

Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 6278

To link to this article : DOI:10.1109/TIP.2012.2199505
URL : http://dx.doi.org/10.1109/TIP.2012.2199505

To cite this version : Park, Se Un and Dobigeon, Nicolas and Hero, Alfred O. *Semi-blind sparse image reconstruction with application to MRFM.* (2012) IEEE Transactions on Image Processing, vol. 21 (n° 9). pp. 3838-3849. ISSN 1057-7149

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes.diff.inp-toulouse.fr

Semi-Blind Sparse Image Reconstruction With Application to MRFM

Se Un Park, Nicolas Dobigeon, *Member, IEEE*, and Alfred O. Hero, *Fellow, IEEE*

***Abstract*— We propose a solution to the image deconvolution problem where the convolution kernel 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 PSF uncertainty in a high-dimensional space. Unlike recent developments on blind deconvolution of natural images, we assume the image is sparse in the pixel basis, a natural sparsity arising in magnetic resonance force microscopy (MRFM). Our approach adopts a Bayesian Metropolis-within-Gibbs sampling framework. The performance of our Bayesian semi-blind algorithm for sparse images is superior to previously proposed semi-blind algorithms such as the alternating minimization algorithm and blind algorithms developed for natural images. We illustrate our myopic algorithm on real MRFM tobacco virus data.**

***Index Terms*— Bayesian inference, magnetic resonance force microscopy (MRFM) experiment, Markov chain Monte Carlo (MCMC) methods, semi-blind (myopic) sparse deconvolution.**

I. Introduction

RECENTLY, a new 3-D imaging technology called magnetic resonance force microscopy (MRFM) has been developed. The principles of MRFM were introduced by Sidles [1]–[3], who described its potential for achieving 3-D atomic scale resolution. In 1992 and 1996, Rugar *et al.* [4], [5] reported experiments that demonstrated the feasibility of MRFM and produced the first MRFM images. More recently, MRFM volumetric spatial resolutions of less than 10 nm have been demonstrated for imaging a biological sample [6]. The signal provided by MRFM is a so-called force map that is the 3-D 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 filters [5], [8], [9] whereas others are based on iterative least-squares reconstruction approaches [6], [7], [10]. More recently, this problem has been addressed 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 physics of the device including mass of cantilever probe, ferromagnetic constant of probe tip, and external field strength. Unfortunately, in practice the physical parameters that tune the response of the MRFM tip are only partially known. In such circumstances, the PSF used in the reconstruction algorithm is mismatched to the true PSF of the microscope and the quality of standard MRFM image reconstruction will suffer if one does not account for this mismatch. Estimating the unknown image and the partially known 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. The myopic image restoration problem was previously studied within a hierarchical Bayesian framework [18] with partially known blur functions in many applications, including natural and astronomical imaging [19], [20]. This previous work [19], [20] models the deviation of the PSF as uncorrelated zero mean Gaussian noise. The authors of [21] considered an extension of this model to a non-sparse, simultaneous autoregression prior model for both the image and PSF. Compared to [21], recent papers on single motion deblurring in photography [22], [23] use heavier tailed distributions for the gradient of images and an exponential distribution for the PSF. In addition, the algorithm in [22] separately identifies the PSF using a multiscale approach to perform conventional image restoration. The authors of [23] proposed an image prior to reduce ringing artifacts from blind deconvolution of photo images. This paper considers an alternative model, appropriate to MRFM but significantly different from photography, that imposes sparsity on the image and an empirical Bayes prior on the PSF.

To mitigate the effects of PSF mismatch on MRFM sparse image reconstruction, a non-Bayesian alternating minimization (AM) algorithm [24] was proposed by Herrity *et al.* which showed robust performance. In this paper, we propose a hierarchical Bayesian approach to semi-blind image deconvolution that exploits prior information on the PSF model. This is a semi-blind modification of the Bayesian MRFM reconstruction approach of Dobigeon *et al.* [12] that uses an adaptive tuning scheme to produce a Bayesian estimate of the PSF and a Bayesian reconstruction of the image. The contribution of this paper is a novel Bayesian approach to a joint estimation of

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

N. Dobigeon is with the École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique et des Télécommunications, University of Toulouse, Toulouse 31071, France (e-mail: nicolas.dobigeon@enseeiht.fr).

PSFs and images. We represent the PSF on a truncated orthogonal basis, where the basis elements are the singular vectors in the singular value decomposition of the family of perturbed nominal PSFs. A Gaussian prior model specifies a log quadratic Bayes prior on deviations from the nominal PSF. Our approach is related to the recent papers of Tzikas *et al.* [25] and Orieux *et al.* [26]. In [25], a pixel-wise, space-invariant Gaussian kernel basis is assumed with a gradient-based image prior. Orieux *et al.* introduced a Metropolis-within-Gibbs algorithm to estimate the parameters that tune the device response. The strategy [26] 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. Since we only need to estimate linear perturbations about the nominal PSF relative to a low-dimensional precomputed and truncated basis set, this leads to reduction in computational complexity and an improvement in convergence as compared to [25] and [26]. We approximate the full posterior distribution of the PSF and the image using samples generated by a Markov chain Monte Carlo (MCMC) algorithm. Simulations and comparisons to other state-of-the-art blind deconvolution algorithms are presented and quantify the advantages of our algorithm for myopic sparse image reconstruction. We then apply it to real MRFM tobacco virus data made available by our IBM collaborators.

This paper is organized as follows. Section II formulates the problem. Section III covers the Bayesian framework of image modeling and Section IV proposes a solution in this framework. Section V shows simulation results and an application to the real MRFM data.

II. Forward Imaging and PSF Model

Let $\mathbf{X}$ denote the $l_1 \times \cdots \times l_n$ unknown nD positive spin density image to be recovered (e.g., $n = 2$ or $n = 3$) and $\mathbf{x} \in \mathbb{R}^M$ denote the vectorized version of $\mathbf{X}$ with $M = l_1 l_2, \ldots, l_n$. This image is to be reconstructed from a collection of $P(= M)$ measurements $\mathbf{y} = [y_1, \ldots, y_P]^T$ via the following noisy transformation:

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

where $T(\cdot, \cdot)$ is the n-dimensional convolution operator or the mean response function $\mathbb{E}[\mathbf{y}|\boldsymbol{\kappa}, \mathbf{x}]$, $\mathbf{n}$ is a $P \times 1$ observation noise vector, and $\boldsymbol{\kappa}$ is the kernel modeling the response of the imaging device.

A typical PSF for MRFM is shown by Mamin *et al.* [13] for horizontal and vertical MRFM tip configurations. In (1) $\mathbf{n}$ is an additive Gaussian noise sequence, independent of $\mathbf{x}$, 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 $\boldsymbol{\kappa}_0$

$$\boldsymbol{\kappa} = \boldsymbol{\kappa}_0 + \Delta\boldsymbol{\kappa}. \tag{2}$$

In the MRFM application, the PSF is described by an approximate parametric function that depends on the experimental setup. Based on the physical parameters (gathered in the vector $\boldsymbol{\zeta}$) 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 $\boldsymbol{\kappa}_0$.

If a vector of the nominal values of parameters $\boldsymbol{\zeta}_0$ for the parametric PSF model $\boldsymbol{\kappa}_{\text{gen}}(\boldsymbol{\zeta})$ is known, then direct estimation of a parameter deviation $\Delta\boldsymbol{\zeta}$ can be performed by evaluation of $\boldsymbol{\kappa}_{\text{gen}}(\boldsymbol{\zeta}_0 + \Delta\boldsymbol{\zeta})$. However, in MRFM, as shown by Mamin *et al.* [13], $\boldsymbol{\kappa}_{\text{gen}}(\boldsymbol{\zeta})$ is a nonlinear function with many parameters that are required to satisfy "resonance conditions" to produce a meaningful MRFM PSF. This makes direct estimation of the PSF difficult.

In this paper, we take a similar approach to the "basis expansions" in [27, Ch. 5], [25] for approximation of the PSF deviation $\Delta\boldsymbol{\kappa}$ as linear models. We propose to model the deviation $\Delta\boldsymbol{\kappa}$ as a linear combination of elements of an *a priori* known basis $\mathbf{v}_k$, $k = 1, \ldots, K$

$$\Delta\boldsymbol{\kappa} = \sum_{k=1}^{K} \lambda_k \mathbf{v}_k \tag{3}$$

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

$$\mathbf{y} = T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x}) + \mathbf{n} = \mathbf{H}(\boldsymbol{\lambda})\mathbf{x} + \mathbf{n} \tag{4}$$

where $\mathbf{H}(\boldsymbol{\lambda})$ is an $P \times M$ matrix that describes convolution by the PSF kernel $\boldsymbol{\kappa}(\boldsymbol{\lambda})$.

We next 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 $T(\cdot, \cdot)$.

III. Hierarchical Bayesian Model

A. Likelihood Function

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

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

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

B. Parameter Prior Distributions

In this section, we introduce prior distributions for the parameters $\boldsymbol{\theta}$.

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

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

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 $\mathbb{E}$

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

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

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

This image prior is similar to the LAZE distribution [weighted average of a Laplacian probability density function (pdf) and an atom at zero] used, for example, by Ting *et al.* [11], [28]. As motivated by Ting *et al.* and Dobigeon *et al.* [11], [12], the image prior in (6) has the interesting property of enforcing some pixel values 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.

2) PSF Parameter Prior: We assume that the parameters $\lambda_1, \ldots, \lambda_K$ are *a priori* independent and uniformly distributed over known intervals associated with error tolerances centered at 0. Specifically, define the interval

$$\mathcal{S}_k = [-\Delta\lambda_k, \Delta\lambda_k] \tag{9}$$

and define the distribution of $\boldsymbol{\lambda}$ as

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

with $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_K]^T$. In our experiment, the $\Delta\lambda_k$ are set to be large enough to be non-informative, i.e., an improper, flat prior.

3) Noise Variance Prior: A noninformative Jeffreys' prior is selected as prior distribution for the noise variance

$$f\left(\sigma^2\right) \propto \frac{1}{\sigma^2}. \tag{11}$$

This model is equivalent to an inverse gamma prior with a non-informative Jeffreys' hyperprior, which can be seen by integrating out the variable corresponding to the hyperprior [12].

C. Hyperparameter Priors

Define the hyperparameter vector associated with the image and noise variance prior distributions as $\boldsymbol{\Phi} = \{a, w\}$. In our hierarchical Bayesian framework, the estimation of these hyperparameters requires prior distributions in the hyperparameters. These priors are defined in [12] but for completeness their definitions are reproduced below.

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

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

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

2) Hyperparameter w: A uniform distribution on the simplex $[0,1]$ is selected as prior distribution for the mean proportion of nonzero pixels

$$w \sim \mathcal{U}([0,1]). \tag{13}$$

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

$$\begin{aligned} f(\boldsymbol{\Phi}|\boldsymbol{\alpha}) &= f(w) f(a) \\ &\propto \frac{1}{a^{\alpha_0+1}} \exp\left(-\frac{\alpha_1}{a}\right) \mathbf{1}_{[0,1]}(w) \mathbf{1}_{\mathbb{R}^+}(a). \end{aligned} \tag{14}$$

D. Posterior Distribution

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

$$f(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta}) f(\boldsymbol{\theta}|\boldsymbol{\Phi}) f(\boldsymbol{\Phi}) \tag{15}$$

with

$$f(\boldsymbol{\theta}|\boldsymbol{\Phi}) = f(\mathbf{x}|a,w) f(\boldsymbol{\lambda}) f\left(\sigma^2\right) \tag{16}$$

where $f(\mathbf{y}|\boldsymbol{\theta})$ and $f(\boldsymbol{\Phi})$ have been defined in (5) and (14). The conjugacy of priors in this hierarchical structure allows one to integrate out the parameters σ^2, and the hyperparameter $\boldsymbol{\Phi}$ in the full posterior distribution (15), yielding

$$\begin{aligned} f(\mathbf{x}, \boldsymbol{\lambda}|\mathbf{y}, \alpha_0, \alpha_1) &\propto \int f(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{y}) \, \mathrm{d}w \mathrm{d}a \mathrm{d}\sigma^2 \\ &\propto \frac{\mathcal{B}e(1+n_1, 1+n_0)}{\|\mathbf{y} - T(\boldsymbol{\kappa}(\boldsymbol{\lambda}), \mathbf{x})\|^N} \frac{\Gamma(n_1+\alpha_0)}{(\|\mathbf{x}\|_1 + \alpha_1)^{n_1+\alpha_0}} \prod_{k=1}^{K} \frac{1}{2\Delta\lambda_k} \mathbf{1}_{\mathcal{S}_k}(\lambda_k) \end{aligned} \tag{17}$$

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

The next section presents the Metropolis-within-Gibbs algorithm [29] 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 $\boldsymbol{\lambda}$.

IV. Metropolis-Within-Gibbs Algorithm for Semi-blind Sparse Image Reconstruction

We describe in this section a Metropolis-within-Gibbs sampling strategy that allows one to generate samples $\{\mathbf{x}^{(t)}, \boldsymbol{\lambda}^{(t)}\}_{t=1,\ldots}$ distributed according to the posterior distribution in (17). As sampling directly from (17) 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 $\mathbf{x}, \boldsymbol{\lambda}$, and σ^2 conditioned on all other variables [12], [30]. The contribution of this paper to [12] is to present an algorithm that simultaneously estimates both the image and PSF. The algorithm results in consistently stable output images and PSFs.

The main steps of our proposed sampling algorithm are given in Sections IV-A–IV-C (see also Algorithm 1).

A. 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 three-step procedure.

Algorithm 1 Metropolis-Within-Gibbs Sampling Algorithm for Semi-blind Sparse Image Reconstruction

1: *% Initialization:*
2: Sample the unknown image $\mathbf{x}^{(0)}$ from pdf in (8),
3: Sample the noise variance $\widetilde{\sigma}^{2(0)}$ from the pdf in (11),
4: *% Iterations:*
5: **for** $t = 1, 2, \ldots,$ **do**
6: Sample hyperparameter $w^{(t)}$ from the pdf in (19),
7: Sample hyperparameter $a^{(t)}$ from the pdf in (20),
8: For $i = 1, \ldots, M$, sample the pixel intensity $x_i^{(t)}$ from the pdf in (21),
9: For $k = 1, \ldots, K$, sample the PSF parameter $\lambda_k^{(t)}$ from the pdf in (23), by using Metropolis-Hastings step (see Algo. 2),
10: Sample the noise variance $\widetilde{\sigma}^{2(t)}$ from the pdf in (26),
11: **end for**

1) Generation of Samples According to $f(w|\mathbf{x}, \alpha_0, \alpha_1)$: The conditional posterior distribution of w is

$$f(w|\mathbf{x}) \propto (1-w)^{n_0+1-1} w^{n_1+1-1} \tag{18}$$

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:

$$w|\mathbf{x} \sim \mathcal{B}e(1+n_1, 1+n_0). \tag{19}$$

2) Generation of Samples According to $f(a|\mathbf{x})$: The joint posterior distribution (15) yields

$$a|\mathbf{x}, \alpha_0, \alpha_1 \sim \mathcal{IG}(\|\mathbf{x}\|_0 + \alpha_0, \|\mathbf{x}\|_1 + \alpha_1). \tag{20}$$

3) Generation of Samples According to $f(\mathbf{x}|w, a, \lambda, \sigma^2, \mathbf{y}.)$: The posterior distribution of each component x_i $(i = 1, \ldots, M)$ given all other variables is derived as

$$f\left(x_i|w, a, \lambda, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}\right) \propto (1-w_i)\delta(x_i) + w_i \phi_+\left(x_i|\mu_i, \eta_i^2\right) \tag{21}$$

where $\mathbf{x}_{-i}$ stands for the vector $\mathbf{x}$ whose ith component has been removed and μ_i and η_i^2 are defined as follows:

$$\eta_i^2 = \frac{\sigma^2}{\|\mathbf{h}_i\|^2}, \quad \mu_i = \eta_i^2\left(\frac{\mathbf{h}_i^T \mathbf{e}_i}{\sigma^2} - \frac{1}{a}\right) \tag{22}$$

with $\mathbf{h}_i$ and $\mathbf{e}_i$ defined in Appendix A.

In (21), $\phi_+(\cdot, m, s^2)$ stands for the pdf of the truncated Gaussian distribution defined on $\mathbb{R}_+^*$ with hidden mean m and hidden variance s^2. Therefore, from (21), $x_i|w, a, \lambda, \sigma^2, \mathbf{x}_{-i}, \mathbf{y}$ is a Bernoulli truncated-Gaussian variable with parameter (w_i, μ_i, η_i^2).

To summarize, generation of samples distributed according to $f(\mathbf{x}|w, \sigma^2, a, , \mathbf{y}.)$ can be performed by updating the coordinates of $\mathbf{x}$ using M Gibbs moves (requiring generation of Bernoulli truncated-Gaussian variables). A pixel-wise fast and recursive sampling strategy is presented in Appendix A and an accelerated sparsity enforcing simulation scheme is described in Appendix B.

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

1: Sample $\varepsilon \sim \mathcal{N}\left(0, s_p^2\right)$,
2: Propose $\lambda_k^\star$ according to $\lambda_k^\star = \lambda_k^{(t)} + \varepsilon$,
3: Draw $u \sim \mathcal{U}([0,1])$,
4: Set $\lambda_k^{(t+1)} = \begin{cases} \lambda_k^\star, & \text{if } u \le \rho_{\lambda_k^{(t)} \to \lambda_k^\star}, \\ \lambda_k^{(t)}, & \text{otherwise.} \end{cases}$

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

B. 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

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

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 Metropolis–Hastings step with a random walk proposition [29] from a centered Gaussian distribution. In order to sample efficiently, the detailed procedure of how to choose an appropriate value of the variance s_k^2 of the proposal distribution for λ_k is described in Appendix C. At iteration t of the algorithm, the acceptance probability of a proposed state $\lambda_k^\star$ is

$$\rho_{\lambda_k^{(t)} \to \lambda_k^\star} = \min\left(1, a_k \mathbf{1}_{\mathcal{S}_k}\left(\lambda_k^\star\right)\right) \tag{24}$$

with

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

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

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

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

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

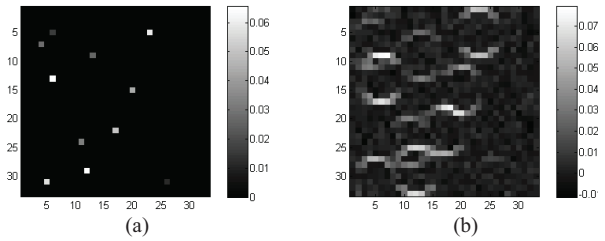
V. Experiments

In this section, we present simulation results that compare the proposed semi-blind Bayesian deconvolution algorithms with the nonblind method [12], the AM algorithm [24], and other blind deconvolution methods. Here, a nominal PSF $\boldsymbol{\kappa}_0$ was selected that corresponds to the mathematical MRFM point response model proposed by Mamin *et al.* [13].

Using our MCMC algorithm described in Section IV, the MMSE estimators of image and PSF are approximated

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_0	6.0 nm
Cantilever tip moment	m	4.6×10^5 emu
Peak cantilever oscillation oscillation	x_{pk}	0.8 nm
Maximum magnetic field gradient	G_{max}	125



(a) (b)

Fig. 1. Simulated true image and MRFM raw image exhibiting superposition of point responses (see Fig. 2) and noise. (a) Sparse true image ($\|\mathbf{x}\|_0 = 11$). (b) Raw MRFM observation.

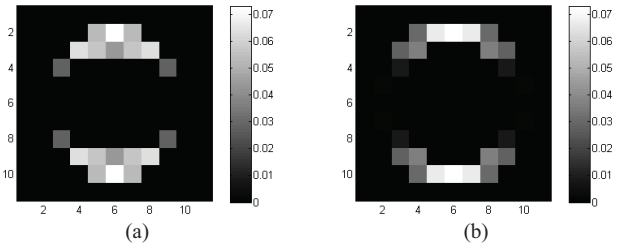
by ensemble averages over the generated samples after the burn-in period. The joint MAP estimator is selected among the drawn samples, after the stationary distribution is achieved, such that it maximizes the posterior distribution $f(\mathbf{x}, \boldsymbol{\lambda}|\mathbf{y})$ [31].

A. Simulation on Simulated Sparse Images

We performed simulations of MRFM measurements for PSF and image models similar to those described by Dobigeon *et al.* [12]. The signal-to-noise ratio was set to SNR = 10 dB. 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 $\boldsymbol{\kappa}_0$ is generated following the physical model described by Mamin *et al.* [13] when the physical parameters are tuned to the values displayed in Table I. This yields a 11×11 2-D convolution kernel represented in Fig. 2(a). We assume that the true PSF $\boldsymbol{\kappa}$ comes from the same physical model where the radius of the tip and the distance from the tip to the sample have been mis-specified with 2% error as $R = R_0 - 2\% = 3.92$ and $d = d_0 + 2\% = 6.12$. This leads to the convolution kernel depicted in Fig. 2(b). The observed measurements $\mathbf{y}$, shown in Fig. 1(b), are a 32×32 image of size $P = 1024$.

The proposed algorithm requires the definition of K basis vectors $\mathbf{v}_k$, $k = 1, \ldots, K$, that span a subspace representing possible perturbations $\Delta\boldsymbol{\kappa}$. We empirically determined this basis using the following PSF variational eigendecomposition approach. A set of 5000 experimental PSFs $\tilde{\boldsymbol{\kappa}}_j$, $j = 1, \ldots, 5000$, were generated following the model described by Mamin *et al.* [13] with parameters d and R randomly drawn according to Gaussian distribution[1] centered at the nominal values d_0, R_0, respectively. Then a standard principal component analysis (PCA) of the residuals $\{\tilde{\boldsymbol{\kappa}}_j - \boldsymbol{\kappa}_0\}_{j=1,\ldots,5000}$, by allowing the maximum variance over the parameters that produce valid MRFM PSFs, was 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, was determined empirically by detecting a knee at the scree plot shown in Fig. 3. The first four eigenfunctions, corresponding to the first four largest eigenvalues, explain 98.69% of the observed perturbations. The four principal patterns of basis vectors are depicted in Fig. 4.

[1]We used a PSF generator provided by Dan Rugar's group at IBM [13]. The variances of the Gaussian distributions are carefully tuned so that their standard deviations produce a minimal volume ellipsoid that contains the set of valid PSFs of the form specified in [13].



(a) (b)

Fig. 2. Assumed PSF $\boldsymbol{\kappa}_0$ and actual PSF $\boldsymbol{\kappa}$. (a) Nominal MRFM PSF. (b) True MRFM PSF.

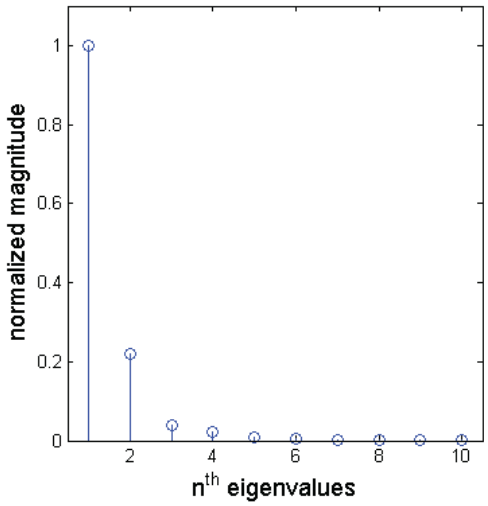


Fig. 3. Scree plot of residual PCA approximation error in l_2 norm (magnitude is normalized up to the largest value, i.e., $\lambda_{\text{max}} := 1$).

The proposed semi-blind Bayesian reconstruction algorithm was applied to estimate both the sparse image and the PSF coefficients of $\mathbf{v}_k$'s, using the prior in (6). From the observation shown in Fig. 1(b) the PSF estimated by the proposed algorithm is shown in Fig. 5(a) and is in good agreement with the true one. The corresponding maximum *a posteriori* estimate (MAP) of the unknown image is depicted in Fig. 6. The obtained coefficients of the PSF eigenfunctions are close to true coefficients (Fig. 7).

B. Comparison to Other Methods

For comparison to a nonblind method, Fig. 6(b) shows the estimate using the Bayesian nonblind method [12] with a mismatched PSF. Fig. 6(c) shows the estimate generated by the AM algorithm. The nominal PSF described in Section V-A is used in the AM algorithm and hereafter

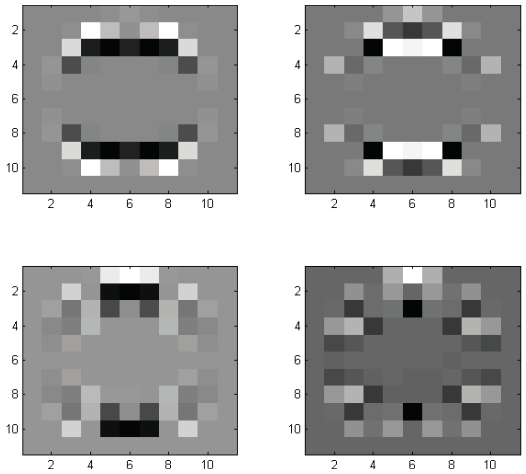


Fig. 4. $K = 4$ principal vectors ($\mathbf{v}_k$) of the perturbed PSF, identified by PCA.

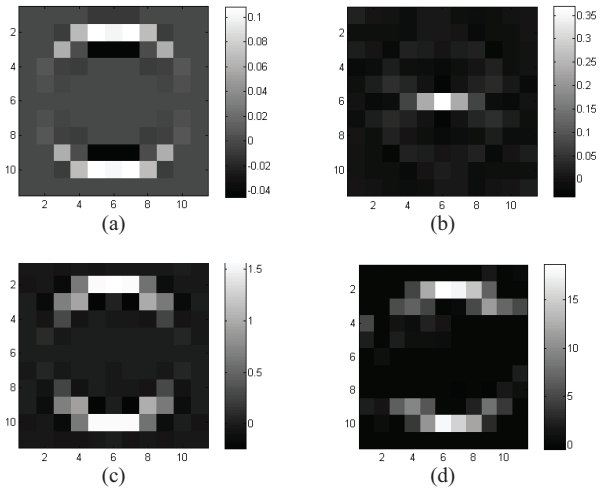


Fig. 5. Estimated PSF $\hat{\boldsymbol{\kappa}}$ of MRFM tip using our semi-blind method, Amizic's method (using TV prior), Almeida's method (using progressive regularization), and Tzikas' method (using the kernel basis PSF model), respectively. For a fair comparison, we used sparse image priors for the methods. (See Section V-B for details on the methods.) (a) Proposed method. (b) Amizic's method. (c) Almeida's method. (d) Tzikas' method.

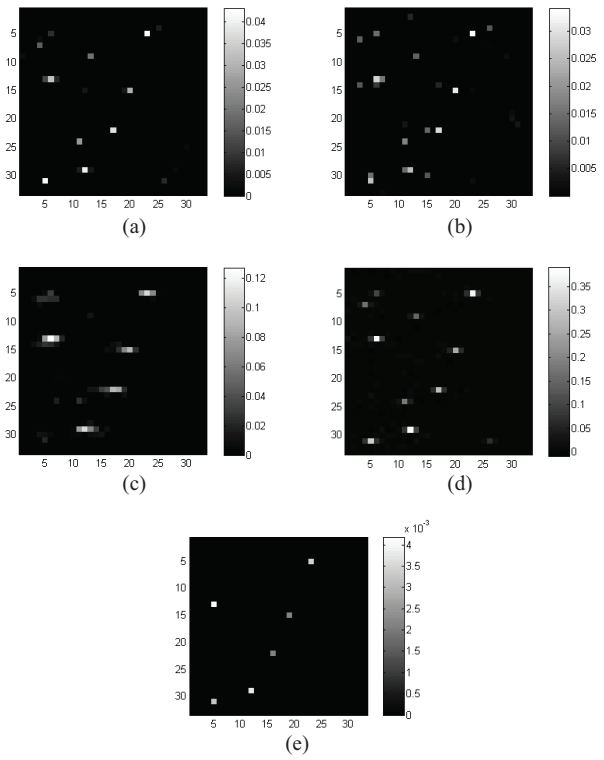


Fig. 6. True sparse image and estimated images from Bayesian nonblind, AM, our semi-blind, Almeida's, and Tzikas' methods. (a) MAP, proposed method. (b) MAP, Bayesian non blind method with $\boldsymbol{\kappa}_0$. (c) AM. (d) Almeida's method. (e) Tzikas' method.

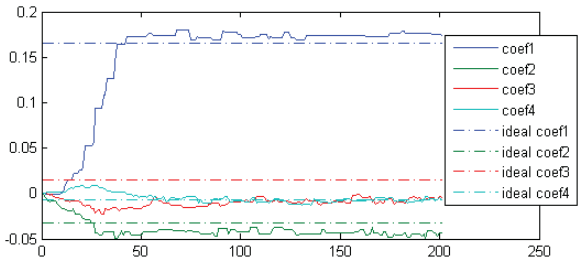


Fig. 7. Estimated PSF coefficients for four PCs over 200 iterations. These curves show fast convergence of our algorithm. "Ideal coefficients" are the projection values of the true PSF onto the space spanned by four principal PSF bases.

in other semi-blind algorithms, and the parameter values of AM algorithm were set empirically according to the procedure in [24]. Our proposed algorithm appears to outperform the others (Fig. 6) while preserving fast convergence (Fig. 7).

Quantitative comparisons were obtained by generating different noises in 100 independent trials for a fixed true image. Here, six different true images with six corresponding different sparsity levels ($\|\mathbf{x}\|_0 = 6, 11, 18, 30, 59, 97$) were tested. Fig. 8 presents the two histograms of the results with the six sets in the corresponding two error criteria, $\|\hat{\mathbf{x}} - \mathbf{x}\|^2$, $\|\hat{\mathbf{x}}\|_0$, respectively, both of which indicate that our method performs better and is more stable than the other two methods.

Fig. 9 shows reconstruction error performance for several measures of error used by 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 nonblind method, our method reveals a significant performance gain under most of the investigated performance criteria and sparsity conditions.

In addition to the AM and nonblind comparisons shown in Fig. 8, we made direct comparisons between our sparse MRFM reconstruction method and several state-of-the-art blind image reconstruction methods [22], [23], [25], [32], [33]. In all cases, the algorithms were initialized with the nominal, mismatched PSF and were applied to a sparse MRFM-type image like in Fig. 1. For a fair comparison, we made a sparse prior modification in the image model of other algorithms. The total variation (TV)-based prior for the PSF suggested by Amizic *et al.* [32] was also implemented. The obtained PSF from this method was considerably worse than the one

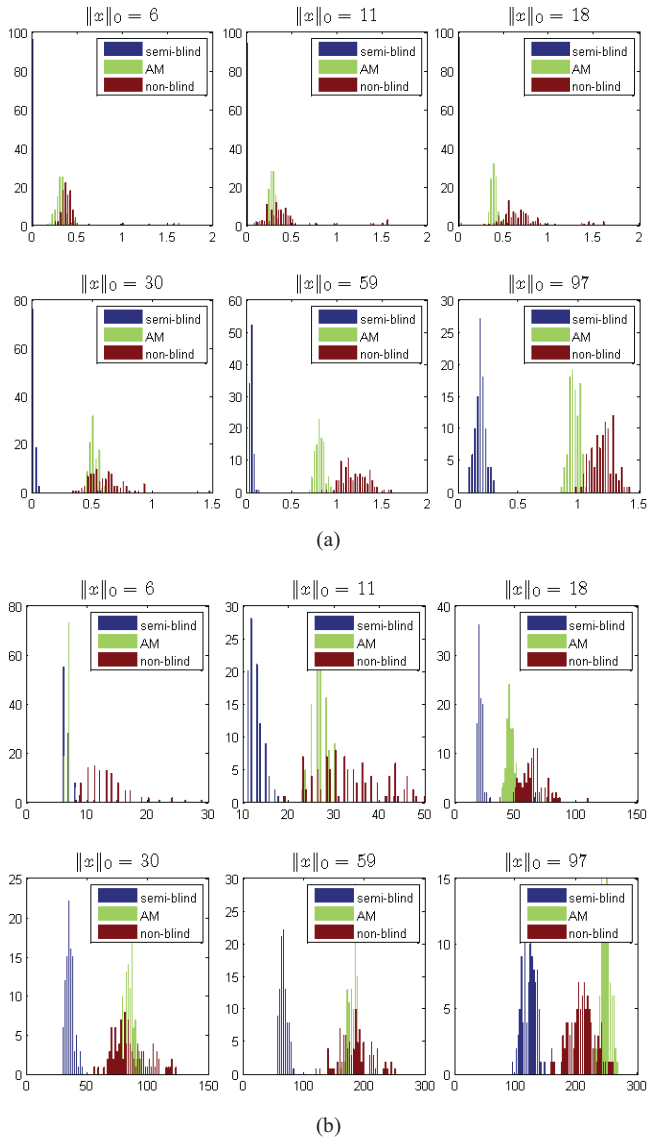


Fig. 8. Histograms of l_2 and l_0 norm of the reconstruction error. Note that the proposed semi-blind reconstructions exhibit smaller mean error and more concentrated error distribution than the nonblind method in [12] and the AM method in [24]. (a) Histograms of the normalized l_2 norm errors. x-axis is $\|(\mathbf{x}/\|\mathbf{x}\|) - (\hat{\mathbf{x}}/\|\hat{\mathbf{x}}\|)\|_2^2/\|\mathbf{x}\|_0$. (b) Histograms of the l_0 measures. x-axis is $\|\hat{\mathbf{x}}\|_0$.

estimated by our proposed method [see Fig. 5(b)] resulting in a very poor quality image deconvolution.[2]

The recent blind deconvolution method proposed by Almeida *et al.* [33] utilizes the "sharp Edge" property in natural images, with initial, high-regularization in order to effectively evaluate the PSF. This iterative approach has a sequentially decreasing regularization parameter to reconstruct fine details of the image. Adapted to sparse images, this method performs worse than our method, in terms of image and PSF estimation errors. The PSF and image estimates from Almeida's method are presented in Figs. 5(c) and 6(d), respectively.

[2]Because this PSF is wrongly estimated and similar to the 2-D delta function, the image estimate looks similar to the denoised version of observation, so we omit the image estimate.

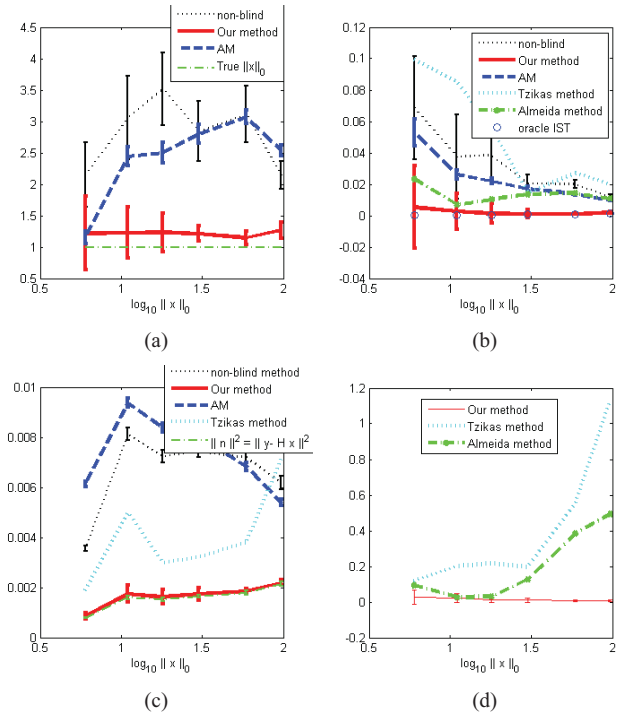


Fig. 9. Error bar graphs of results from our myopic deconvolution algorithm. For several image $\mathbf{x}$s of different sparsity levels, errors are illustrated with standard deviations. (Some of the sparsity measure and residual errors are too large to be plotted together with results from other algorithms.) (a) $\|\hat{\mathbf{x}}\|_0/\|\mathbf{x}\|_0$: estimated sparsity. Normalized true level is 1. (b) $\|(\mathbf{x}/\|\mathbf{x}\|) - (\hat{\mathbf{x}}/\|\hat{\mathbf{x}}\|)\|_2^2/\|\mathbf{x}\|_0$: normalized error in reconstructed image. (c) $\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2/\|\mathbf{x}\|_0$: residual (projection) error. (d) $\|(\hat{\boldsymbol{\kappa}}/\|\hat{\boldsymbol{\kappa}}\|) - (\boldsymbol{\kappa}/\|\boldsymbol{\kappa}\|)\|_2^2$, as a performance gauge of our myopic method. At the initial stage of the algorithm, $\|(\boldsymbol{\kappa}_0/\|\boldsymbol{\kappa}_0\|) - (\boldsymbol{\kappa}/\|\boldsymbol{\kappa}\|)\|_2^2 = 0.5627$.

Tzikas *et al.* [25] use a similar PSF model to our method using basis kernels. However, no sparse image prior was assumed in [25] making it unsuitable for sparse reconstruction problems such as the MRFM problem considered in the paper. For a fair comparison, we applied the suggested PSF model [25] along with our sparse image prior. The results of PSF and image estimation and the performance graph are shown in Figs. 5(d), 6(e), and 9, respectively. In terms of PSF estimation error, our algorithm outperforms the others.

We also compared against the mixture model-based algorithm of Fergus *et al.* [22], and the related method of Shan *et al.* [23], which are proposed for blind deconvolution of shaking/motion blurs and do not incorporate any sparsity penalization. When applied to the sparse MRFM reconstruction problem, the algorithms in [22] and [23] performed very poorly (produced divergent or trivial solutions, not shown due to space limitations). This poor performance is likely due to the fact that the shape of the MRFM PSF and the sparse image model are significantly different from those in blind deconvolution of camera shaking/motion blurs. The generalized PSF model in [22] and [23] with the sparse image prior is Tzikas' model [25], which is described above.

We used the iterative shrinkage/thresholding (IST) algorithm with a true PSF to lower bound our myopic reconstruction algorithm. The IST algorithm effectively

TABLE II
COMPUTATION TIME OF ALGORITHMS (IN SECONDS), FOR THE DATA IN FIG. 1

Proposed method	19.06
Bayesian nonblind [12]	3.61
IST [34]	0.09
AM [24]	0.40
Almeida's method [33]	5.63
Amizic's method [32]	5.69
Tzikas' method [25]	20.31

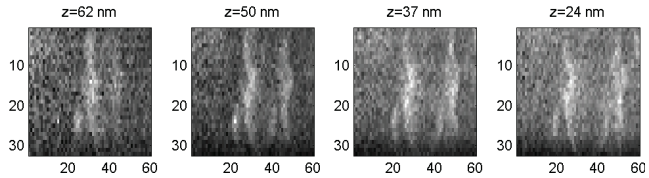


Fig. 10. Observed data at various tip-sample distances z.

reconstructs images with a sparsity constraint [34]. From Fig. 9(b) the performance of our result is as good as that of the oracle IST. In Table II, we present comparison of computation time[3] of the proposed sparse semi-blind Bayes reconstruction to that of several other algorithms.

C. Application to 3-D MRFM Image Reconstruction

In this section, we apply the semi-blind Bayesian reconstruction algorithm to the 3-D MRFM tobacco virus data [6] shown in Fig. 10. The necessary modification for our algorithm to apply to 3-D data is simple because the extension of our 2-D pixel-wise sampling method requires only one more added dimension to extend to 3-D basis vectors and 3-D convolution kernel. As seen in Appendix A, the voxel-wise update of a vectorized image $\mathbf{x}$ can be generalized to n-D data. This scalability is another benefit of our algorithm. The implementation of the AM algorithm is impractical due to its slow convergence rates [24]. Here, we only consider Bayesian methods. The additive noise is assumed Gaussian consistently with [4] and [6], so the noise model in Section III-A is applied here.

The PSF basis vectors were obtained from a standard PCA and the number of principal components in the PSF perturbation was selected as 4 based on detecting the knee in a scree plot. The same interpolation method as used in [12] was adopted to account for unequal spatial sampling rates in the supplied data for the PSF domain and the image domain.

In the PSF and image domains, along the z-axis, the grid in PSF signal space is three times finer than the observation sampling density, because the PSF sampling rate along the z-axis is three times higher than the data sampling rate. To interpolate this lower sampled data, we implemented a version of the Bayes MC reconstruction that compensates for unequal projection sampling in the z-directions using the interpolation procedure of Dobigeon *et al.* [12].

[3]MATLAB is used under Windows 7 Enterprise and HP-Z200 (Quad 2.66 GHz) platform.

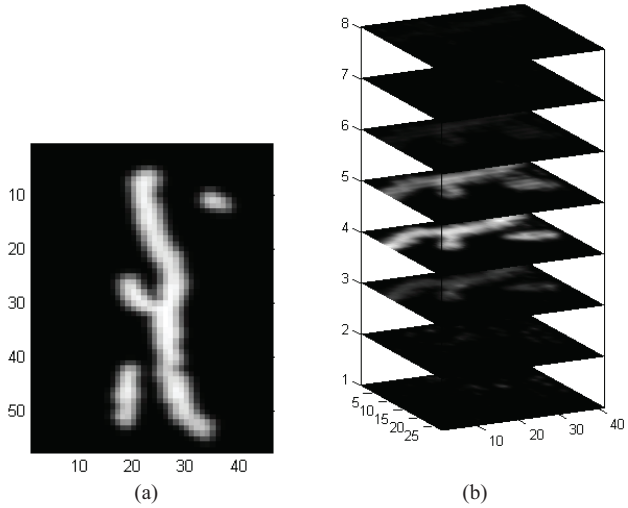


Fig. 11. Results of applying the proposed semi-blind sparse image reconstruction algorithm to the synthetic 3-D MRFM virus image. (a) Ground truth synthetic virus image obtained from data by Degen *et al*. [6]. (b) Semi-blind reconstruction of the synthetic virus data. Only the z-planes that have nonzero image intensity are shown.

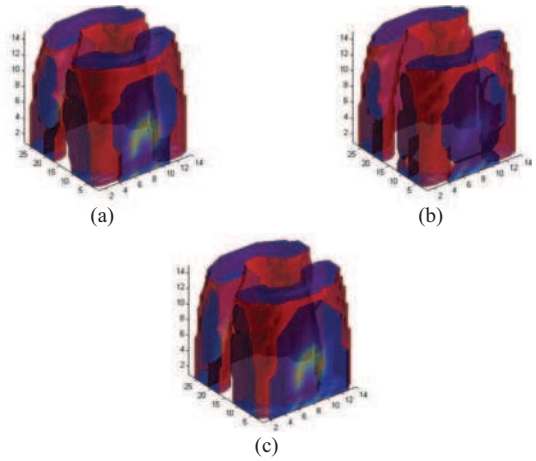


Fig. 12. PSF estimation result. (a) True PSF. (b) Initial mismatched PSF. (c) Estimated PSF.

To demonstrate that the proposed 3-D MCMC semi-blind reconstruction algorithm is capable of reconstruction in the presence of significant MRFM PSF mismatch, we first applied it to a simulated version of the experimental data shown in Fig. 10. We used the scanning electron microscope virus image reported by Degen *et al.* [6] to create a synthetic 3-D MRFM virus image, one slice of which is shown in Fig. 11(a). This 3-D image was then passed through the MRFM forward model, shown in Fig. 12(a), and 10 dB Gaussian noise was added. The mismatched PSF depicted in Fig. 12(b) was used to initialize our proposed semi-blind reconstruction algorithm. After 40 iterations, the algorithm reduced the initial normalized PSF error $\|(\boldsymbol{\kappa}_0/\|\boldsymbol{\kappa}_0\|) - (\boldsymbol{\kappa}/\|\boldsymbol{\kappa}\|)\|^2$ from 0.7611 to $\|(\hat{\boldsymbol{\kappa}}/\|\hat{\boldsymbol{\kappa}}\|) - (\boldsymbol{\kappa}/\|\boldsymbol{\kappa}\|)\|^2 = 0.0295$. Figs. 11(b) and 12(c) show the estimated image and the estimated PSF, respectively.

We next applied the proposed semi-blind reconstruction algorithm to the actual experimental data shown in Fig. 10 for which there is neither ground truth on the MRFM image

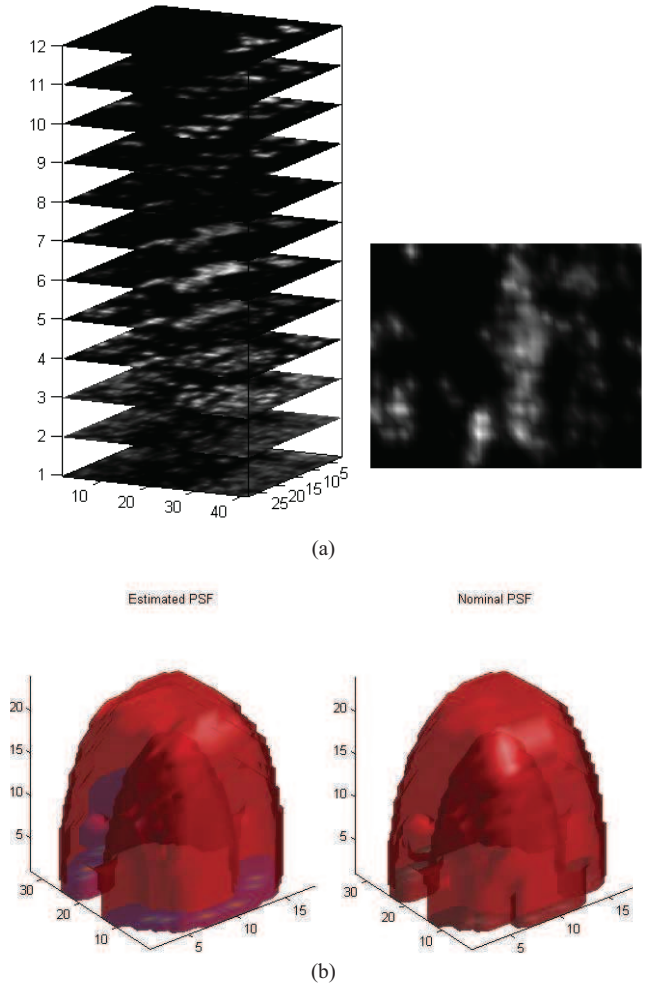


Fig. 13. Semi-blind MC Bayes method results and PSF coefficient curves. $\Delta z = 4.3$ nm, and pixel spacing is 8.3 nm $\times$ 16.6 nm in $x \times y$, respectively. The size of (x, y) plane is 498 nm $\times$ 531.2 nm. Smoothing is applied for visualization. (a) MAP estimate in 3-D and the estimated image on the sixth plane, showing a virus particle. (b) Estimated (left) and nominal (right) PSFs. $\|(\hat{\boldsymbol{\kappa}}/\|\hat{\boldsymbol{\kappa}}\|) - (\boldsymbol{\kappa}_0/\|\boldsymbol{\kappa}_0\|)\|^2 = 0.0212$. The difference between the two is small. (Hard thresholding with level $= \max(PSF) \times 10^{-4}$ is applied for visualization.)

nor on the MRFM PSF. The image reconstruction results are shown in Fig. 13. The small difference between the nominal PSF and the estimated PSF indicates that the estimated PSF is close to the assumed PSF. We empirically validated this small difference by verifying that multiple runs of the Gibbs sampler gave low variance PSF residual coefficients. We conclude from this finding that the PSF model of Degen *et al.* [6] is in fact nearly Bayes optimal for these experimental data. The blind image reconstruction shown in Fig. 13 is similar to the image reconstruction in Degen *et al.* [6] obtained from applying the Landweber reconstruction algorithm with the nominal PSF.

Using the MCMC generated posterior distribution obtained from the experimental MRFM data, we generated confidence intervals, posterior mean, and posterior variance of the pixel intensities of the unknown virus image. The posterior mean and variance are presented in Fig. 14 for selected pixels. In addition, to demonstrate the match between the estimated region occupied by the virus particle and the actual region, we evaluated Bayesian p-values for object regions. The

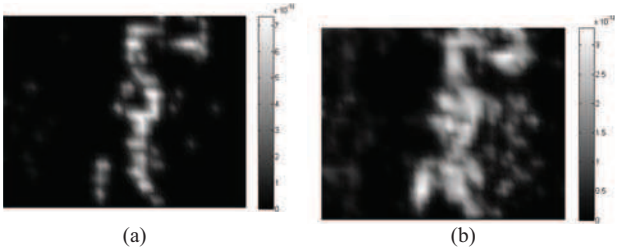


Fig. 14. Posterior mean and variance at the sixth plane of the estimated image, as shown in Fig. 13(a). (a) MMSE solution. Gray level image intensity values range from 0 (black) to 7.34×10^{-12} (white). (b) Pixel-wise square root of the image variance. White color indicates a high-variance. Gray level image intensity values range from 0 (black) to 3.29×10^{-12} (white).

Bayesian p-value for a specific region R_i having nonzero intensity is $pv(R_i) = \mathbb{P}(\{I_k = 1\}_{k \in R_i} | \mathbf{y})$ where $\mathbb{P}$ is a probability measure and I_k is an indicator function at the kth voxel. Assuming voxel-wise independence, the p-values are easily computed from the posterior distribution and provide a level of a posteriori confidence in the statistical significance of the reconstruction. We found that over 95% of the Bayesian p-values were greater than 0.5 for the nonzero regions of the reconstruction.

D. Discussion

Joint identifiability is a common issue underlying all blind deconvolution methods. (e.g., scale ambiguity.) Even though the unicity of our solution is not proven, given the conditions that: 1) $\text{span}(\boldsymbol{\kappa}) = \boldsymbol{\kappa}_0 + \text{span}(\sum \boldsymbol{\kappa}_i)$ does not cover a kernel of a delta function, $\boldsymbol{\kappa} = \delta(\cdot)$ and 2) the PSF solution is restricted to this linear space of κ_0, κ_i, the MAP estimate computed from (17) is a reasonable solution that is not trivial the MAP criteria promises a reasonable solution that is not trivial such as $\hat{\mathbf{x}} = \mathbf{y}$. Due to this restriction and the sparse nature of the image to be estimated, we can reasonably expect that the solution provided by the algorithm is close to the true PSF. A study of unicity of the solution would be worthwhile but is beyond the scope of this paper as it would require study of the complicated and implicit fixed points of the proposed Bayes objective function.

Note that proposed sparse image reconstruction algorithm can be extended to exploit sparsity in other domains, such as the wavelet domain. In this case, if we define $\mathbf{W}$ to be the transformation matrix on $\mathbf{x}$, the proposed semi-blind approach can be applied to reconstruct the transformed signal $\mathbf{z} = \mathbf{W}\mathbf{x}$. However, instead of assigning the single-sided exponential distribution as prior for $\mathbf{z}$, a double-sided Laplacian distribution might be used to cover the negative values of the pixels. The estimation procedure for PSF coefficients, noise level, and hyperparameters would be the same. For image estimation, the vector $\mathbf{h}_i$ used in (22) would be replaced with the ith column of $\mathbf{H}\mathbf{W}^{-1}$.

VI. Conclusion

We have proposed an extension of the method of the nonblind Bayes reconstruction by Dobigeon *et al.* [12] that simultaneously estimates a 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, with an adaptive mechanism for tuning random-walk step size for fast convergence. Our approach can be used to empirically evaluate the accuracy of assumed nominal PSF models in the presence of model uncertainty. In our sparse reconstruction simulations, we demonstrated that the semi-blind Bayesian algorithm has improved performance as compared to the AM reconstruction and other blind deconvolution algorithms and nonblind Bayes method under several criteria.

Possible extensions of the proposed method may include enforcing sparsity constraints on the result PSF and the eigenfunctions, by using sparse PCA type algorithms. Also, even with the selective sampling strategy that speeds up the sampling, the MCMC methods are slower than nonstochastic methods. This will be addressed in future 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 $T(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i)$ where $\tilde{\mathbf{x}}_i$ is the vector $\mathbf{x}$ whose ith element has been replaced by 0. Moreover, sampling according to the conditional posterior distributions of σ^2 and λ_k (23) and (26) requires computations of $T(\boldsymbol{\kappa}, \mathbf{x})$.

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

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

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

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

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

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

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

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,

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

At iteration t of the Gibbs sampler, for $i = 1, \ldots, M$, 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 $\mathbf{h}_i = \mathbf{h}_i^{(0)} + \sum_{k=1}^{K} \lambda_k \mathbf{h}_i^{(k)}$,
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 [6],
5: draw $x_i^{(t)}$ according to (21),
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 $T\left(\boldsymbol{\kappa}, \mathbf{x}^{(t,i)}\right) = T\left(\boldsymbol{\kappa}, \tilde{\mathbf{x}}_i^{(t,i-1)}\right) + x_i^{(t)} \mathbf{h}_i$.

evaluating $T\left(\boldsymbol{\kappa}\left(\lambda_k^\star\right), \mathbf{x}\right)$ in the acceptance probability (25) can be recursively evaluated as follows:

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

Appendix B
Sparsity Enforcing Selective Sampling

Since we have estimated the "overall sparsity," $1 - \hat{w}$ of $\mathbf{x}$ from (19), we can expedite the sampling procedure of $\mathbf{x}$ by selectively sampling only significant portions of entire pixels of $\mathbf{x}$. As a result, we expect $(1 - \hat{w}) \times 100\%$ of pixel domain of $\mathbf{x}$ to be zero, which will not need to be resampled.

At time t, in order to approximate the quantile $(1 - \hat{w})$ of $\{w_i^{(t)}\}_{i=1,\ldots,M}$ in (21) we evaluate the $(1-\hat{w})$ quantile value of the previously obtained $\{w_i^{(t-1)}\}_{i=1,\ldots,M}$. This approximation accelerates the computation because the exact calculation of $\{w_i^{(t)}\}_{i=1,\ldots,M}$ requires sampling of all x_is. Let $q = \text{quantile}(\{w_i^{(t-1)}\}_{i=1,\ldots,M}, 1 - \hat{w})$ and $w_{\text{thr}} = \max(q, 1 - \hat{w})$. When $w_i^{(t)}$ for $x_i^{(t)}$ from (21) is less than w_{thr}, then $x_i^{(t)}$ is not updated or is set to zero. Because MCMC sampling is computationally expensive, especially for large size images, this suggestion can be restricted to the burn-in period to save computations.

In our experiment, the selective sampling of $\mathbf{x}$ applied after third or fourth iterations produces equally good results compared to the conventional MCMC sampling methods, while reducing computation time by 30%–50% for nonblind sparse reconstruction with a fixed PSF and by 10%–30% for the semi-blind sparse reconstruction.

Appendix C
Adaptive Tuning of an Acceptance Rate in the Random-Walk Sampling

For an efficient sampling of $\lambda_k, k = 1, \ldots, K$, from the desired distribution $\pi(\lambda_k) = f(\lambda_k|\boldsymbol{\lambda}_{-k}, \mathbf{x}, \sigma^2, \mathbf{y})$, we need to

Algorithm 4 Tuning s in the Gaussian Proposal Density $q(\cdot,\cdot)$

Select upper and lower limits acc_H and acc_L. At each time $t = W, 2W, 3W, \ldots$, tune s via the following steps:

1: Evaluate acc_s using (31) for the given time-frame window,

2: Update $s \leftarrow \begin{cases} s \times c, & \text{if } \text{acc}_s > \text{acc}_H, \\ s \div c, & \text{if } \text{acc}_s < \text{acc}_L, \\ s, & \text{otherwise.} \end{cases}$

properly tune the acceptance rate of the samples from the proposal distribution. A careful selection of a step size is critical for convergence of the method. For example, if the step size is too large, most of the iterations will be rejected and the sampling algorithm will be inefficient. On the other hand, if the step size is too small, most of the random walk moves are accepted but these moves are slow to cover the probable space of the desired distribution, and the method is once again inefficient.

The transition density of Metropolis–Hastings sampling is $q(\lambda^{(t)}, \lambda^{\star(t)})\text{acc}(\lambda^{(t)}, \lambda^{\star(t)})$, where $q(\lambda^{(t)}, \lambda^{\star(t)})$ is the proposal density from $\lambda^{(t)}$ and $\text{acc}(\lambda^{(t)}, \lambda^{\star(t)})$ is the acceptance probability for the move from $\lambda^{(t)}$ to $\lambda^{\star(t)}$. Here, we denote λ_k by λ without a subscript for simplicity. We set $q(\lambda^{(t)}, \lambda^{\star(t)})$ to be a Gaussian density function of $\lambda^{\star(t)}$, denoted by $q(\lambda^{(t)}, \lambda^{\star(t)}) = q(\lambda^{\star(t)} - \lambda^{(t)}) = \phi(\lambda^{\star(t)}; \lambda^{(t)}, s^2)$ with a mean $\lambda^{(t)}$ and a variance s^2, which produces a random walk sample path. Since $q(\cdot,\cdot)$ is symmetrical, $\text{acc}_s(\lambda^{(t)}, \lambda^{\star(t)}) = \min(1, (\pi(\lambda^{\star(t)})q(\lambda^{\star(t)}, \lambda^{(t)})/\pi(\lambda^{(t)})q(\lambda^{(t)}, \lambda^{\star(t)}))) = \min(1, (\pi(\lambda^{\star(t)})/\pi(\lambda^{(t)}))) = \rho_{\lambda^{(t)} \to \lambda^{\star(t)}}$, as derived in (24). Then the acceptance probability from a parameter value $\lambda^{(t)}$ is $\text{acc}_s(\lambda^{(t)}) = \int_{\lambda^{\star(t)}} q(\lambda^{(t)}, \lambda^{\star(t)})\text{acc}_s(\lambda^{(t)}, \lambda^{\star(t)})d\lambda^{\star(t)}$. The acceptance rate with a scale parameter s, acting as a step size, can be expressed as $\text{acc}_s = \int_\lambda \pi(\lambda)\text{acc}_s(\lambda)d\lambda$.

We evaluate these integrations by using Monte Carlo methods, $\text{acc}_s \approx (1/n_1)\sum_{t=1}^{n_1} \text{acc}_s(\lambda^{(t)})$ with $\lambda^{(t)} \sim \pi(\lambda^{(t)})$, and $\text{acc}_s(\lambda^{(t)}) \approx (1/n_2)\sum_{t=1}^{n_2} \text{acc}_s(\lambda^{(t)}, \lambda^{\star(t)})$ with $\lambda^{\star(t)} \sim q(\lambda^{(t)}, \lambda^{\star(t)})$. In practice, this empirical version of the integration value is evaluated as

$$\text{acc}_s \approx \frac{1}{W}\sum_{t=1}^{W} \text{acc}_s\left(\lambda^{(t)}\lambda^{\star(t)}\right) \tag{31}$$

after the burn-in period. Therefore, we can evaluate the acceptance rate with s by averaging the Boolean variables of $\text{acc}_s(\lambda^{(t)}, \lambda^{\star(t)}), t = 1, \ldots, W$, over a time-frame window of length W with realizations $\{\lambda^{(t)}, \lambda^{\star(t)}\}_t$. In short, we iteratively update s to achieve an appropriate acceptance rate acc_s as described in Algorithm 4.

In practice, we fix the variance of the instrumental distribution at the end of a burn-in period. Consequently, the transition kernel will be fixed and this guarantees both ergodicity and stationary distribution. In our experiment, we set a conservative acceptance range, $\text{acc}_H = 0.6, \text{acc}_L = 0.4$, referring to [29], and $W = 20, c = 4$. This strategy can also be applied to the direct parameter estimation described in Appendix D.

Appendix D
Direct Sampling of PSF Parameter Values

As described in Section I, in the MRFM experiments, the direct estimation of PSF parameters is difficult because of the nonlinearity of $\boldsymbol{\kappa}_{\text{gen}}$ and the slow evaluation of $\boldsymbol{\kappa}_{\text{gen}}(\boldsymbol{\zeta}')$ given a candidate value $\boldsymbol{\zeta}'$. However, if $\boldsymbol{\kappa}_{\text{gen}}$ is simple and evaluated quickly, then a direct sampling of parameter values can be performed. To apply this sampling method, instead of calculating a linearized convolution kernel $\boldsymbol{\kappa}(\boldsymbol{\lambda})$, we evaluate the exact model PSF, $\boldsymbol{\kappa}_{\text{gen}}(\boldsymbol{\zeta})$, in (23) and (25). Also, the proposed parameter vector $\boldsymbol{\zeta}^\star$ correspondingly replaces a coefficient vector $\lambda^\star$ and the updated PSF is used in the estimation of other variables. This strategy turns out to be similar to the approach adopted by Orieux *et al.* [26].

Acknowledgment

The authors would like to thank D. Rugar for providing the tobacco virus data and insightful comments on this paper. They would also like to thank the reviewers for their comments, which helped the authors significantly improve the quality of this paper.

References

[1] J. A. Sidles, "Noninductive detection of single-proton magnetic resonance," *Appl. Phys. Lett.*, vol. 58, no. 24, pp. 2854–2856, Jun. 1991.

[2] J. A. Sidles, "Folded Stern-Gerlach experiment as a means for detecting nuclear magnetic resonance in individual nuclei," *Phys. Rev. Lett.*, vol. 68, no. 8, pp. 1124–1127, Feb. 1992.

[3] J. A. Sidles, J. L. Garbini, K. J. Bruland, D. Rugar, O. Züger, S. Hoen, and C. S. Yannoni, "Magnetic resonance force microscopy," *Rev. Mod. Phys.*, vol. 67, no. 1, pp. 249–265, Jan. 1995.

[4] D. Rugar, C. S. Yannoni, and J. A. Sidles, "Mechanical detection of magnetic resonance," *Nature*, vol. 360, no. 6404, pp. 563–566, Dec. 1992.

[5] 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.

[6] 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.

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

[8] O. Züger and D. Rugar, "First images from a magnetic resonance force microscope," *Appl. Phys. Lett.*, vol. 63, no. 18, pp. 2496–2498, 1993.

[9] 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.

[10] 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. 1–5, Feb. 2009.

[11] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction for molecular imaging," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1215–1227, Jun. 2009.

[12] N. Dobigeon, A. O. Hero, and J.-Y. Tourneret, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2059–2070, Sep. 2009.

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

[14] S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline, "Semi-blind deconvolution of neural impulse response in fMRI using a Gibbs sampling method," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal*, vol. 5. May 2004, pp. 601–604.

[15] G. Pillonetto and C. Cobelli, "Identifiability of the stochastic semi-blind deconvolution problem for a class of time-invariant linear systems," *Automatica*, vol. 43, no. 4, pp. 647–654, Apr. 2007.

[16] P. Sarri, G. Thomas, E. Sekko, and P. Neveux, "Myopic deconvolution combining Kalman filter and tracking control," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal*, vol. 3. May 1998, pp. 1833–1836.
[17] G. Chenegros, L. M. Mugnier, F. Lacombe, and M. Glanc, "3D phase diversity: A myopic deconvolution method for short-exposure images: Application to retinal imaging," *J. Opt. Soc. Amer. A*, vol. 24, no. 5, pp. 1349–1357, May 2007.
[18] R. Molina, "On the hierarchical Bayesian approach to image restoration: Applications to astronomical images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 11, pp. 1122–1128, Nov. 1994.
[19] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, and A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially known blurs," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1784–1797, Oct. 2000.
[20] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, A. K. Katsaggelos, and J. Mateos, "Hyperparameter estimation in image restoration problems with partially known blurs," *Opt. Eng.*, vol. 41, no. 8, pp. 1845–1854, 2002.
[21] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.
[22] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Proc. ACM SIGGRAPH Papers*, 2006, pp. 787–794.
[23] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, p. 1, 2008.
[24] K. Herrity, R. Raich, and A. O. Hero, "Blind reconstruction of sparse images with unknown point spread function," *Proc. SPIE Comput. Imaging*, vol. 6814, no. 1, pp. 68140K-1–68140K-11, Jan. 2008.
[25] D. Tzikas, A. Likas, and N. Galatsanos, "Variational Bayesian sparse kernel-based blind image deconvolution with Student's-t priors," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 753–764, Apr. 2009.
[26] F. Orieux, J.-F. Giovannelli, and T. Rodet, "Bayesian estimation of regularization and point spread function parameters for Wiener-Hunt deconvolution," *J. Opt. Soc. Amer. A*, vol. 27, no. 7, pp. 1593–1607, Jul. 2010.
[27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag, Aug. 2003.
[28] M. Ting, R. Raich, and A. O. Hero, "Sparse image reconstruction using sparse priors," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 1261–1264.
[29] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.
[30] P. Bremaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag, 1999.
[31] J.-M. Marin and C. P. Robert, *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. New York: Springer-Verlag, 2007.
[32] B. Amizic, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Sparse Bayesian blind image deconvolution with parameter estimation," in *Proc. Eur. Signal Process. Conf.*. Aalborg, Denmark, Aug. 2010, pp. 626–630.
[33] M. Almeida and L. Almeida, "Blind and semi-blind deblurring of natural images," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 36–52, Jan. 2010.
[34] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.



Se Un Park received the B.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2004, and the M.S. and M.A. degrees in electrical engineering and statistics from the University of Michigan, Ann Arbor, in 2009 and 2011, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering.

His current research interests include statistical signal processing, image processing, and machine learning.



Nicolas Dobigeon (S'05–M'08) was born in Angoulême, France, in 1981. He received the Eng. degree in electrical engineering from École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIHT), Toulouse, France, and the M.Sc. and Ph.D. degrees in signal processing from the National Polytechnic Institute of Toulouse, Toulouse, in 2004, 2004, and 2007, respectively.

He was a Post-Doctoral Research Associate with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, from 2007 to 2008. Since 2008, he has been an Assistant Professor with ENSEEIHT, National Polytechnic Institute of Toulouse, University of Toulouse, Toulouse, and with the Signal and Communication Group, IRIT Laboratory, Toulouse. His current research interests include statistical signal and image processing with a particular interest in Bayesian inference and Markov chain Monte Carlo methods.



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

He has been with the University of Michigan, Ann Arbor, since 1984, where he is the R. Jamison and B. Williams Professor of engineering. His primary appointment is with the Department of Electrical Engineering and Computer Science, and he has appointments with the Department of Biomedical Engineering and the Department of Statistics. He has held visiting positions with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, in 2006, Boston University in 2006, I3S University of Nice, Sophia-Antipolis, France, in 2001, École Normale Supérieure de Lyon, Lyon, France, in 1999, École Nationale Supérieure des Télécommunications, Paris, France, in 1999, Lucent Bell Laboratories, Murray Hill, NJ, in 1999, Scientific Research Laboratory, Ford Motor Company, Dearborn, MI, in 1993, École Nationale Supérieure des Techniques Avancees, École Supérieure d'Electricite, Paris, in 1990, and M.I.T. Lincoln Laboratory, Cambridge, from 1987 to 1989. His current research interests include detection, classification, pattern analysis, adaptive sampling for spatiotemporal data, applications to network security, multimodal sensing and tracking, biomedical imaging, and genomic signal processing.

Dr. Hero was a recipient of the Digiteo Chaire d'Excellence from Digiteo Research Park, École Supérieure d'Electricite, France, and the University of Michigan Distinguished Faculty Achievement Award in 2011. He received several Best Paper Awards, including the IEEE Signal Processing Society Best Paper Award in 1998, the Best Original Paper Award from the Journal of Flow Cytometry in 2008, and the Best Magazine Paper Award from the IEEE Signal Processing Society in 2010. He received the IEEE Signal Processing Society Meritorious Service Award in 1998, the IEEE Third Millennium Medal in 2000, and the IEEE Signal Processing Society Distinguished Lecturership in 2002. He has been a plenary and keynote speaker at major workshops and conferences. He was the President of the IEEE Signal Processing Society from 2006 to 2007. He was on the Board of Directors of the IEEE from 2009 to 2011, where he served as the Director Division IX of signals and applications.