mathbf{1}_{\mathbb{R}_+^*}\left(x_i\right)\right] \tag{34}$$

where[6]

$$\begin{cases} \mathbf{e}_i = \mathbf{y} - T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i\right) \\ \mathbf{h}_i = T\left(\boldsymbol{\kappa}, \mathbf{u}_i\right). \end{cases} \tag{35}$$

An efficient way to compute $\mathbf{e}_i$ within the Gibbs sampler scheme is reported in Appendix B. Then, straightforward computations similar to those in [7] and [37, Annex B], yield to the following distribution:

$$f\left(x_i | w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right) \propto (1-w_i)\delta\left(x_i\right) + w_i \phi_+\left(x_i | \mu_i, \eta_i^2\right) \tag{36}$$

[6]It can be noticed that, for deblurring applications, $\mathbf{h}_i$ is also the ith column of the matrix $\mathbf{H}$ introduced in (2).

with

$$\begin{cases} \eta_i^2 = \frac{\sigma^2}{\|\mathbf{h}_i\|^2} \\ \mu_i = \eta_i^2 \left(\frac{\mathbf{h}_i^T \mathbf{e}_i}{\sigma^2} - \frac{1}{a} \right) \end{cases} \tag{37}$$

and

$$\begin{cases} u_i = \frac{w}{a} C\left(\mu_i, \eta_i^2\right) \exp\left(\frac{\mu_i^2}{2\eta_i^2}\right) \\ w_i = \frac{u_i}{u_i + (1-w)}. \end{cases} \tag{38}$$

The distribution in (36) is a Bernoulli-truncated Gaussian distribution with hidden mean μ_i and hidden variance η_i^2.

Appendix B
Fast Recursive Computations for Simulating According to $f\left(\mathbf{x} \middle| w, a, \sigma^2, \mathbf{y}\right)$

In the Gibbs sampling strategy presented in Section IV, the main computationally expensive task is the generation of samples distributed according to $f\left(x_i \middle| w, a, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right)$. Indeed, the evaluation of the hidden mean and hidden variance in (37) of the Bernoulli-truncated Gaussian distribution may be costly, especially when the bilinear application $T(\cdot,\cdot)$ is not easily computable. In this appendix, an appropriate recursive strategy is proposed to accelerate the Gibbs sampling by efficiently updating the coordinate i of the vector $\mathbf{x}$ at iteration t of the Gibbs sampler.

Let $\mathbf{x}^{(t,i-1)}$ denote the current Monte Carlo state of the unknown vectorized image $\mathbf{x}$ $(i = 1, \ldots, M)$

$$\mathbf{x}^{(t,i-1)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, x_i^{(t-1)}, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T. \tag{39}$$

with, by definition, $\mathbf{x}^{(t,0)} = \mathbf{x}^{(t-1,M)}$. Updating $\mathbf{x}^{(t,i-1)}$ consists of drawing $x_i^{(t)}$ according to the Bernoulli-truncated Gaussian distribution $f\left(x_i \middle| w, a, \sigma^2, \mathbf{x}_{-i}^{(t,i-1)}, \mathbf{y}\right)$ in (23) with

$$\mathbf{x}_{-i}^{(t,i-1)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T. \tag{40}$$

The proposed strategy to simulate efficiently according to (23) is based on the following property.

Property: Given the quantity $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(0)}\right)$ and the vectors $\{\mathbf{h}_i\}_{i=1,\ldots,M}$, simulating according to $f\left(x_i \middle| w, a, \sigma^2, \mathbf{x}_{-i}^{(t,i)}, \mathbf{y}\right)$ can be performed without evaluating the bilinear function $T(\cdot,\cdot)$.

Proof: Simulating according to (23) mainly requires to compute the vector $\mathbf{e}_i$ introduced by (35)

$$\mathbf{e}_i = \mathbf{y} - T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) \tag{41}$$

with

$$\tilde{\mathbf{x}}_i^{(t,i-1)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, 0, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T. \tag{42}$$

Moreover, by using the decomposition in (32) and by exploiting the linear property of $T(\boldsymbol{\kappa}, \cdot)$, the vector $T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right)$ on the right-hand side of (41) can be rewritten as

$$T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) = T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i-1)}\right) - x_i^{(t-1)} \mathbf{h}_i \tag{43}$$

where $\mathbf{h}_i$ has been introduced in (35). Consequently, to prove the property, we have to demonstrate that the vector series $\left\{T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,k)}\right)\right\}_{k=1,\ldots,M}$ can be computed recursively without using $T(\cdot,\cdot)$. Assume that $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i-1)}\right)$ is available at this stage of the Gibbs sampling and that $x_i^{(t)}$ has been drawn. The new Monte Carlo state is then

$$\mathbf{x}^{(t,i)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, x_i^{(t)}, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T. \tag{44}$$

Similarly to (43), the vector $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}\right)$ can be decomposed as follows:

$$T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}\right) = T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) + x_i^{(t)} \mathbf{h}_i. \tag{45}$$

Therefore, combining (43) and (45) allow one to state

$$T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}\right) = T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i-1)}\right) + \left(x_i^{(t)} - x_i^{(t-1)}\right) \mathbf{h}_i.$$

■

The bilinear function $T(\cdot,\cdot)$ only needs to be used at the very beginning of the Gibbs sampling algorithm to evaluate $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(0)}\right)$ and the vectors $\{\mathbf{h}_i\}_{i=1,\ldots,M}$. The resulting simulation scheme corresponding to step 3 of Algorithm 1 is shown in Algorithm 2.

Algorithm 2: Efficient Simulation According to $f\left(\mathbf{x} \middle| w, a, \sigma^2, \mathbf{y}\right)$*:* For $i = 1, \ldots, M$, update the ith coordinate of the vector

$$\mathbf{x}^{(t,i-1)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, x_i^{(t-1)}, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T$$

via the following steps:

1. Compute $\|\mathbf{h}_i\|^2$.
2. Set $T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) = T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i-1)}\right) - x_i^{(t-1)} \mathbf{h}_i$.
3. Set $\mathbf{e}_i = \mathbf{x} - T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right)$.
4. Compute μ_i, η_i^2 and w_i as defined in (37) and (38).
5. Draw $x_i^{(t)}$ according to (23).
6. Set $\mathbf{x}^{(t,i)} = \left[x_1^{(t)}, \ldots, x_{i-1}^{(t)}, x_i^{(t)}, x_{i+1}^{(t-1)}, \ldots, x_M^{(t-1)}\right]^T$.
7. Set $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}\right) = T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) + x_i^{(t)} \mathbf{h}_i$.

Appendix C
Simulation According to a Bernoulli-Truncated Gaussian Distribution

This appendix describes how we generate random variables distributed according to a Bernoulli-truncated Gaussian distribution with parameters $\left(w_i, \mu_i, \eta_i^2\right)$ whose pdf is

$$f\left(x_i|w_i,\mu_i,\eta_i^2\right) = (1-w_i)\,\delta(x_i) + \frac{w_i}{C\left(\mu_i,\eta_i^2\right)}\exp\left[-\frac{\left(x_i-\mu_i\right)^2}{2\eta_i^2}\right]\mathbf{1}_{\mathbb{R}_+^*}(x_i)$$

where $C\left(\mu_i,\eta_i^2\right)$ has been defined in (25). Monte Carlo draws from this density can be obtained by using an auxiliary binary variable z_i following the strategy shown in Algorithm 3. This indicator variable takes the value 0 (resp. 1) if the pixel x_i is zero (resp. nonzero).

Algorithm 3: Simulation According to a Bernoulli-Truncated Gaussian Distribution:

1. Generate z_i according to $z_i \sim \mathcal{B}er\left(w_i\right)$.
2. Set $\begin{cases} x_i = 0, & \text{if } z_i = 0; \\ x_i \sim \mathcal{N}^+\left(\mu_i,\eta_i^2\right), & \text{if } z_i = 1. \end{cases}$

In Algorithm 3, $\mathcal{B}er\left(\cdot\right)$ and $\mathcal{N}^+\left(\cdot,\cdot\right)$ denote the Bernoulli and the positive truncated Gaussian distributions respectively. In step 2, samples distributed according to the truncated Gaussian distribution can be generated by using an appropriate accept-reject procedure with instrumental distributions [18], [36], [42].

ACKNOWLEDGMENT

The authors would like to thank M. Ting (Seagate Technology) for providing the code to generate point spread functions of MRFM tip and H. Snoussi (University of Technology of Troyes) for interesting suggestions regarding this work. The authors would also like to thank Dr. D. Rugar who provided the real data used in Section VI, as well as valuable feedback on this paper.

REFERENCES

[1] S. Alliney and S. A. Ruzinsky, "An algorithm for the minimization of mixed L1 and L2 norms with application to Bayesian estimation," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 618–627, Mar. 1994.

[2] H. Andrews and B. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[3] S. Bourguignon and H. Carfantan, "Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data," in *Proc. IEEE Workshop on Stat. Signal Processing (SSP)*, Bordeaux, France, Jul. 2005, pp. 811–816.

[4] S. Bourguignon, H. Carfantan, and J. Idier, "A sparsity-based method for the estimation of spectral lines from irregularly sampled data," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 575–585, Dec. 2007.

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

[6] S. Chao, W. M. Dougherty, J. L. Garbini, and J. A. Sidles, "Nanometer-scale magnetic resonance imaging," *Review Sci. Instrum.*, vol. 75, no. 5, pp. 1175–1181, Apr. 2004.

[7] Q. Cheng, R. Chen, and T.-H Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 377–384, Mar. 1996.

[8] C. L. Degen, M. Poggio, H. J. Mamin, C. T. Rettner, and D. Rugar, "Nanoscale magnetic resonance imaging," *Proc. Nat. Acad. Sci.*, vol. 106, no. 5, pp. 1313–1317, Feb. 2009.

[9] C. L. Degen, M. Poggio, H. J. Mamin, C. T. Rettner, and D. Rugar, "Nanoscale magnetic resonance imaging. Supporting information," *Proc. Nat. Acad. Sci.*, vol. 106, no. 5, Feb. 2009.

[10] J. Diebolt and E. H. S. Ip, "Stochastic EM: method and application," in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. London, U.K.: Chapman & Hall, 1996, pp. 259–273.

[11] N. Dobigeon and J.-Y. Tourneret, "Joint segmentation of wind speed and direction using a hierarchical model," *Comput. Stat. Data Anal.*, vol. 51, no. 12, pp. 5603–5621, Aug. 2007.

[12] N. Dobigeon and J.-Y. Tourneret, "Bayesian sampling of structured noise covariance matrix for hyperspectral imagery," Tech. Rep., Univ. Toulouse, France, Dec. 2008 [Online]. Available: http://dobigeon.perso.enseeiht.fr/publis.html

[13] 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 Process.*, vol. 56, no. 7, pp. 2684–2695, Jul. 2008.

[14] N. Dobigeon, J.-Y. Tourneret, and M. Davy, "Joint segmentation of piecewise constant autoregressive processes processes by using a hierarchical model and a Bayesian sampling approach," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1251–1263, Apr. 2007.

[15] 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 Process.*, vol. 55, no. 2, pp. 414–423, Feb. 2007.

[16] A. Doucet and P. Duvaut, "Bayesian estimation of state-space models applied to deconvolution of Bernoulli-Gaussian processes," *Signal Process.*, vol. 57, no. 2, pp. 147–161, Mar. 1997.

[17] C. Févotte, B. Torrésani, L. Daudet, and S. J. Godsill, "Sparse linear regression with structured priors and application to denoising of musical audio," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 174–185, Jan. 2008.

[18] J. Geweke, E. M. Keramidas, Ed., "Efficient simulation from the multivariate normal and student-T distributions subject to linear constraints," in *Proc. 23rd Symp. Interface Computing Science and Statistics*, Fairfax, VA, 1991, pp. 571–578.

[19] S. J. Godsill and P. J. W. Rayner, "Statistical reconstruction and analysis of autoregressive signals in impulsive noise using the Gibbs sampler," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 4, pp. 352–372, Jul. 1998.

[20] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.

[21] P. C. Hammel, D. V. Pelekhov, P. E. Wigen, T. R. Gosnell, M. M. Midzor, and M. L. Roukes, "The magnetic-resonance force microscope: A new tool for high-resolution, 3-D, subsurface scanned probe imaging," *Proc. IEEE*, vol. 91, no. 5, pp. 789–798, May 2003.

[22] K. Herrity, R. Raich, and A. O. Hero, "Blind deconvolution for sparse molecular imaging," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Las Vegas, NV, Apr. 2008, pp. 545–548.

[23] K. Herrity, R. Raich, and A. O. Hero, C. A. Bouman, E. L. Miller, and I. Pollak, Eds., "Blind reconstruction of sparse images with unknown point spread function," in *Proc. Computational Imaging Conf. IS&TSPIE Symp. Electronic Imaging Science and Technology*, San Jose, CA, USA, Jan. 2008, vol. 6814, pp. 68140K-1–68140K-11.

[24] J. Idier and Y. Goussard, "Stack algorithm for recursive deconvolution of Bernoulli-Gaussian processes," *IEEE Trans. Signal Process.*, vol. 28, no. 9, pp. 67–79, Sep. 1990.

[25] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proc. Roy. Soc. Lond. A*, vol. 186, no. 1007, pp. 453–461, 1946.

[26] H. Jeffreys, *Theory of Probability*, 3rd ed. London, U.K.: Oxford Univ. Press, 1961.

[27] I. M. Johnstone and B. W. Silverman, "Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences," *Ann. Statist.*, vol. 32, no. 4, pp. 1594–1649, 2004.

[28] 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.

[29] S. Kuehn, S. A. Hickman, and J. A. Marohn, "Advances in mechanical detection of magnetic resonance," *J. Chem. Phys.*, vol. 128, no. 5, pp. 052208-1–052208-19, Feb. 2008.

[30] E. Kuhn and M. Lavielle, "Coupling a stochastic approximation version of EM with an MCMC procedure," *ESAIM Probab. Statist.*, vol. 8, pp. 115–131, 2004.

[31] L. Landweber, "An iteration formula for fredholm integral equations of the first kind," *Amer. J. Math.*, vol. 73, no. 3, pp. 615–624, Jul. 1951.

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

[33] M. Lavielle, "Bayesian deconvolution of Bernoulli-Gaussian processes," *Signal Process.*, vol. 33, no. 1, pp. 67–79, Jul. 1993.

[34] J. Mamin, R. Budakian, and D. Rugar, "Point response function of an MRFM tip," Tech. Rep., IBM Research Division, Oct. 2003.
[35] J.-M. Marin and C. P. Robert, *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. New York, NY: Springer, 2007.
[36] V. Mazet, D. Brie, and J. Idier, "Simulation of postive normal variables using several proposal distributions," in *Proc. IEEE Workshop on Statistical Signal Processing (SSP)*, Bordeaux, France, Jul. 2005, pp. 37–42.
[37] V. Mazet, "Développement De Méthodes De Traitement De Signaux Spectroscopiques: Estimation De La Ligne De Base Et Du Spectre De Raies," Ph.D. dissertation, Univ. Henri Poincaré, Nancy, France, 2006.
[38] D. Mounce, "Magnetic resonance force microscopy," *IEEE Instrum. Meas. Mag.*, vol. 8, no. 2, pp. 20–26, Jun. 2005.
[39] T. Olofsson and E. Wennerström, "Sparse deconvolution of B-scan images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, no. 8, pp. 1634–1641, Aug. 2007.
[40] E. Punskaya, C. Andrieu, A. Doucet, and W. Fitzgerald, "Bayesian curve fitting using MCMC with applications to signal segmentation," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 747–758, Mar. 2002.
[41] S. E. Reichenbach, D. E. Koehler, and D. W. Strelow, "Restoration and reconstruction of AVHRR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 997–1007, Jul. 1995.
[42] C. P. Robert, "Simulation of truncated normal variables," *Statist. Comput.*, vol. 5, no. 2, pp. 121–125, Jun. 1995.
[43] C. P. Robert, "The Bayesian choice: from decision-theoretic motivations to computational implementation," in *Ser. Springer Texts in Statistics*, 2nd ed. New York: Springer-Verlag, 2007.
[44] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer, 2004.
[45] O. Rosec, J.-M. Boucher, B. Nsiri, and T. Chonavel, "Blind marine seismic deconvolution using statistical MCMC methods," *IEEE J. Ocean. Eng.*, vol. 28, no. 3, pp. 502–512, Jul. 2003.
[46] 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, Jul. 2004.
[47] J. C. Russ, *The Image Processing Handbook*, 5th ed. Boca Raton, FL: CRC, 2006.
[48] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 32–45, May 2006.
[49] J.-L. Starck and F. Murtagh, *Astronomical Image and Data Analysis*, 2nd ed. Berlin, Germany: Springer-Verlag, 2006.
[50] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, Nov. 1981.
[51] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Statist. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.
[52] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction using sparse priors," in *Proc. IEEE Int. Conf. Image Processing*, Oct. 2006, pp. 1261–1264.
[53] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction for molecular imaging," *IEEE Trans. Image Process.*, 2009.
[54] M. Y. Ting, "Signal Processing for Magnetic Resonance Force Microscopy," Ph.D. dissertation, Univ. Michigan, Ann Arbor, 2006.
[55] F, Sroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1094–1106, Sep. 2003.
[56] O. Züger, S. T. Hoen, C. S. Yannoni, and D. Rugar, "Three-dimensional imaging with a nuclear magnetic resonance force microscope," *J. Appl. Phys.*, vol. 79, no. 4, pp. 1881–1884, Feb. 1996.
[57] O. Züger and D. Rugar, "First images from a magnetic resonance force microscope," *Appl. Phys. Lett.*, vol. 63, no. 18, pp. 2496–2498, 1993.
[58] O. Züger and D. Rugar, "Magnetic resonance detection and imaging using force microscope techniques," *J. Appl. Phys.*, vol. 75, no. 10, pp. 6211–6216, May 1994.

Nicolas Dobigeon (S'05-M'08) was born in Angoulême, France, in 1981. He received the Eng. degree in electrical engineering from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse, both in 2004. In 2007, he received the Ph.D. degree in signal processing also from the National Polytechnic Institute of Toulouse.

From 2007 to 2008, he was a postdoctoral research associate at the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. Since 2008, he has been an Assistant Professor with the National Polytechnic Institute of Toulouse (ENSEEIHT—University of Toulouse), within the Signal and Communication Group of the IRIT Laboratory. His research interests are centered around statistical signal and image processing with a particular interest to Bayesian inference and Markov chain Monte Carlo methods.

Alfred O. Hero (S'79–M'84–SM'96–F'98) received the B.S. degree (summa cum laude) from Boston University, Boston, MA, in 1980, and the Ph.D. degree from Princeton University, Princeton, NJ, 1984, both in electrical engineering.

Since 1984, he has been with the University of Michigan, Ann Arbor, where he is the R. Jamison and Betty Professor of engineering. His primary appointment is in the Department of Electrical Engineering and Computer Science and he also has appointments, by courtesy, in the Department of Biomedical Engineering and the Department of Statistics. In 2008, he was awarded the the Digiteo Chaire d'Excellence, sponsored by Digiteo Research Park in Paris, located at the Ecole Superieure d'Electricite, Gif-sur-Yvette, France. He has held other visiting positions at LIDS Massachussets Institute of Technology (2006), Boston University (2006), I3S University of Nice, Sophia-Antipolis, France (2001), Ecole Normale Supérieure de Lyon (1999), Ecole Nationale Supérieure des Télécommunications, Paris (1999), Lucent Bell Laboratories (1999), Scientific Research Labs of the Ford Motor Company, Dearborn, MI (1993), Ecole Nationale Superieure des Techniques Avancees (ENSTA), Ecole Superieure d'Electricite, Paris (1990), and MIT Lincoln Laboratory (1987–1989). His recent research interests have been in detection, classification, pattern analysis, and adaptive sampling for spatio-temporal data. Of particular interest are applications to network security, multimodal sensing and tracking, biomedical imaging, and genomic signal processing.

Dr. Hero is member of Tau Beta Pi, the American Statistical Association (ASA), the Society for Industrial and Applied Mathematics (SIAM), and the U.S. National Commission (Commission C) of the International Union of Radio Science (URSI). He has received a IEEE Signal Processing Society Meritorious Service Award (1998), an IEEE Signal Processing Society Best Paper Award (1998), a IEEE Third Millenium Medal, and a 2002 IEEE Signal Processing Society Distinguished Lecturership. He was President of the IEEE Signal Processing Society (2006–2007) and during his term served on the TAB Periodicals Committee (2006). He was a member of the IEEE TAB Society Review Committee (2008) and is Director-elect of IEEE for Division IX (2009).

Jean-Yves Tourneret (M'94–SM'08) received the ingénieur degree in electrical engineering from École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique et d'Hydraulique, Toulouse (ENSEEIHT), France, in 1989, and the Ph.D. degree from the National Polytechnic Institute of Toulouse in 1992.

He is currently a Professor at the ENSEEIHT and a member of the IRIT laboratory (UMR 5505, CNRS). His research activities are centered around statistical signal processing with a particular interest to classification and Markov Chain Monte Carlo methods.

Dr. Tourneret was the program chair of the European conference on signal processing (EUSIPCO), which was held in Toulouse in 2002. He was also member of the organizing committee for the international conference ICASSP'06 which was held in Toulouse in 2006. He has been a member of different technical committees including the Signal Processing Theory and Methods (SPTM) committee of the IEEE Signal Processing Society (2001–2007). He is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.