

Myopic sparse image reconstruction with application to MRFM

Se Un Park^{a,*}, Nicolas Dobigeon^{b,*} and Alfred O. Hero^a

^aUniversity of Michigan, EECS Department, Ann Arbor, USA;

^bUniversity of Toulouse, IRIT/INP-ENSEEIH, Toulouse, France

ABSTRACT

We propose a solution to the image deconvolution problem where the convolution operator or point spread function (PSF) is assumed to be only partially known. Small perturbations generated from the model are exploited to produce a few principal components explaining the uncertainty in a high dimensional space. Specifically, we assume the image is sparse corresponding to the natural sparsity of magnetic resonance force microscopy (MRFM). Our approach adopts a Bayesian Metropolis-within-Gibbs sampling framework. The performance of our Bayesian myopic algorithm is superior to previously proposed algorithms such as the alternating minimization (AM) algorithm for sparse images. We illustrate our myopic algorithm on real MRFM tobacco virus data.

Keywords: Myopic deconvolution, Bayesian inference, Markov Chain Monte Carlo (MCMC) methods, MRFM experiment

1. INTRODUCTION

Recently, a new 3D imaging technology called magnetic resonance force microscopy (MRFM) has been developed. The principles of MRFM were introduced by Sides¹⁻³ who described its potential for achieving 3D atomic scale resolution. In 1992 and 1996, Rugar *et al.*^{4,5} reported experiments that demonstrated the practicality of MRFM and produced the first MRFM images. More recently, MRFM volumetric spatial resolutions of less than 10nm have been demonstrated for imaging a biological sample.⁶ The signal provided by MRFM is a so-called force map that is the 3D convolution of the atomic spin distribution and the point spread function (PSF).⁷ This formulation casts the estimation of the spin density from the force map as an inverse problem. Several approaches have been proposed to solve this inverse problem, i.e., to reconstruct the unknown image from the measured force map. Basic algorithms rely on Wiener^{5,8,9} filters whereas others are based on iterative least squares reconstruction approaches.^{6,7,10} More recently, promising works addressed this problem within the Bayesian estimation framework.^{11,12}

However, all of these reconstruction techniques require prior knowledge of the device response, namely the PSF. As shown by Mamin *et al.*,¹³ this PSF is a function of several parameters specified by the physical model of the device. Unfortunately, in many practical situations of MRFM imaging, the physical parameters that tune the response of the MRFM tip are only partially known, even totally unknown. In such circumstances, the PSF used in the reconstruction algorithm is mismatched to the true PSF and the quality of standard image reconstruction technique will suffer if one does not account for this mismatch. Estimating the unknown image and the PSF jointly is usually referred to as semi-blind^{14,15} or myopic^{16,17} deconvolution, and this is the approach taken in this paper.

This work was partially supported by ARO, grant number W911NF-05-1-0403.

Further author information: (Send correspondence to A.O.H.)

S.U.P.: E-mail: seunpark@umich.edu

N.D.: E-mail: nicolas.dobigeon@enseeiht.fr

A.O.H.: E-mail: hero@umich.edu

*contributed equally to this work.

To mitigate the effects of PSF mismatch on MRFM image reconstruction, an alternating minimization (AM) algorithm¹⁸ was proposed by Herrity *et al.* which showed robust performance but does not account for prior information on the PSF or on the image, e.g., through Bayesian priors. In this paper, we propose a hierarchical Bayesian approach to myopic image deconvolution that uses prior information on the PSF model. We propose a myopic modification of the Bayesian MRFM reconstruction approach in Dobigeon *et al.*,¹² whereby one performs a simple additional step in the initial Gibbs sampler, producing a Bayesian estimate of the PSF and a Bayesian reconstruction of the image. Our approach can be related to the recent paper of Orioux *et al.*¹⁹ who introduced a Metropolis-within-Gibbs algorithm to estimate the parameters that tune the device response. This strategy focuses on reconstruction with smoothness constraints and requires recomputation of the entire PSF at each step of the algorithm. This is computationally expensive, especially for complex PSF models such as in the MRFM instrument. Here, we propose an alternative that consists of estimating the deviation from a given nominal PSF. More precisely, the nominal point response of the device is assumed known and the true PSF is modeled as a small perturbation about the nominal response. We approximate the full posterior distribution of the PSF and the image using samples generated by a Markov Chain Monte Carlo algorithm. Simulations are presented that quantify the advantages of our algorithm for myopic sparse image reconstruction. We then apply it to the real MRFM tobacco virus data made available by our IBM collaborators.

This paper is organized as follows: Section 2 formulates the problem. Section 3 covers Bayesian framework of image modeling and the following Section 4 proposes a solution in this framework. Section 5 shows simulation results and an application to the real MRFM data.

2. PROBLEM FORMULATION

We adopt the notation of Dobigeon *et al.*¹² Let \mathbf{X} denote the $l_1 \times \dots \times l_n$ unknown n -D original image to be recovered (e.g. $n = 2$ or $n = 3$). This image is to be reconstructed from a collection of P measurements $\mathbf{y} = [y_1, \dots, y_P]^T$ via the following noisy transformation:

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

where $T(\cdot, \cdot)$ is the mean response function $\mathbb{E}[\mathbf{y}|\boldsymbol{\kappa}, \mathbf{x}]$, \mathbf{n} is an $P \times 1$ observation noise vector and $\boldsymbol{\kappa}$ is the kernel modeling the response of the imaging device. A typical PSF for MRFM is shown in Mamin *et al.*¹³ for horizontal and vertical MRFM tip configurations. In (1), \mathbf{n} is an additive Gaussian noise sequence distributed according to $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_P)$. The PSF is assumed to be known up to a perturbation $\Delta\boldsymbol{\kappa}$ about a known nominal $\boldsymbol{\kappa}_0$:

$$\boldsymbol{\kappa} = \boldsymbol{\kappa}_0 + \Delta\boldsymbol{\kappa}. \quad (2)$$

In the MRFM application the PSF is described by an approximate parametric function that depends on the experimental setup. Based on the physical parameters tuned during the experiment (external magnetic field, mass of the probe, etc.), an approximation $\boldsymbol{\kappa}_0$ of the PSF can be derived. However, due to model mismatch and experimental errors, the true PSF $\boldsymbol{\kappa}$ may deviate from the nominal PSF $\boldsymbol{\kappa}_0$. Our model for this deviation is that $\Delta\boldsymbol{\kappa}$ can be expressed as a linear combination of elements of an *a priori* known basis \mathbf{v}_k , $k = 1, \dots, K$,

$$\Delta\boldsymbol{\kappa} = \sum_{k=1}^K \lambda_k \mathbf{v}_k, \quad (3)$$

where $\{\mathbf{v}_k\}_{k=1, \dots, K}$ is a set of basis functions for the PSF perturbations and λ_k , $k = 1, \dots, K$ are unknown coefficients. To emphasize the influence of these coefficients on the actual PSF, $\boldsymbol{\kappa}$ will be denoted $\boldsymbol{\kappa}(\boldsymbol{\lambda})$ with $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_K]^T$.

Note that in standard deblurring, the function $T(\cdot, \cdot)$ represents the standard n D-convolution operator \otimes . In this case, the image \mathbf{X} can be vectorized as $\mathbf{x} \in \mathbb{R}^M$ with $M = P = l_1 l_2 \dots l_n$. With these notations, (1) can be rewritten:

$$\mathbf{y} = \mathbf{H}(\boldsymbol{\lambda}) \mathbf{x} + \mathbf{n} \quad \text{or} \quad \mathbf{Y} = \boldsymbol{\kappa}(\boldsymbol{\lambda}) \otimes \mathbf{X} + \mathbf{N} \quad (4)$$

where \mathbf{y} (resp. \mathbf{n}) stands for the vectorized version of \mathbf{Y} (resp. \mathbf{N}) and $\mathbf{H}(\boldsymbol{\lambda})$ is a $P \times M$ matrix that describes the convolution by the PSF kernel $\boldsymbol{\kappa}(\boldsymbol{\lambda})$.

We address the problem of estimating the unobserved image \mathbf{x} and the PSF perturbation $\Delta\boldsymbol{\kappa}$ under sparsity constraints given the measurement \mathbf{y} and the bilinear function $T(\cdot, \cdot)$.

3. HIERARCHICAL BAYESIAN MODEL

3.1 Likelihood function

Under the hypothesis that the noise in (1) is Gaussian, the observation model likelihood function takes the form

$$f(\mathbf{y}|\mathbf{x}, \boldsymbol{\lambda}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{P}{2}} \exp\left(-\frac{\|\mathbf{y} - T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x})\|^2}{2\sigma^2}\right), \quad (5)$$

where $\|\cdot\|$ denotes the standard ℓ_2 norm: $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$. This function will be denoted $f(\mathbf{y}|\boldsymbol{\theta})$, where $\boldsymbol{\theta} = \{\mathbf{x}, \boldsymbol{\lambda}, \sigma^2\}$.

3.2 Parameter prior distributions

In this section, we introduce prior distributions for the parameters $\boldsymbol{\theta}$. Except PSF parameter prior, the rest parameter prior distributions are given in Dobigeon *et al.*¹²

3.2.1 Image prior

As the prior distribution for x_i , we adopt a mixture of a mass at zero and a single-sided exponential distribution:

$$f(x_i|w, a) = (1-w)\delta(x_i) + \frac{w}{a} \exp\left(-\frac{x_i}{a}\right) \mathbf{1}_{\mathbb{R}_+^*}(x_i), \quad (6)$$

where $w \in [0, 1]$, $a \in [0, \infty)$, $\delta(\cdot)$ is the Dirac function, \mathbb{R}_+^* is a set of real open interval $(0, \infty)$ and $\mathbf{1}_{\mathbb{E}}(x)$ is the indicator function of the set \mathbb{E} :

$$\mathbf{1}_{\mathbb{E}}(x) = \begin{cases} 1, & \text{if } x \in \mathbb{E}, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

By assuming the components x_i to be a conditionally independent ($i = 1, \dots, M$) given w, a, σ , the following conditional prior distribution is obtained for the image \mathbf{x} :

$$f(\mathbf{x}|w, a) = \prod_{i=1}^M \left[(1-w)\delta(x_i) + \frac{w}{a} \exp\left(-\frac{x_i}{a}\right) \mathbf{1}_{\mathbb{R}_+^*}(x_i) \right]. \quad (8)$$

This image prior is similar to the LAZE distribution (weighted average of a Laplacian pdf and an atom at zero) used, for example, in Ting *et al.*^{11,20} As motivated by Dobigeon *et al.*,¹² the image prior in (6) has the interesting property of enforcing the pixel value to be zero, reflecting the natural sparsity of the MRFM images. Furthermore, the proposed prior in (6) ensures positivity of the pixel values (spin density) to be estimated.

3.2.2 PSF parameter prior

We assume that the parameters $\lambda_1, \dots, \lambda_K$ are *a priori* independent and uniformly distributed over known intervals associated with some error tolerances centered at 0. Define the interval

$$\mathcal{S}_k = [-\Delta\lambda_k, \Delta\lambda_k] \quad (9)$$

and assume the distribution of $\boldsymbol{\lambda}$ has density

$$f(\boldsymbol{\lambda}) = \prod_{k=1}^K \frac{1}{2\Delta\lambda_k} \mathbf{1}_{\mathcal{S}_k}(\lambda_k), \quad (10)$$

with $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_K]^T$.

3.2.3 Noise variance prior

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

$$\sigma^2 | \nu, \gamma \sim \mathcal{IG} \left(\frac{\nu}{2}, \frac{\gamma}{2} \right). \quad (11)$$

In the following, ν will be fixed to $\nu = 2$ and γ will be an hyperparameter to be estimated (see^{21–23} for similar choices).

3.3 Hyperparameter priors

Define the hyperparameter vector associated with the image and noise variance prior distributions as $\Phi = \{a, \gamma, w\}$. In our hierarchical Bayesian framework, the estimation of these hyperparameters requires prior distributions in the hyperparameters. These priors are defined in Dobigeon *et al.*¹² but for completeness of this paper brief definitions of them are reproduced below.

3.3.1 Hyperparameter a

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

$$a | \alpha \sim \mathcal{IG} (\alpha_0, \alpha_1), \quad (12)$$

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

3.3.2 Hyperparameter γ

A non informative Jeffreys' prior is selected as prior distribution for hyperparameter γ :

$$f(\gamma) \propto \frac{1}{\gamma} \mathbf{1}_{\mathbb{R}^+}(\gamma). \quad (13)$$

3.3.3 Hyperparameter w

A uniform distribution on the simplex $[0, 1]$ is selected as prior distribution for the mean proportion of non-zero pixels:

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

Assuming that the individual hyperparameters are independent the full hyperparameter prior distribution for Φ can be expressed as:

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

3.4 Posterior distribution

The posterior distribution of $\{\theta, \Phi\}$ is:

$$f(\theta, \Phi | \mathbf{y}) \propto f(\mathbf{y} | \theta) f(\theta | \Phi) f(\Phi), \quad (16)$$

with

$$f(\theta | \Phi) = f(\mathbf{x} | \lambda, \sigma^2, a, w) f(\lambda) f(\sigma^2 | \gamma), \quad (17)$$

where $f(\mathbf{y}|\boldsymbol{\theta})$ and $f(\boldsymbol{\Phi})$ have been defined in (5) and (15). This hierarchical structure allows one to integrate out the parameters σ^2 , and the hyperparameter $\boldsymbol{\Phi}$ in the full posterior distribution (16), yielding:

$$f(\mathbf{x}, \boldsymbol{\lambda} | \mathbf{y}, \alpha_0, \alpha_1) \propto \frac{\mathcal{B}e(\alpha_1 + n_1, \alpha_0 + n_0) \Gamma(n_1)}{\|\mathbf{y} - T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x})\|^N \|\mathbf{x}\|_1^{n_1}} \prod_{k=1}^K \frac{1}{2\Delta\lambda_k} \mathbf{1}_{\mathcal{S}_k}(\lambda_k), \quad (18)$$

where $\mathcal{B}e$ is the beta function and Γ is the gamma function.

The next section presents the Metropolis-within-Gibbs algorithm²⁴ that generates samples distributed according to the posterior distribution $f(\mathbf{x}, \boldsymbol{\lambda} | \mathbf{y})$. These samples are then used to estimate \mathbf{x} and $\boldsymbol{\lambda}$.

4. METROPOLIS-WITHIN-GIBBS ALGORITHM FOR MYOPIC SPARSE IMAGE RECONSTRUCTION

We describe in this section a Metropolis-within-Gibbs sampling strategy that allows one to generate samples $\{\mathbf{x}^{(t)}, \boldsymbol{\lambda}^{(t)}\}_{t=1, \dots}$ distributed according to the posterior distribution in (18). As sampling directly from (18) is a difficult task, we will instead generate samples distributed according to the joint posterior $f(\mathbf{x}, \boldsymbol{\lambda}, \sigma^2 | \mathbf{y}, \alpha_0, \alpha_1)$. Sampling from this posterior distribution is done by alternatively sampling one of $\mathbf{x}, \boldsymbol{\lambda}, \sigma^2$ conditioned on all other variables.^{12, 25}

The main steps of our proposed sampling algorithm are given in subsections 4.1 through 4.3 (see also Algorithm 1).

Algorithm 1 Metropolis-within-Gibbs sampling algorithm for myopic sparse image reconstruction

- 1: % Initialization:
 - 2: Sample the unknown image $\mathbf{x}^{(0)}$ from pdf in (8),
 - 3: Sample the noise variance $\tilde{\sigma}^{2(0)}$ from the pdf in (11),
 - 4: % Iterations:
 - 5: **for** $t = 1, 2, \dots$, **do**
 - 6: Sample hyperparameter $w^{(t)}$ from the pdf in (20),
 - 7: Sample hyperparameter $a^{(t)}$ from the pdf in (21),
 - 8: For $i = 1, \dots, M$, sample the pixel intensity $x_i^{(t)}$ from the pdf in (22),
 - 9: For $k = 1, \dots, K$, sample the PSF parameter $\lambda_k^{(t)}$ from the pdf in (23) (see Algo. 2),
 - 10: Sample the noise variance $\tilde{\sigma}^{2(t)}$ from the pdf in (26),
 - 11: **end for**
-

4.1 Generation of samples according to $f(\mathbf{x} | \boldsymbol{\lambda}, \sigma^2, \mathbf{y}, \alpha_0, \alpha_1)$

To generate samples distributed according to $f(\mathbf{x} | \boldsymbol{\lambda}, \sigma^2, \mathbf{y}, \alpha_0, \alpha_1)$, it is convenient to sample according to $f(\mathbf{x}, w, a | \boldsymbol{\lambda}, \sigma^2, \mathbf{y}, \alpha_0, \alpha_1)$ by the following 3-step procedure.

4.1.1 Generation of samples according to $f(w | \mathbf{x}, \alpha_0, \alpha_1)$

The conditional posterior distribution of w is

$$f(w | \mathbf{x}, \alpha_0, \alpha_1) \propto (1 - w)^{n_0 + \alpha_0 - 1} w^{n_1 + \alpha_1 - 1}, \quad (19)$$

where $n_1 = \|\mathbf{x}\|_0$ and $n_0 = M - \|\mathbf{x}\|_0$. Therefore, generation of samples according to $f(w | \mathbf{x})$ is achieved as follows:

$$w | \mathbf{x}, \alpha_0, \alpha_1 \sim \mathcal{B}e(\alpha_1 + n_1, \alpha_0 + n_0). \quad (20)$$

4.1.2 Generation of samples according to $f(a|\mathbf{x})$

The joint posterior distribution (16) yields:

$$a|\mathbf{x} \sim \mathcal{IG}(\|\mathbf{x}\|_0, \|\mathbf{x}\|_1). \quad (21)$$

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

The posterior distribution of each component x_i ($i = 1, \dots, M$) given all other variables is easily derived:¹²

$$f(x_i|w, a, \lambda, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}) \propto (1 - w_i)\delta(x_i) + w_i\phi_+(x_i|\mu_i, \eta_i^2), \quad (22)$$

where \mathbf{x}_{-i} stands for the vector \mathbf{x} whose i th component has been removed and μ_i and η_i^2 are given in equation (36) in Dobigeon *et al.*¹² In (22), $\phi_+(\cdot, m, s^2)$ stands for the pdf of the truncated Gaussian distribution defined on \mathbb{R}_+^* with hidden mean m and hidden variance s^2 . Therefore, from (22), $x_i|w, a, \lambda, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}$ is a Bernoulli-truncated Gaussian variable with parameter (w_i, μ_i, η_i^2) .

To summarize, generation of samples distributed according to $f(\mathbf{x}|\sigma^2, a, \lambda, \mathbf{y})$ can be performed by updating the coordinates of \mathbf{x} using M Gibbs moves (requiring to generate Bernoulli-truncated Gaussian variables). An efficient simulation scheme is presented in Appendix A.

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

The posterior distribution of the parameter λ_k conditioned on the unknown image \mathbf{x} , the noise variance σ^2 and the other PSF parameters $\{\lambda_j\}_{j \neq k}$ is

$$f(\lambda_k|\boldsymbol{\lambda}_{-k}, \mathbf{x}, \sigma^2, \mathbf{y}) \propto \exp\left[-\frac{\|\mathbf{y} - T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x})\|^2}{2\sigma^2}\right] \mathbf{1}_{S_k}(\lambda_k), \quad (23)$$

with $\boldsymbol{\lambda}_{-k} = \{\lambda_j\}_{j \neq k}$. We summarize in Algorithm 2 a procedure for generating samples distributed according to the posterior in (23) using a simple acceptance-rejection sampling and random walk procedure²⁴ with a proposed centered Gaussian distribution having variance s_k^2 . Thus, at iteration t of the algorithm, the acceptance probability of a proposed state λ_k^* is:

$$\rho_{\lambda_k^{(t)} \rightarrow \lambda_k^*} = \min(1, a_k \mathbf{1}_{S_k}(\lambda_k^*)), \quad (24)$$

with

$$\log a_k = \frac{1}{2\sigma^2} \left(\left\| \mathbf{y} - T(\boldsymbol{\kappa}(\lambda_k^{(t)}), \mathbf{x}) \right\|^2 \right) - \frac{1}{2\sigma^2} \left(\left\| \mathbf{y} - T(\boldsymbol{\kappa}(\lambda_k^*), \mathbf{x}) \right\|^2 \right). \quad (25)$$

Computing the transformation $T(\cdot, \cdot)$ at each step of the sampler can be computationally costly. Appendix A provides a recursive strategy to sample according to $f(\boldsymbol{\lambda}|\mathbf{x}, \sigma^2, \mathbf{y})$ efficiently.

Algorithm 2 Sampling according to $f(\lambda_k|\boldsymbol{\lambda}_{-k}, \mathbf{x}, \sigma^2, \mathbf{y})$

- 1: Sample $\varepsilon \sim \mathcal{N}(0, s_p^2)$,
- 2: Propose λ_k^* according to $\lambda_k^* = \lambda_k^{(t)} + \varepsilon$,
- 3: Draw $w_k \sim \mathcal{U}([0, 1])$,
- 4: Set $\lambda_k^{(t+1)} = \begin{cases} \lambda_k^*, & \text{if } w_k \leq \rho_{\lambda_k^{(t)} \rightarrow \lambda_k^*}, \\ \lambda_k^{(t)}, & \text{otherwise.} \end{cases}$

where $\mathcal{U}(\mathbb{E})$ stands for the uniform distribution on the set \mathbb{E} .

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

Samples $(\sigma^2)^{(t)}$ are generated according to the posterior

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

5. EXPERIMENTS

In this section we present simulation results that compare the proposed myopic Bayesian deconvolutions with the AM algorithm¹⁸ and the non-myopic Bayesian method.¹² Here an nominal PSF κ_0 was assumed such that it corresponds to the mathematical MRFM point response model proposed by Mamin *et al.*¹³ This nominal PSF is used in AM algorithm and the parameter values of AM algorithm were set empirically according to the procedure in Herrity *et al.*¹⁸

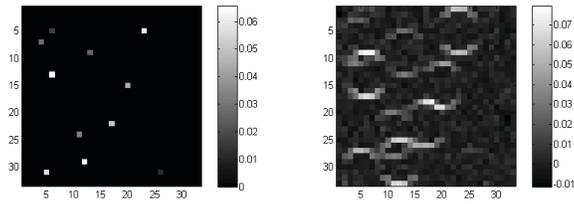
5.1 Simulation on synthetic sparse images

We performed simulations of MRFM measurements for PSF and image models similar to those described in Dobigeon *et al.*¹² The signal-to-noise ratio was set to SNR = 10dB. Several 32×32 synthetic sparse images, one of which is depicted in Fig. 1(a), were used to produce the data and were estimated using the proposed Bayesian method. The assumed PSF κ_0 , generated following the physical model described in Mamin *et al.*¹³ when the physical parameters are tuned to the values displayed in Table 1, is a 11×11 2-dimensional convolution kernel, represented in Fig. 2(a). We assume that the true PSF κ comes from the same physical model where the radius of the tip and the distance from tip to sample have been mis-specified as values $R = R_0 - 2\% = 3.92$ and $d = d_0 + 2\% = 6.12$, leading to the convolution kernel depicted in Fig. 2(b). The observed measurements \mathbf{y} , shown Fig. 1(b) are a 32×32 image of size $P = 1024$.

Table 1. Parameters used to compute the MRFM PSF.

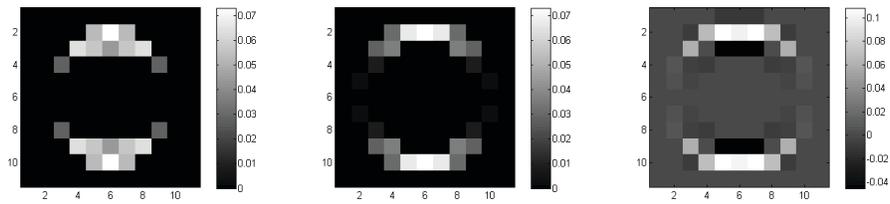
Parameter		Value
Description	Name	
Amplitude of external magnetic field	B_{ext}	9.4×10^3 G
Value of B_{mag} in the resonant slice	B_{res}	1.0×10^4 G
Radius of tip	R_0	4.0 nm
Distance from tip to sample	d_0	6.0 nm
Cantilever tip moment	m	4.6×10^5 emu
Peak cantilever oscillation	x_{pk}	0.8 nm
Maximum magnetic field gradient	G_{max}	125

The proposed algorithm requires the definition of K basis vectors \mathbf{v}_k , $k = 1, \dots, K$, that span a subspace representing possible perturbations $\Delta\boldsymbol{\kappa}$. We empirically determined this basis using the following PSF variation eigendecomposition approach. A set of 5000 experimental PSFs $\tilde{\boldsymbol{\kappa}}_j$, $j = 1, \dots, 5000$, were generated following the model described in Mamin *et al.*¹³ with parameters d and R randomly drawn according to Gaussian distribution centered at the nominal values d_0 , R_0 , respectively. Then a standard principal component analysis (PCA) of the residuals $\{\tilde{\boldsymbol{\kappa}}_j - \boldsymbol{\kappa}_0\}_{j=1, \dots, 5000}$ is used to identify $K = 4$ principal axes that are associated with the basis vectors \mathbf{v}_k . The necessary number of basis vectors, $K = 4$ here, is determined empirically by looking at the scree plot in Fig. 3(b) which explains proportions of the variance of the perturbations. The first four eigen-functions,



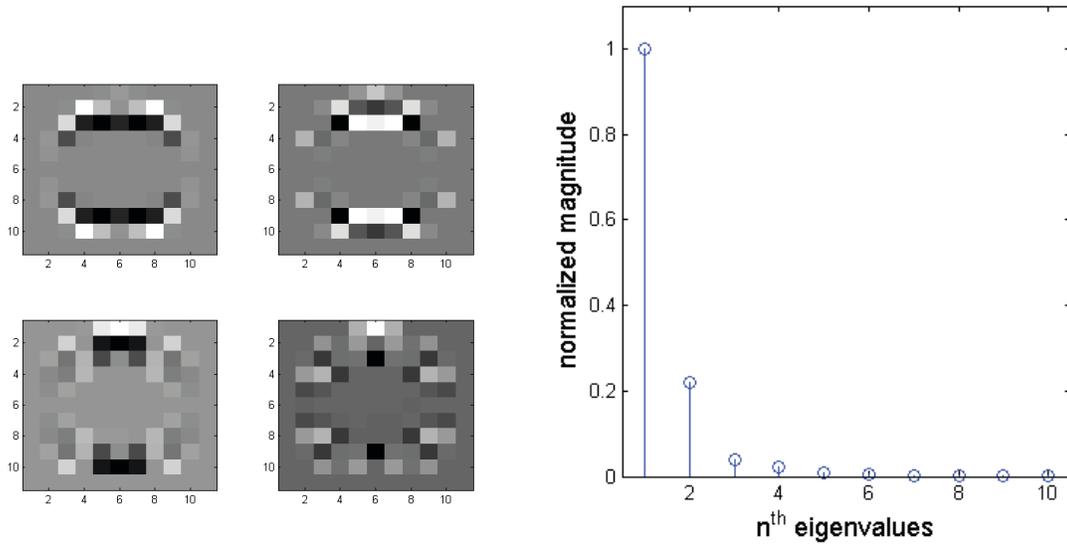
(a) Sparse true image ($\|\mathbf{x}\|_0 = 11$) (b) Noisy observation

Figure 1. True image and noisy observation



(a) Assumed PSF (b) True PSF (c) Estimated PSF

Figure 2. Assumed PSF κ_0 (left), actual PSF κ (middle), and estimated PSF $\hat{\kappa}$ (right) of MRFM tip.



(a) The $K = 4$ patterns \mathbf{v}_k identified by PCA. (b) Scree plot (magnitude is normalized up to the largest value, i.e. $\lambda_{max} := 1$.)

Figure 3. PCA of the PSF perturbations.

corresponding to the first four largest eigenvalues, seemed enough to explain major perturbations. The patterns of basis vectors are depicted in Fig. 3(a).

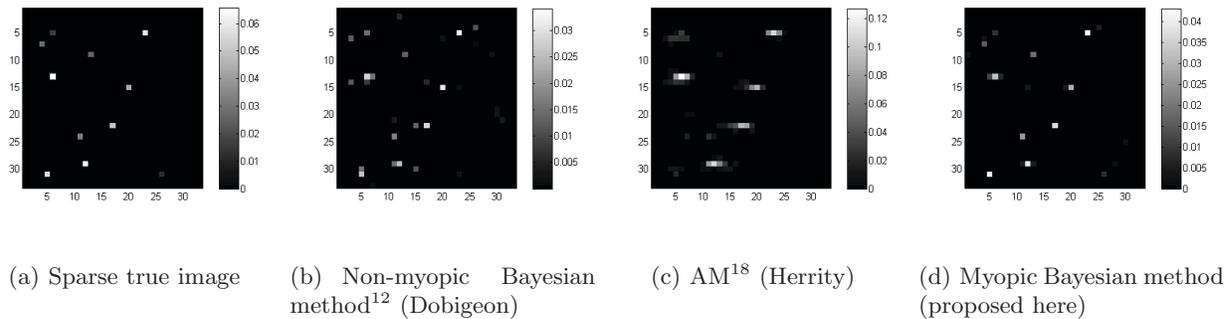


Figure 4. Estimated image from previous Bayesian, AM, and myopic Bayesian methods.

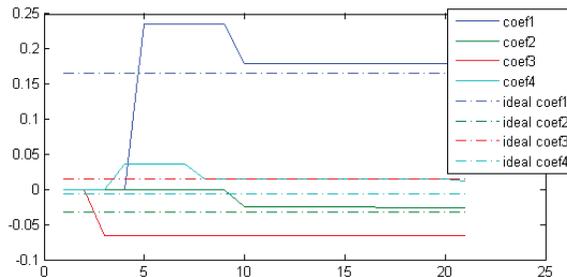


Figure 5. Estimated PSF coefficients for 4 PCs over 21 iterations

The proposed Bayesian algorithm was applied to estimate both the sparse image and the PSF coefficients of \mathbf{v}_k 's, using the prior in (6) with parameter $a = 1$ and $w = 0.02$. From the observation in Fig. 1(b) the PSF estimated by the proposed algorithm is shown in Fig. 2(c) and is in good agreement with the true one. The corresponding maximum *a posteriori* estimate (MAP) of the unknown image is depicted in Fig. 4(d). The obtained coefficients of the PSF-eigenfunctions are close to true coefficients (Fig. 5). For comparison, Fig. 4(b) shows the estimate from the method of Dobigeon *et al.*¹² with a mismatched PSF and Fig. 4(c) shows the estimate generated by the AM algorithm. Our proposed algorithm seems outperform the others visually while preserving fast convergence.

Quantitative comparisons were obtained, by generating different noises for 100 independent trials with each true image. Here, six true images with six corresponding different sparsity levels ($\|\mathbf{x}\|_0 = 6, 11, 18, 30, 59, 97$) were tested. Fig. 6 shows reconstruction error performance for several measures of error used in Ting *et al.*¹¹ and Dobigeon *et al.*¹² to compare different reconstruction algorithms for sparse MRFM images. Notably, compared to the AM algorithm that aims to compensate ‘blindness’ of the unknown PSF and the previous Bayesian method, our method reveals a significant performance gain under most of the displayed performance criteria and sparsity conditions.

5.2 Application to MRFM data

In this section, we apply the myopic Bayesian reconstruction algorithm to the MRFM tobacco virus data⁶ shown in Fig. (7). Here the implementation of AM algorithm is impractical, having “overwhelmingly” slow convergence rates as noted in Herrity *et al.*,¹⁸ due to the large volume of the data, so we only consider Bayesian methods.

The number of principal components (PC) in the PSF perturbation was selected as 4 based on a scree plot. In the data, along z axis, the grid in PSF signal space is 3 times finer than the observation sampling density. We implemented a version of the Bayes MC reconstruction that compensates for unequal projection sampling in x, y, z directions using the interpolation procedure of Dobigeon *et al.*¹²

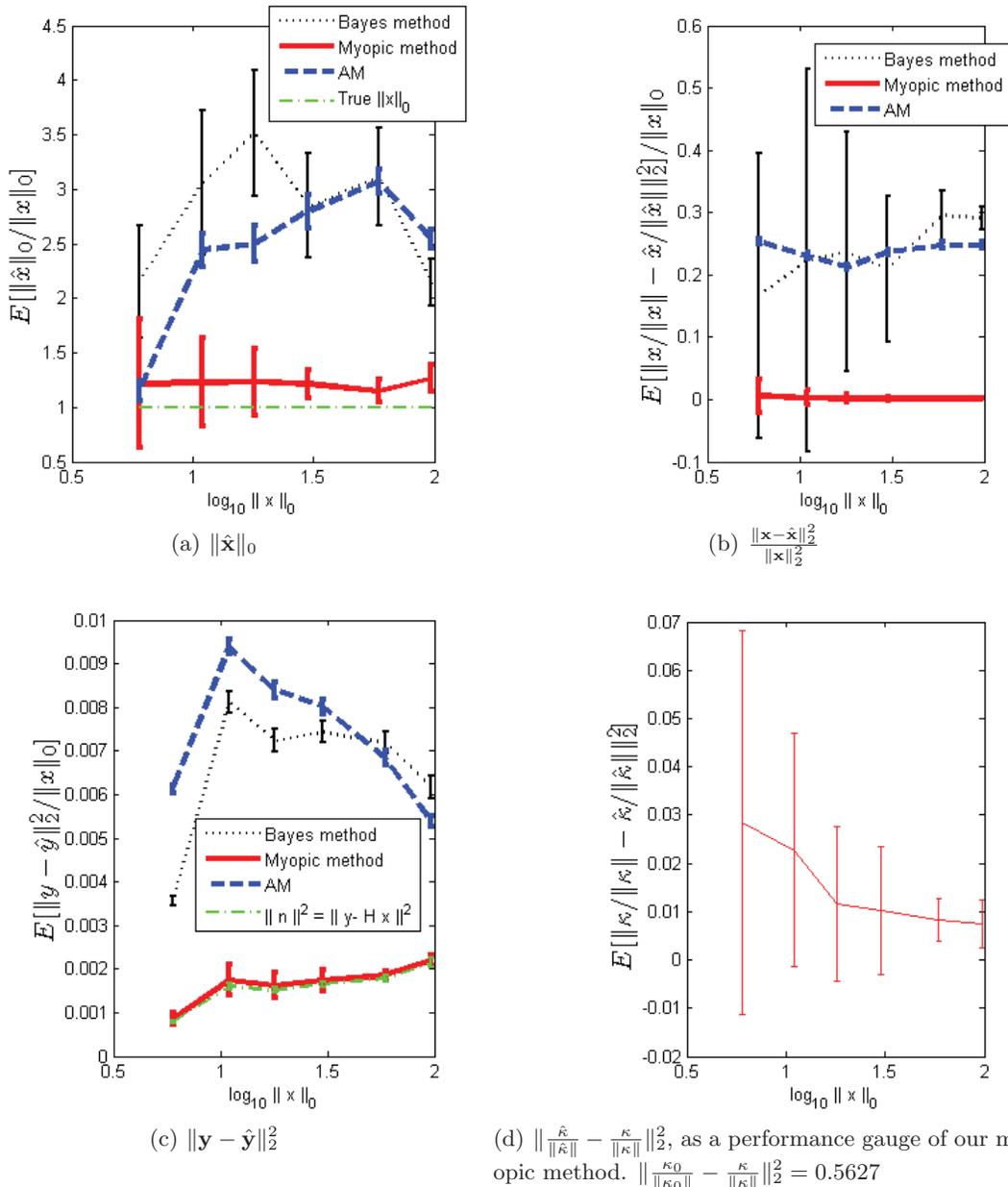


Figure 6. Error bar graphs of results from myopic deconvolution algorithm. For several image x 's of different sparsity level, errors are illustrated with standard deviations.

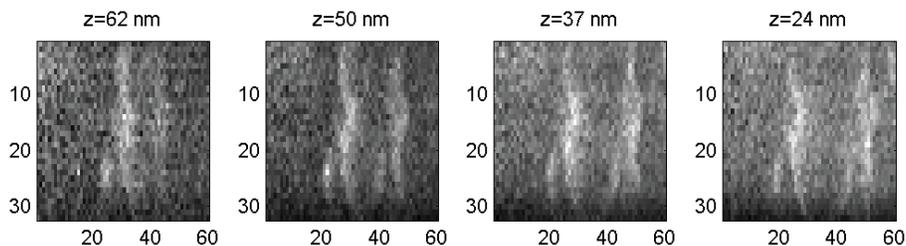
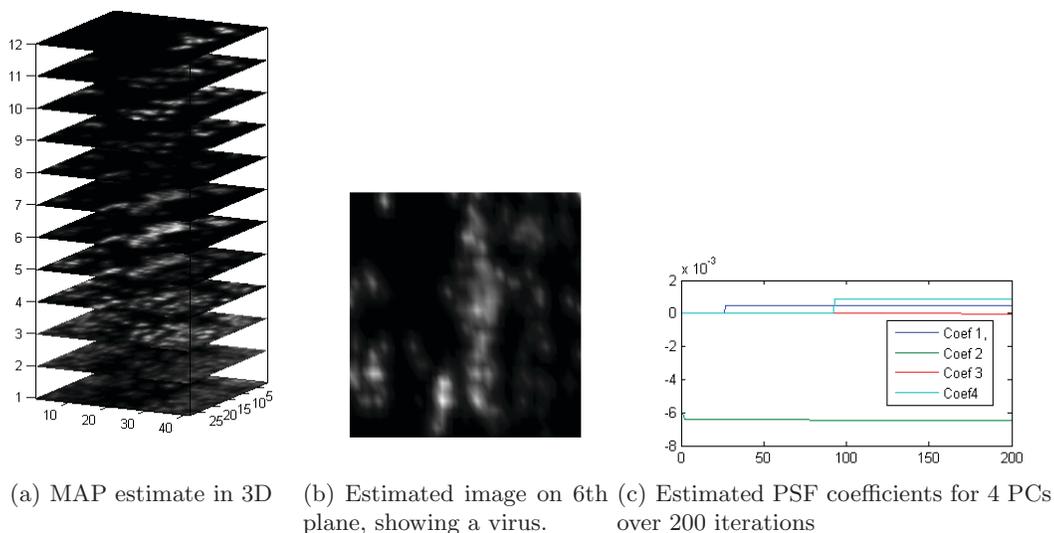
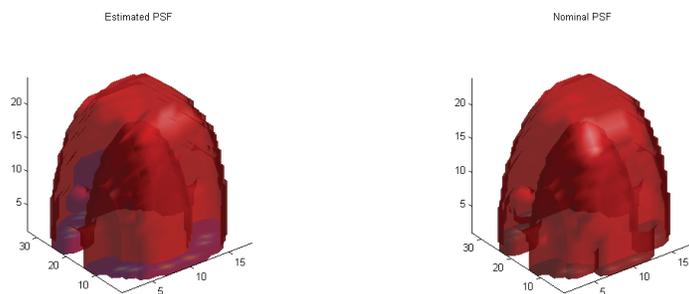


Figure 7. Virus data. z in nm here indicates tip-sample distance.



(a) MAP estimate in 3D (b) Estimated image on 6th plane, showing a virus. (c) Estimated PSF coefficients for 4 PCs over 200 iterations



(d) Estimated PSF. (Hard thresholding with level = $\max(PSF) \times 10^{-4}$ is applied for visualization, as in the Fig. 8(e).)

(e) nominal PSF

Figure 8. Myopic MC Bayes method results and PSF coefficient curves. $\Delta z = 4.3nm$, pixel spacing is $8.3nm \times 16.6nm$ in $x \times y$, respectively. The size of (x, y) plane is $498nm \times 531.2nm$. Smoothing is applied for visualization.

The reconstruction results are shown in Fig. 8(c), 8(d), and 8(e). The small magnitude of PC coefficients indicates that the estimated PSF is close to the assumed PSF. We empirically validated these results by multiple runs of the Gibbs sampler, establishing low standard errors on the estimated PSF coefficients. This suggests that the nominal PSF used in Degen *et al.*⁶ is sufficiently accurate since the myopic reconstruction produces a PSF estimate which is not significantly different from the nominal.

6. CONCLUSION

We have proposed an extension of the method of Bayes reconstruction in Dobigeon *et al.*¹² that simultaneously estimates partially known PSF and the unknown but sparse image. The method uses a prior model on the PSF that reflects a nominal PSF and uncertainty about the nominal PSF. In our algorithm the values of the parameters of the convolution kernel were estimated by a Metropolis-within-Gibbs algorithm. Our approach can be used to empirically evaluate the accuracy of assumed nominal PSF models in the presence of model uncertainty. In our simulation, we showed the myopic Bayesian algorithm has improved performance as compared to the AM reconstruction algorithm and non-myopic Bayes method¹² under several criteria.

Acknowledgement

The authors gratefully acknowledge Dr. Dan Rugar for providing the tobacco virus data and his insightful comments on this work.

APPENDIX A. FAST RECURSIVE SAMPLING STRATEGY

In iterative MRFM algorithms such as AM and the proposed Bayesian method, repeated evaluations of the transformation $T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x})$ can be computationally difficult. For example, at each iteration of the proposed Bayesian myopic deconvolution algorithm, one must generate x_i from its conditional distribution $f(x_i|w, a, \lambda, \sigma^2, \mathbf{x}_{-i}, \mathbf{y})$, which requires the calculation of $T(\kappa, \tilde{\mathbf{x}}_i)$ where $\tilde{\mathbf{x}}_i$ is the vector \mathbf{x} whose i th element has been replaced by 0. Moreover, sampling according to the conditional posterior distributions of σ^2 and λ_k (26) and (23) requires computations of $T(\kappa, \mathbf{x})$.

Exploiting the bilinearity of the transformation $T(\cdot, \cdot)$ can be used to reduce the complexity of the algorithm. We describe below a strategy, similar to those presented in [Appendix B] of Ref. 12 that only requires to compute $T(\cdot, \cdot)$ at most $M \times (K + 1)$ times. First, let \mathbf{I}_M denote the $M \times M$ identity matrix and \mathbf{u}_i its i th column. In a first step of the analysis, the M vectors $\mathbf{h}_i^{(0)}$ ($i = 1, \dots, M$)

$$\mathbf{h}_i^{(0)} = T(\boldsymbol{\kappa}_0, \mathbf{u}_i), \quad (27)$$

and KM vectors $\mathbf{h}_i^{(k)}$ ($i = 1, \dots, M, k = 1, \dots, K$)

$$\mathbf{h}_i^{(k)} = T(\mathbf{v}_k, \mathbf{u}_i), \quad (28)$$

are computed. Then one can compute the quantity $T(\kappa, \tilde{\mathbf{x}}_i)$ and $T(\kappa, \mathbf{x})$ at any stage of the Gibbs sampler without evaluating $T(\cdot, \cdot)$, based on the following decomposition

$$T(\kappa, \mathbf{x}) = \sum_{i=1}^M x_i \mathbf{h}_i^{(0)} + \sum_{k=1}^K \lambda_k \sum_{i=1}^M x_i \mathbf{h}_i^{(k)}. \quad (29)$$

The resulting procedure to update the i th coordinate of the vector \mathbf{x} is described in Algorithm 3 below.

Note that in step 7. of the algorithm above, $T(\boldsymbol{\kappa}, \mathbf{x})$ is recursively computed. Once all the coordinates have been updated, the current $T(\boldsymbol{\kappa}, \mathbf{x})$ can be directly used to sample according to the posterior distribution of the noise variance in (26). Moreover, this quantity can be used to sample according to the conditional posterior distribution of λ_k in (23). More precisely, evaluating $T(\boldsymbol{\kappa}(\lambda_k^*), \mathbf{x})$ in the acceptance probability (25) can be recursively evaluated as follows

$$T(\boldsymbol{\kappa}(\lambda_k^*), \mathbf{x}) = T\left(\boldsymbol{\kappa}\left(\lambda_k^{(t)}\right), \mathbf{x}\right) - \left(\lambda_k^{(t)} - \lambda_k^*\right) \sum_{i=1}^M x_i \mathbf{h}_i^{(k)}. \quad (30)$$

Algorithm 3 Efficient simulation according to $f(\mathbf{x} | w, a, \sigma^2, \mathbf{y})$

At iteration t of the Gibbs sampler, for $i = 1, \dots, M$, update the i th coordinate of the vector

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

via the following steps:

- 1: compute $\mathbf{h}_i = \mathbf{h}_i^{(0)} + \sum_{k=1}^K \lambda_k \mathbf{h}_i^{(k)}$,
 - 2: set $T(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}) = T(\boldsymbol{\kappa}, \mathbf{x}^{(t,i-1)}) - x_i^{(t-1)} \mathbf{h}_i$,
 - 3: set $\mathbf{e}_i = \mathbf{x} - T(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)})$,
 - 4: compute μ_i, η_i^2 and w_i as defined in [6],
 - 5: draw $x_i^{(t)}$ according to (22),
 - 6: set $\mathbf{x}^{(t,i)} = \left[x_1^{(t)}, \dots, x_{i-1}^{(t)}, x_i^{(t)}, x_{i+1}^{(t-1)}, \dots, x_M^{(t-1)} \right]^T$,
 - 7: set $T(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}) = T(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}) + x_i^{(t)} \mathbf{h}_i$.
-

REFERENCES

- [1] Sidles, J. A., “Noninductive detection of single-proton magnetic resonance,” *Appl. Phys. Lett.* **58**, 2854–2856 (June 1991).
- [2] Sidles, J. A., “Folded stern-gerlach experiment as a means for detecting nuclear magnetic resonance in individual nuclei,” *Phys. Rev. Lett.* **68**, 1124–1127 (Feb 1992).
- [3] Sidles, J. A., Garbini, J. L., Bruland, K. J., Rugar, D., Züger, O., Hoen, S., and Yannoni, C. S., “Magnetic resonance force microscopy,” *Rev. Mod. Phys.* **67**, 249–265 (Jan 1995).
- [4] Rugar, D., Yannoni, C. S., and Sidles, J. A., “Mechanical detection of magnetic resonance,” *Nature* **360**, 563–566 (Dec. 1992).
- [5] Züger, O., Hoen, S. T., Yannoni, C. S., and Rugar, D., “Three-dimensional imaging with a nuclear magnetic resonance force microscope,” *J. Appl. Phys.* **79**, 1881–1884 (Feb. 1996).
- [6] Degen, C. L., Poggio, M., Mamin, H. J., Rettner, C. T., and Rugar, D., “Nanoscale magnetic resonance imaging,” *Proc. Nat. Academy of Science* **106**, 1313–1317 (Feb. 2009).
- [7] Chao, S., Dougherty, W. M., Garbini, J. L., and Sidles, J. A., “Nanometer-scale magnetic resonance imaging,” *Review Sci. Instrum.* **75**, 1175–1181 (April 2004).
- [8] Züger, O. and Rugar, D., “First images from a magnetic resonance force microscope,” *Applied Physics Letters* **63**(18), 2496–2498 (1993).
- [9] Züger, O. and Rugar, D., “Magnetic resonance detection and imaging using force microscope techniques,” *J. Appl. Phys.* **75**, 6211–6216 (May 1994).
- [10] Degen, C. L., Poggio, M., Mamin, H. J., Rettner, C. T., and Rugar, D., “Nanoscale magnetic resonance imaging. Supporting information,” *Proc. Nat. Academy of Science* **106** (Feb. 2009).
- [11] Ting, M., Raich, R., and Hero, A. O., “Sparse image reconstruction for molecular imaging,” *IEEE Trans. Image Processing* **18**, 1215–1227 (June 2009).
- [12] Dobigeon, N., Hero, A. O., and Tourneret, J.-Y., “Hierarchical Bayesian sparse image reconstruction with application to MRFM,” *IEEE Trans. Image Processing* **18**, 2059–2070 (Sept. 2009).
- [13] Mamin, J., Budakian, R., and Rugar, D., “Point response function of an MRFM tip,” tech. rep., IBM Research Division (Oct. 2003).
- [14] Makni, S., Ciuciu, P., Idier, J., and Poline, J.-B., “Semi-blind deconvolution of neural impulse response in fMRI using a Gibbs sampling method,” in [*Proc. IEEE Int. Conf. Acoust., Speech, and Signal (ICASSP)*], **5**, 601–604 (May 2004).

- [15] Pillonetto, G. and Cobelli, C., “Identifiability of the stochastic semi-blind deconvolution problem for a class of time-invariant linear systems,” *Automatica* **43**, 647–654 (April 2007).
- [16] Sarri, P., Thomas, G., Sekko, E., and Neveux, P., “Myopic deconvolution combining Kalman filter and tracking control,” in [*Proc. IEEE Int. Conf. Acoust., Speech, and Signal (ICASSP)*], **3**, 1833–1836 (1998).
- [17] Chenegros, G., Mugnier, L. M., Lacombe, F., and Glanc, M., “3D phase diversity: a myopic deconvolution method for short-exposure images: application to retinal imaging,” *J. Opt. Soc. Am. A* **24**, 1349–1357 (May 2007).
- [18] Herrity, K., Raich, R., and Hero, A. O., “Blind reconstruction of sparse images with unknown point spread function,” in [*Proc. Computational Imaging Conference in IS&T/SPIE Symposium on Electronic Imaging Science and Technology*], Bouman, C. A., Miller, E. L., and Pollak, I., eds., **6814**, SPIE, San Jose, CA, USA (Jan. 2008).
- [19] Orieux, F., Giovannelli, J.-F., and Rodet, T., “Bayesian estimation of regularization and point spread function parameters for wiener–hunt deconvolution,” *J. Opt. Soc. Am. A* **27**(7), 1593–1607 (2010).
- [20] Ting, M., Raich, R., and Hero, A. O., “Sparse image reconstruction using sparse priors,” in [*Proc. IEEE Int. Conf. Image Processing (ICIP)*], 1261–1264 (Oct. 2006).
- [21] Punsakaya, E., Andrieu, C., Doucet, A., and Fitzgerald, W., “Bayesian curve fitting using MCMC with applications to signal segmentation,” *IEEE Trans. Signal Processing* **50**, 747–758 (March 2002).
- [22] Dobigeon, N., Tourneret, J.-Y., and Davy, M., “Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach,” *IEEE Trans. Signal Processing* **55**, 1251–1263 (April 2007).
- [23] Dobigeon, N., Tourneret, J.-Y., and Chang, C.-I., “Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery,” *IEEE Trans. Signal Processing* **56**, 2684–2695 (July 2008).
- [24] Robert, C. P. and Casella, G., [*Monte Carlo Statistical Methods*], Springer-Verlag, New York, NY, USA (1999).
- [25] Bremaud, P., [*Markov Chains: Gibbs fields, Monte Carlo Simulation, and Queues.*], Springer Verlag (1999).