NONNEGATIVE SIGNAL RECONSTRUCTION FROM COMPRESSIVE SAMPLES VIA A DIFFERENCE MAP ECME ALGORITHM

Kun Qiu and *Aleksandar Dogandzic*

ECpE Department, Iowa State University, Ames, IA 50011
email: {kqiu,ald}@iastate.edu

ABSTRACT

We develop an approximate maximum likelihood (ML) scheme for reconstructing nonnegative sparse signals from compressive samples. The measurements follow an underdetermined linear model, where the regression vector is modeled as the sum of an unknown deterministic nonnegative signal with sparse transform coefficients and a zero-mean white Gaussian component with an unknown variance. We first derive an expectation-conditional maximization either (ECME) algorithm that aims at maximizing the likelihood function with respect to the unknown parameters and then employ a difference map iteration to approximate the maximization (M) step of the ECME iteration. We compare the proposed and existing large-scale sparse signal reconstruction methods via numerical simulations and demonstrate that, by exploiting both the nonnegativity of the underlying image and the sparsity of its wavelet coefficients, we can reconstruct this image using a significantly smaller number of measurements than the existing methods.

***Index Terms*—** Compressive sampling, sparse signal reconstruction, nonnegative signal, expectation-conditional maximization either (ECME) algorithm, difference map.

1. INTRODUCTION

Compressive sampling or compressed sensing exploits the fact that most natural signals are well described by only a few significant coefficients in some [e.g. discrete wavelet transform (DWT)] domain, where the number of significant coefficients is much smaller than the signal size [1–3]. Therefore, for an $m \times 1$ vector $\boldsymbol{x}$ representing the signal and an appropriate $m \times m$ sparsifying transform matrix Ψ, we have $\boldsymbol{x} = \Psi\, \boldsymbol{s}$, where $\boldsymbol{s}$ is an $m \times 1$ signal transform-coefficient vector with most elements having negligible magnitudes. The idea behind compressive sampling is to *sense* the significant components of $\boldsymbol{s}$ using a small number of linear measurements: $\boldsymbol{y} = \Phi\, \boldsymbol{x} = \Phi\, \Psi\, \boldsymbol{s}$, where $\boldsymbol{y}$ is an $N \times 1$ measurement vector and Φ is a known $N \times m$ *sampling matrix* with $N \leq m$. We refer to the composite matrix $\Phi\, \Psi$ as the *sensing matrix*.

This work was supported by NSF under Grant CCF-0545571.

In [4] and [5], the signal transform coefficients $\boldsymbol{s}$ were assumed to be *both* nonnegative and sparse. In this paper, we consider nonnegative signals $\boldsymbol{x} = \Psi\, \boldsymbol{s}$ with sparse transform coefficients $\boldsymbol{s}$. This scenario is of significant practical interest and has immediate applications in X-ray computed tomography (CT) and magnetic resonance imaging (MRI). For example, in X-ray CT, the underlying image $\boldsymbol{x}$ corresponds to the tissue or material attenuation coefficients [6, Ch. 3.2], which are clearly nonnegative. Many inherently nonnegative signals are also encountered in spectroscopy, tomography, DNA microarrays, network monitoring, and hidden Markov models [5, 7, 8]. However, most such signals are *not* approximately sparse in the identity basis $\Psi = I_m$, where I_m denotes the identity matrix of size m. Therefore, the nonnegative sparse signal model with the general sparsifying transform Ψ is *practically more useful and challenging* than that in [4] and [5]: It allows the signal of interest to be nonnegative *as well as* sparse in the appropriate transform domain. In [8], Harmany *et al*. have recently considered a similar nonnegative sparse signal model and developed a convex-relaxation sparse Poisson-intensity reconstruction algorithm (SPIRAL) that assumes Poisson measurements and nonnegative elements of the sampling matrix. The SPIRAL method in [8] can be adapted to the standard compressive sampling scenario by replacing the Poisson measurement model with the Gaussian signal-plus-noise model. Here, we propose a probabilistic measurement model and derive a *difference map expectation-conditional maximization either (DM-ECME)* algorithm that approximately computes the maximum likelihood (ML) parameter estimates of the unknown parameters. We apply our algorithm to simulated CT data and demonstrate that, by exploiting both the nonnegativity of the underlying image and the sparsity of its wavelet coefficients, we can achieve perfect reconstruction using a significantly smaller number of measurements than the state-of-art sparse reconstruction methods that exploit signal sparsity only.

We introduce the notation: $\mathcal{N}(\boldsymbol{y}\,;\, \boldsymbol{\mu}, \Sigma)$ denotes the multivariate probability density function (pdf) of a real-valued Gaussian random vector $\boldsymbol{y}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ; $\|\cdot\|_{\ell_p}$ and "^{T}" are the ℓ_p norm and transpose; I_m and $\mathbf{0}_{m\times 1}$ are the identity matrix of size m and the $m \times 1$ vector of zeros; $\mathcal{P}_{\mathcal{S}}(\boldsymbol{a}) = \arg\min_{\boldsymbol{s}\in\mathcal{S}} \|\boldsymbol{a} - \boldsymbol{s}\|_{\ell_2}^2$ denotes

the minimum-distance projection of a vector $\boldsymbol{a}$ onto the constraint set $\mathcal{S}$; the sparse thresholding operator $\mathcal{T}_r(\boldsymbol{s})$ keeps the r largest-magnitude elements of a vector $\boldsymbol{s}$ intact and sets the rest to zero; the nonnegative thresholding operator $\boldsymbol{s}^+$ keeps the nonnegative elements of $\boldsymbol{s}$ intact and sets the rest to zero.

2. PROBABILISTIC MEASUREMENT MODEL AND THE ECME ALGORITHM

We model a $N \times 1$ real-valued measurement vector $\boldsymbol{y}$ as

$$\boldsymbol{y} = \Phi\, \boldsymbol{z} \tag{1a}$$

where Φ is the known $N \times m$ full-rank sampling matrix ($N \leq m$) and $\boldsymbol{z}$ is an $m \times 1$ multivariate Gaussian vector with pdf

$$p_{\boldsymbol{z}\,|\,\boldsymbol{\theta}}(\boldsymbol{z}\,|\,\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{z};\ \Psi\, \boldsymbol{s}, \sigma^2\, I) \tag{1b}$$

Ψ is the known $m \times m$ orthogonal sparsifying transform matrix satisfying $\Psi\ \Psi^T = \Psi^T\ \Psi = I_m$, $\boldsymbol{s}$ is an *unknown* $m \times 1$ real-valued sparse signal transform coefficients vector and σ^2 is an unknown *variance-component parameter*. The set of unknown parameters is $\boldsymbol{\theta} = (\boldsymbol{s}, \sigma^2) \in \Theta$ with the parameter space $\Theta = \mathcal{S}_r^+ \times [0, +\infty)$ where

$$\mathcal{S}_r^+ = \{\boldsymbol{s} \in \mathbb{R}^m :\ \|\boldsymbol{s}\|_{\ell_0} \leq r\,,\ \Psi\, \boldsymbol{s} \succeq \boldsymbol{0}_{m\times 1}\} \tag{2}$$

is the signal parameter space that enforces the signal nonnegativity and sparsity of the signal transform coefficients. We refer to r as the *signal sparsity level* and assume that it is *known*. The marginal likelihood function of $\boldsymbol{\theta}$ is obtained by *integrating* $\boldsymbol{z}$ *out* [see (1)]:

$$p_{\boldsymbol{y}\,|\,\boldsymbol{\theta}}(\boldsymbol{y}\,|\,\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{y};\ \Phi\ \Psi\ \boldsymbol{s}, \sigma^2\ \Phi\ \Phi^T). \tag{3}$$

Finding the exact ML estimate $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}} = \big(\widehat{\boldsymbol{s}}_{\mathrm{ML}}, \widehat{\sigma}^2_{\mathrm{ML}}\big) = \arg\max_{\boldsymbol{\theta}\in\Theta} p_{\boldsymbol{y}\,|\,\boldsymbol{\theta}}\big(\boldsymbol{y}\,|\,\boldsymbol{\theta}\big)$ involves a combinatorial search and is therefore intractable in practice. In the following section, we present an ECME iteration that aims at maximizing (3) and is the basis of the DM-ECME algorithm in Section 3.

2.1. ECME Algorithm

We treat $\boldsymbol{z}$ as the *missing (unobserved) data*, and the ECME algorithm maximizes *either* the expected complete-data log-likelihood function (where the expectation is computed with respect to the conditional distribution of the unobserved data given the observed measurements) *or* the actual observed-data log-likelihood (see [9, Ch. 5.7]) for estimating the unknown parameters $\boldsymbol{\theta}$.

Assume that the parameter estimate $\boldsymbol{\theta}^{(p)} = \big(\boldsymbol{s}^{(p)}, (\sigma^2)^{(p)}\big)$ is available, where p denotes the iteration index. The $(p+1)$ iteration of the ECME algorithm proceeds as follows:

- update the estimate of the signal transform coefficients using the expectation (E) step:

$$\boldsymbol{z}^{(p+1)} = \Psi\, \boldsymbol{s}^{(p)} + \Phi^T \left(\Phi\ \Phi^T\right)^{-1} \left(\boldsymbol{y} - \Phi\ \Psi\ \boldsymbol{s}^{(p)}\right) \tag{4a}$$

  followed by the maximization (M) step:

$$\boldsymbol{s}^{(p+1)} = \mathcal{P}_{\mathcal{S}_r^+}(\,\Psi^T\, \boldsymbol{z}^{(p+1)}); \tag{4b}$$

- update the variance component estimate using the conditional M (CM) step: $(\sigma^2)^{(p+1)} = (\boldsymbol{y} - \Phi\ \Psi\ \boldsymbol{s}^{(p+1)})^T\ (\Phi\Phi^T)^{-1}(\boldsymbol{y} - \Phi\Psi\boldsymbol{s}^{(p+1)})/N$ obtained by maximizing (3) with respect to σ^2 for a fixed $\boldsymbol{s} = \boldsymbol{s}^{(p+1)}$.

To derive the M step (4b), we have used the fact that Ψ is an orthogonal matrix. The M step requires finding the minimum-distance projection of the vector $\Psi^T\, \boldsymbol{z}^{(p+1)}$ onto the signal parameter space $\mathcal{S}_r^+$, which is the intersection of the sparsity and nonnegativity constraint sets: $\mathcal{S}_r^+ = \mathcal{S}_r \cap \mathcal{S}^+$, where the sparsity and nonnegativity constraint sets are given by

$$\mathcal{S}_r = \{\boldsymbol{s} \in \mathbb{R}^m :\ \|\boldsymbol{s}\|_{\ell_0} \leq r\,\} \tag{5a}$$

$$\mathcal{S}^+ = \{\boldsymbol{s} \in \mathbb{R}^m :\ \Psi\, \boldsymbol{s} \succeq \boldsymbol{0}_{m\times 1}\}. \tag{5b}$$

The intersection $\mathcal{S}_r \cap \mathcal{S}^+$ is not empty because it contains the zero vector $\boldsymbol{0}_{m\times 1}$ for any sparsity level $r \geq 0$. Although the M step is always well-defined, it is computationally intractable: In the following section, we approximate it via a difference map iteration.

3. DM-ECME ALGORITHM

We first describe the difference map iteration for approximating the M step and then summarize the proposed DM-ECME iteration.

The difference map algorithm, first proposed in [10] to solve the phase retrieval problem, is an iterative scheme which, upon convergence, finds a point in the intersection of two constraint sets $\mathcal{A}$ and $\mathcal{B}$, see also [11]. Typically, this algorithm is applied when the minimum-distance projections onto $\mathcal{A}$ and $\mathcal{B}$ individually are simple. Given the q-th iterate $\boldsymbol{a}^{(q)}$, the $(q+1)$-th difference map iteration proceeds as:

$$\boldsymbol{a}^{(q+1)} = \boldsymbol{a}^{(q)} + \beta\left[\mathcal{P}_{\mathcal{A}}\big(f_{\mathcal{B}}(\boldsymbol{a}^{(q)})\big) - \mathcal{P}_{\mathcal{B}}\big(f_{\mathcal{A}}(\boldsymbol{a}^{(q)})\big)\right] \tag{6}$$

where $f_{\mathcal{A}}(\boldsymbol{a}^{(q)}) = \mathcal{P}_{\mathcal{A}}(\boldsymbol{a}^{(q)}) - [\mathcal{P}_{\mathcal{A}}(\boldsymbol{a}^{(q)}) - \boldsymbol{a}^{(q)}]/\beta$, $f_{\mathcal{B}}(\boldsymbol{a}^{(q)}) = \mathcal{P}_{\mathcal{B}}(\boldsymbol{a}^{(q)}) + [\mathcal{P}_{\mathcal{B}}(\boldsymbol{a}^{(q)}) - \boldsymbol{a}^{(q)}]/\beta$, and β is a tuning parameter with typical value chosen as $0.6 \leq |\beta| \leq 1$. Iterate until convergence to a fixed point $\boldsymbol{a}^\star$, which satisfies $\mathcal{P}_{\mathcal{A}}\big(f_{\mathcal{B}}(\boldsymbol{a}^\star)\big) = \mathcal{P}_{\mathcal{B}}\big(f_{\mathcal{A}}(\boldsymbol{a}^\star)\big)$. If the intersection $\mathcal{A} \cap \mathcal{B}$ is not empty, then $\mathcal{P}_{\mathcal{B}}\big(f_{\mathcal{A}}(\boldsymbol{a}^\star)\big)$ must lie in this intersection and is reported as the output of the difference map iteration. The difference map iteration is known to be much more effective than the naïve alternating projection scheme: $\boldsymbol{a}^{(q+1)} = \mathcal{P}_{\mathcal{B}}\big(\mathcal{P}_{\mathcal{A}}\big(\boldsymbol{a}^{(q)}\big)\big)$. Indeed, the alternating projection scheme is likely to be trapped in stagnation points, whereas difference map is capable of escaping such traps [10, Sec. 1], [11, p. 418].

We now apply the above iteration to approximate the M step in (4b). Assume that $\boldsymbol{z}^{(p+1)}$ is available, computed in the E step (4a). In our problem (4b), the two constraint sets are $\mathcal{A} = \mathcal{S}_r$ and $\mathcal{B} = \mathcal{S}^+$, see (5a) and (5b). Note that the projections of an $m \times 1$ vector $\boldsymbol{a}$ to $\mathcal{S}_r$ and $\mathcal{S}^+$ alone are simple:

$$\mathcal{P}_{\mathcal{S}_r}(\boldsymbol{a}) = \mathcal{T}_r(\boldsymbol{a}), \quad \mathcal{P}_{\mathcal{S}^+}(\boldsymbol{a}) = \Psi^T\,(\Psi\, \boldsymbol{a})^+ \tag{7}$$

Input: $\boldsymbol{y}$, Φ, Ψ, r, β
Output: parameter estimate $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{s}}, \widehat{\sigma}^2)$
1: $\boldsymbol{s} \leftarrow \mathbf{0}_{m\times 1}, \sigma^2 \leftarrow \boldsymbol{y}^T (\Phi\, \Phi^T)^{-1}\, \boldsymbol{y}/N$
2: $\widehat{\boldsymbol{s}} \leftarrow \boldsymbol{s}, \widehat{\sigma}^2 \leftarrow \sigma^2$
3: **repeat**
4: $\quad \boldsymbol{z} \leftarrow \Psi\, \boldsymbol{s} + \Phi^T \left(\Phi\, \Phi^T\right)^{-1} (\boldsymbol{y} - \Phi\, \Psi\, \boldsymbol{s})$
5: $\quad \boldsymbol{a} \leftarrow \Psi^T\, \boldsymbol{z}$ {Initialize the difference map iteration}
6: $\quad$ **repeat**
7: $\quad\quad f_{\mathcal{S}_r}(\boldsymbol{a}) \leftarrow \mathcal{P}_{\mathcal{S}_r}(\boldsymbol{a}) - [\mathcal{P}_{\mathcal{S}_r}(\boldsymbol{a}) - \boldsymbol{a}]/\beta$
8: $\quad\quad f_{\mathcal{S}^+}(\boldsymbol{a}) \leftarrow \mathcal{P}_{\mathcal{S}^+}(\boldsymbol{a}) + [\mathcal{P}_{\mathcal{S}^+}(\boldsymbol{a}) - \boldsymbol{a}]/\beta$
9: $\quad\quad \boldsymbol{a} \leftarrow \boldsymbol{a} + \beta \left[\mathcal{P}_{\mathcal{S}_r}\big(f_{\mathcal{S}^+}(\boldsymbol{a})\big) - \mathcal{P}_{\mathcal{S}^+}\big(f_{\mathcal{S}_r}(\boldsymbol{a})\big)\right]$
10: $\quad$ **until** two consecutive $\boldsymbol{a}$ do not differ significantly
11: $\quad \boldsymbol{s} \leftarrow \mathcal{P}_{\mathcal{S}^+}\big(f_{\mathcal{S}_r}(\boldsymbol{a})\big)$
12: $\quad \sigma^2 \leftarrow (\boldsymbol{y} - \Phi\, \Psi\, \boldsymbol{s})^T (\Phi\, \Phi^T)^{-1} (\boldsymbol{y} - \Phi\, \Psi\, \boldsymbol{s})/N$
13: $\quad$ **if** $\sigma^2 < \widehat{\sigma}^2$ **then**
14: $\quad\quad \widehat{\boldsymbol{s}} \leftarrow \boldsymbol{s}, \widehat{\sigma}^2 \leftarrow \sigma^2$
15: $\quad$ **end if**
16: **until** two consecutive $\boldsymbol{s}$ do not differ significantly or max number of iterations is reached

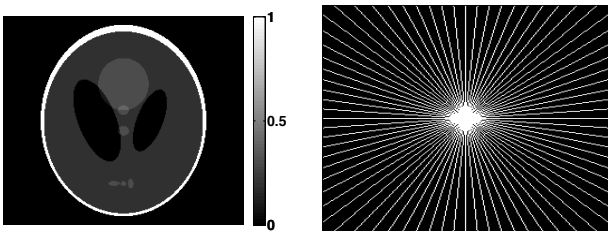
Fig. 1. Pseudo-code describing the DM-ECME algorithm.

where we used the fact that Ψ is an orthogonal matrix. We use (7) and the initial value $\boldsymbol{a}^{(0)} = \Psi^T\, \boldsymbol{z}^{(p+1)}$ to implement the difference map iteration (6). Upon convergence of this iteration, we obtain an estimate of the signal transform coefficient vector $\boldsymbol{s}^{(p+1)}$ that, in our experience, is much closer (in Euclidean distance) to $\Psi^T\, \boldsymbol{z}^{(p+1)}$ than the trivial feasible point $\mathbf{0}_{m\times 1}$. [Recall that the exact M step in (4b) seeks the $\boldsymbol{s}^{(p+1)}$ that is the closest to $\Psi^T\, \boldsymbol{z}^{(p+1)}$.]

Fig. 1 summarizes the DM-ECME algorithm in the pseudo-code format. Since the maximization step (4b) is only approximated by the difference map iteration, the DM-ECME iteration *does not* guarantee monotonic increase of the marginal likelihood function. Consequently, we introduce the condition in step 13 to check if the marginal likelihood function (3) evaluated at the new parameter estimates is higher than the largest likelihood achieved in the previous steps. If this condition holds, we update the parameter estimates.

4. NUMERICAL EXAMPLES

Consider the tomographic reconstruction of the Shepp-Logan phantom of size $m = 256^2$ in Fig. 2(a). All pixel values of the phantom image are nonnegative, representing attenuation coefficients. The tomographic measurements $\boldsymbol{y}$ were simulated using the 2-D discrete Fourier transform (DFT) coefficients of the phantom sampled over a star-shaped domain, as illustrated in Fig. 2(b), see also [2, 3]. The sampling matrix Φ is therefore constructed using selected rows of the DFT matrix that yield the corresponding DFT coefficients of the phantom image within the star-shaped domain. Consequently, the rows of Φ are *orthonormal*, i.e. $\Phi\, \Phi^T = I_N$, obviating the need to compute and store $(\Phi\, \Phi^T)^{-1}$. In this example, we select the inverse Haar (Daubechies-2) DWT matrix to be the orthogonal sparsifying transform matrix Ψ because the Haar wavelet



(a) (b)

Fig. 2. (a) The size-256^2 Shepp-Logan phantom and (b) a typical star-shaped sampling domain in the frequency plane.

transform coefficients of the phantom image in Fig. 2(a) are sparse, with $\|\boldsymbol{s}\|_{\ell_0} = 3769 \approx 0.06\, m$.

Our performance metric is the peak signal-to-noise ratio (PSNR) of a signal estimate [12, eq. (3.7)]. We compare the PSNRs of (i) the DM-ECME algorithm in Section 3, with the difference map parameter set to $\beta = 1$ and maximum number of iterations set to 2000; (ii) the SPIRAL method in [8] adapted to the Gaussian signal-plus-noise model (termed SPIRAL_G), which solves the following optimization problem: $\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{y} - \Phi\, \boldsymbol{x}\|_{\ell_2}^2 + \tau \|\Psi^T\, \boldsymbol{x}\|_{\ell_1}$ subject to $\boldsymbol{x} \succeq \mathbf{0}_{m\times 1}$, with the regularization parameter

$$\tau = 10^{-4} \tag{8}$$

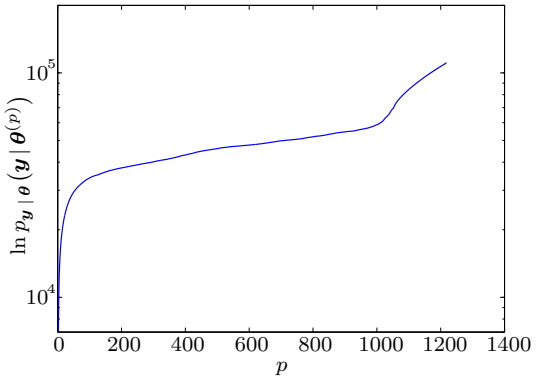
tuned manually for good performance; (iii) the ECME method in [13, Sec. II-A] that exploits signal sparsity only (termed ECME_S) and corresponds to the ECME iteration (4) with $\mathcal{P}_{\mathcal{S}_r^+}(\cdot)$ replaced by $\mathcal{P}_{\mathcal{S}_r}(\cdot)$ in the M step (4b); (iv) the normalized iterative hard thresholding (NIHT) scheme in [14]; (v) the fixed-point continuation active set convex relaxation method (labeled FPC_AS) in [15], which solves $\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{y} - \Phi\, \boldsymbol{x}\|_{\ell_2}^2 + \tau \|\Psi^T\, \boldsymbol{x}\|_{\ell_1}$, with the regularization parameter τ in (8); (vi) the standard filtered backprojection (FBP) that corresponds to setting the unobserved DFT coefficients to zero and taking the inverse DFT, see [2]. DM-ECME, ECME_S, and NIHT require knowledge of the signal sparsity level r; in this example, we set r to the true signal support size: $r = 3769$. DM-ECME, SPIRAL_G, ECME_S, and NIHT are initialized by $\boldsymbol{s}^{(0)} = \mathbf{0}_{m\times 1}$ and employ the following convergence criterion:

$$\|\boldsymbol{s}^{(p+1)} - \boldsymbol{s}^{(p)}\|_{\ell_2}^2 / m < 10^{-14}. \tag{9}$$

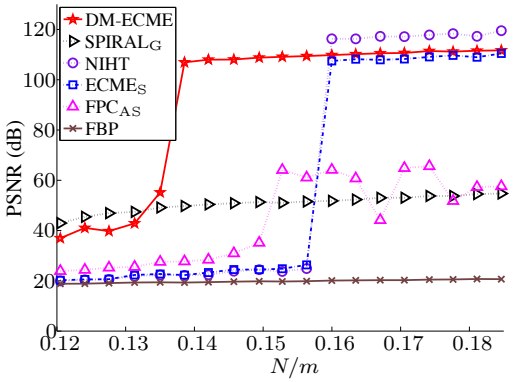
We apply (9) as the convergence criterion for the difference map iteration in step 10 of Fig. 1, where $\boldsymbol{s}^{(p+1)}$ and $\boldsymbol{s}^{(p)}$ are replaced by the difference map iterates $\boldsymbol{a}^{(q+1)}$ and $\boldsymbol{a}^{(q)}$.

We plot in Fig. 3(a) the logarithm of the marginal likelihood (3) of the DM-ECME parameter estimate at each iteration p until convergence for the reconstructions from 38 radial lines, which corresponds to the subsampling factor $N/m = 0.142$. This log likelihood increases strictly monotonically as p grows, which is appealing.

Fig. 3(b) shows the PSNRs of the above methods as we change the subsampling factor N/m by varying the number



(a)



(b)

Fig. 3. (a) The log likelihood achieved by the DM-ECME iteration as a function of the iteration index p for $N/m = 0.142$ and (b) the PSNRs achieved by various methods as functions of the subsampling factor N/m.

of radial lines in the star-shaped partial Fourier sampling pattern. In this example, the hard thresholding methods (DM-ECME, ECME$_S$, and NIHT) have significantly sharper phase transitions than the convex methods SPIRAL$_G$ and FPC$_{AS}$, and effectively achieve perfect reconstructions after the phase transitions. FPC$_{AS}$, ECME$_S$, and NIHT, which take only signal sparsity into account, exhibit the phase transitions at $N/m = 0.153$ (FPC$_{AS}$) and $N/m = 0.160$ (ECME$_S$ and NIHT), where FPC$_{AS}$ *does not* achieve perfect reconstruction after the phase transition. In contrast, the DM-ECME approach achieves an earlier phase transition at $N/m = 0.139$. Hence, in this example, exploiting the nonnegativity of the underlying image leads to 13 % saving in the number of measurements required to achieve perfect reconstruction. For $N/m < 0.153$, the SPIRAL$_G$ scheme has a consistently higher PSNR than the traditional convex FPC$_{AS}$ method by about 20 dB, which is attributed to its accounting for the signal nonnegativity. We also observe that, before its phase transition (i.e. when $N/m < 0.139$), DM-ECME achieves noticeably higher PSNRs than the traditional FBP, FPC$_{AS}$, ECME$_S$, and NIHT methods.

5. REFERENCES

[1] *IEEE Signal Processing Mag. (Special Issue on Sensing, Sampling, and Compression)*, Mar. 2008.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[3] E. Candès and J. Romberg, "Signal recovery from random projections," in *Computational Imaging III*, C. A. Bouman and E. L. Miller, Eds., San Jose, CA, Jan. 2005, vol. 5674 of *Proc. SPIE-IS&T Electron. Imaging*, pp. 76–86.

[4] D. L. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," *Proc. Nat. Acad. Sci*, vol. 102, no. 27, pp. 9446–9451, 2005.

[5] M. A. Khajehnejad, A. Dimakis, W. Xu, and B. Hassibi, "Sparse recovery of nonnegative signals with minimal expansion," *IEEE Trans. Signal Processing*, vol. 59, no. 1, pp. 196–208, 2011.

[6] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, 1988.

[7] D. L. Donoho and J. Tanner, "Precise undersampling theorems," *Proc. IEEE*, vol. 98, pp. 913–924, June 2010.

[8] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "Spiral out of convexity: Sparsity-regularized algorithms for photon-limited imaging," in *Computational Imaging VIII*, C. A. Bouman *et al.*, Eds., San Jose, CA, Jan. 2010, vol. 7533 of *Proc. SPIE-IS&T Electron. Imaging*.

[9] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, Wiley, New York, second edition, 2008.

[10] V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am. A*, vol. 20, pp. 40–55, Jan. 2003.

[11] V. Elser, I. Rankenburg, and P. Thibault, "Searching with iterated maps," *Proc. Nat. Acad. Sci*, vol. 104, pp. 418–423, Jan. 2007.

[12] J.L. Starck, F. Murtagh, and J. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*, Cambridge Univ. Press, New York, 2010.

[13] K. Qiu and A. Dogandžić, "Double overrelaxation thresholding methods for sparse signal reconstruction," in *Proc. 44th Annu. Conf. Inform. Sci. Syst.*, Princeton, NJ, 2010.

[14] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE J. Select. Areas Signal Processing*, vol. 4, pp. 298–309, Apr. 2010.

[15] Z. Wen, W. Yin, D. Goldfarb, and Y. Zhang, "A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization, and continuation," *SIAM J. Sci. Comput.*, vol. 32, no. 4, pp. 1832–1857, 2010.