A Model Based Iterative Reconstruction Algorithm For High Angle Annular Dark Field-Scanning Transmission Electron Microscope (HAADF-STEM) Tomography

S. V. Venkatakrishnan, *Student Member, IEEE*, Lawrence F. Drummy, Michael A. Jackson, Marc De Graef, Jeff Simmons, *Member, IEEE*, and Charles A. Bouman, *Fellow, IEEE*

Abstract—High angle annular dark field (HAADF)-scanning transmission electron microscope (STEM) data is increasingly being used in the physical sciences to research materials in 3D because it reduces the effects of Bragg diffraction seen in bright field TEM data. Typically, tomographic reconstructions are performed by directly applying either filtered back projection (FBP) or the simultaneous iterative reconstruction technique (SIRT) to the data. Since HAADF-STEM tomography is a limited angle tomography modality with low signal to noise ratio, these methods can result in significant artifacts in the reconstructed volume. In this paper, we develop a model based iterative reconstruction algorithm for HAADF-STEM tomography. We combine a model for image formation in HAADF-STEM tomography along with a prior model to formulate the tomographic reconstruction as a maximum *a posteriori* probability (MAP) estimation problem. Our formulation also accounts for certain missing measurements by treating them as nuisance parameters in the MAP estimation framework. We adapt the iterative coordinate descent algorithm to develop an efficient method to minimize the corresponding MAP cost function. Reconstructions of simulated as well as experimental data sets show results that are superior to FBP and SIRT reconstructions, significantly suppressing artifacts and enhancing contrast.

Index Terms—Model based iterative reconstruction (MBIR), transmission electron microscopy, electron tomography.

I. INTRODUCTION

ELECTRON tomography has been widely used in the life sciences to study biological structures in 3D [1]. In the life sciences, electron tomography is dominated by Bright

S. V. Venkatakrishnan and C. A. Bouman are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-2035 USA (e-mail: svenkata@purdue.edu; bouman@purdue.edu).

L. F. Drummy and J. Simmons are with the Materials and Manufacturing Directorate, Air Force Research Laboratory, Dayton, OH 45433 USA (e-mail: lawrence.drummy@wpafb.af.mil; jeff.simmons@wpafb.af.mil).

M. Jackson is with BlueQuartz Software, Springboro, OH 45066 USA (e-mail: mike.jackson@bluequartz.net).

M. De Graef is with the Department of Material Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: degraef@cmu.edu).
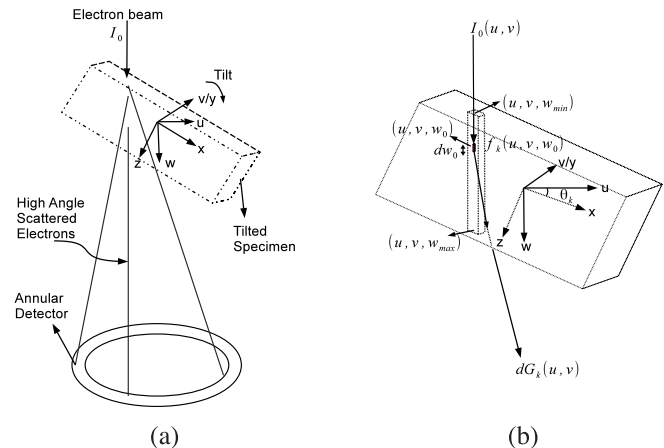
Fig. 1. Illustration of the measurement process in a HAADF-STEM acquisition system. (a) A single measurement is obtained by focusing the electron beam I_0 on the material and measuring electrons scattered into the annular detector. (b) Equivalent model for the measurement process. The scattered beam from location w_0, $dG_k(u, v)$, which is deflected by angles in the range of 50–300 mrad, is detected by the annular detector.

Field (BF) tomography which is a transmission modality. However, direct extension of the BF technique to the material sciences has proved challenging because of diffraction effects in some regions of the acquired projection images (sinogram) caused by crystalline samples [2]. This has led to a growing interest in the use of HAADF-STEM tomography to study materials in 3D (e.g. [3]–[8]). This modality, which is based on measuring electrons scattered by the material, is relatively free from diffraction effects observed in BF TEM data. The HAADF-STEM signal contains information about the atomic number of the region being imaged and hence is useful for extracting chemical information about the material. As a result HAADF-STEM tomography is also referred to as *Z*-contrast tomography [2].

A typical HAADF-STEM acquisition involves focusing an electron probe at a point on the material for a short duration and measuring the electrons scattered into an annular detector as shown in Fig. 1(a). The electron beam is raster scanned and at each point a measurement is made to obtain a projection image of the object. The object is then tilted along a single axis and the process is repeated. At the end of the acquisition, a set of projection images is obtained corresponding to each tilt

of the object. In most cases, due to mechanical constraints the object can only be tilted in the range of approximately $\pm 70°$. Therefore, HAADF-STEM tomography can be classified as a parallel beam, limited angle tomography modality. More details of HAADF-STEM acquisition and pre-processing can be found in [2] and [9].

Most approaches (e.g. [3]–[6]) for HAADF-STEM tomography have directly used the measured HAADF data for tomographic inversion using FBP or SIRT [10]. The reconstruction quality for these algorithms is limited by the fact that they do not incorporate noise statistics of the measurement, and do not incorporate a prior model. Furthermore, algorithms such as SIRT require the selection of an ad-hoc stopping criterion; otherwise the reconstructed image will typically diverge or become excessively noisy. Recently, there have been efforts on using compressed sensing (i.e. total variation prior model) based methods [7], [8] and discrete tomography [11] for HAADF-STEM tomography. However, these examples have not explicitly taken into account specific noise models and do not deal with the case when there are missing calibration parameters associated with the measurement.

Model Based Iterative Reconstruction (MBIR) algorithms have enabled significant qualitative and quantitative improvements in tomography applications like X-Ray CT [12]–[17], positron emission tomography (PET) [18]–[21], bright field electron tomography [22] , optical diffusion tomography (ODT) [23], [24], and atomic resolution tomography [25]. These methods typically require the specification of a model for the measurement system and its associated noise (i.e., a forward model); a model for the unknown image volume being reconstructed (i.e., a prior model); an estimation criteria for finding the reconstruction (typically the MAP estimate); and an algorithm to compute the estimate. However, computing the estimate requires solving a computationally demanding optimization problem, and typically an iterative approach is employed to solve it. A variety of iterative algorithms have been explored for tomographic reconstruction such as Expectation Maximization [26], Ordered Subsets (OS) [27], [28], Conjugate Gradient [29], [30], graph cuts [31] and ICD [32]. It has been shown that ICD has relatively faster convergence than other popular methods especially at higher spatial frequencies in the X-ray CT case [33] when the algorithms were initialized with a FBP reconstruction. Furthermore positivity constraints can be easily incorporated in ICD compared to other gradient based methods.

In this paper, we present a MBIR algorithm for quantitative HAADF-STEM tomography [34]. We begin by specifying a quantitative forward model for the HAADF-STEM acquisition system which depends on the 3D volume of scatter coefficients along with other parameters which can be either estimated or measured. Next, we incorporate knowledge of either sharp or diffuse interfaces between materials by using a q-Generalized Gaussian Markov Random Field (q-GGMRF) [35] prior for the scatter coefficients. If the interfaces between materials are sharp, then the parameters of the model can be adjusted to produce an edge preserving reconstruction; and if they are diffuse, then the parameters can be adjusted to produce more smooth boundaries. Finally, we formulate the reconstruction

as a maximum *a posteriori* probability (MAP) estimation problem and develop a fast multiresolution iterative algorithm to minimize the corresponding cost function.

Our algorithm to minimize the MAP cost function is based on the ICD [32] algorithm. In HAADF-STEM tomography, it is also necessary to estimate certain unknown calibration parameters, in addition to the actual 3D volume of scatter coefficients. In every iteration of the optimization, the voxels are updated, followed by the update of other unknown parameters to lower the value of the MAP cost function. To speed convergence of the algorithm we update voxels in a random order [12], use a substitute function approach and parallelize the updates. The convergence is further sped up by using a multi-resolution initialization for the parameters, which also eliminates the need to compute an analytical reconstruction as an initial condition. We evaluate our method by comparing the reconstructions with those produced by FBP and SIRT for a simulated as well as a real data set. Reconstructions of a simulated data set as well as a real data set show that MBIR produces superior reconstructions by suppressing artifacts and enhancing contrast.

The organization of this paper is as follows. In Section II we discuss the model used for the measurement process in HAADF-STEM tomography. In Section III we present details of the q-GGMRF prior model. In Section IV we use the image formation model along with the prior to formulate the MAP estimation problem and describe our algorithm to find the MAP estimate. In Section V we present results from a simulated data set and follow it up with results on a real data set. Finally in Section VI we draw our conclusions.

II. MEASUREMENT MODEL

In HAADF-STEM tomography, a single measurement is obtained by focusing an electron beam at a point on the surface of the material of interest. As the beam propagates through the material, the electrons which are scattered through high angles, approximately in the range 50–300 mrad, are detected by an annular detector (see Fig. 1). The number of electrons detected depends on the value of the HAADF scatter coefficients along the region being probed which is related to the type and number of atoms per unit volume at each location, the beam energy used as well as the inner and outer angle of the detector. The total number of electrons detected by the annular detector is typically assumed to be proportional to the projected value of the scatter coefficients through the region being probed. The electron beam is raster scanned and at each point we get a single measurement. The measurements resulting from a single raster scan of the material constitute a single electron microscope image. The sample is then tilted around a single axis (y axis in Fig. 1) and the process is repeated. Thus we get 2D parallel beam tomography data which we need to invert in order to reconstruct the HAADF scatter coefficients.

The goal of HAADF-STEM tomography is to reconstruct the HAADF scatter coefficients (units of nm^{-1}) denoted by $f(x, y, z)$ at every point in space. If (x, y, z) is the frame of reference of the object and (u, v, w) is the reference frame for the electron source (see Fig. 1), then any function of space

can be reparameterized so that $f_k(u, v, w) = f([u, v, w]R_{\theta_k})$ where R_{θ_k} is an orthonormal rotation of the spatial coordinates by an angle θ_k. For tomographic reconstruction we require measurements of the projection integral $\int_{-\infty}^{+\infty} f_k(u, v, w)dw$ through the object for every tilt θ_k and every point (u, v). We begin by describing how this measurement can be obtained from the HAADF-STEM signal. Let $I_0(u, v)$ be the source electron flux in units of electrons per nm^2 corresponding to the beam at location (u, v). Under the assumptions that attenuation effects are minimal and the sample is relatively thin, the total HAADF scatter is given by [36]

$$G_k(u, v) = I_0(u, v) \int_{w_{min}}^{w_{max}} f_k(u, v, w)\, dw. \qquad (1)$$

Thus we get

$$\begin{aligned} \frac{G_k(u, v)}{I_0(u, v)} &= \int_{w_{min}}^{w_{max}} f_k(u, v, w)dw \\ &= \int_{w_{min}}^{w_{max}} f\left([u, v, w]R_{\theta_k}\right) dw. \qquad (2) \end{aligned}$$

So the normalized quantity $G_k(u, v)/I_0(u, v)$ is an estimate of the tomographic projection at angle θ_k and at position (u, v).

We note that the above model, though widely used for HAADF-STEM tomography, does not account for the attenuation effects observed during the imaging of thick specimens of heavy elements [37]. While Van den Broek et al. have addressed some of these non-linear effects [38], the topic of developing simple tractable models that describe all the experimentally observed characteristics of the HAADF-STEM signal including contrast reversal [39] remains an open problem. We use the linear model because it is still widely applicable for a range of materials and thicknesses studied in STEM tomography including all the cases considered in this paper.

Next we model the process by which the detector converts the incident electron flux to a measured signal. The HAADF-STEM detector typically consists of a photomultiplier tube (PMT) which converts the detected electron flux to a current which is then converted to a voltage using a preamplifier and read out using an A/D converter [40]. Since the HAADF-STEM detector has a limited dynamic range, the gain associated with the PMT (referred to as contrast) and offset associated with the preamplifier [40], [41] (referred to as brightness) can be adjusted so a wide range of materials and thicknesses can be imaged. We denote the PMT gain associated with the measurements at tilt k by α_k and the additive offset associated with the measurements at tilt k by d_k. Therefore, we model the ith measurement [corresponding to the electron source at (u_i, v_i)] at tilt k by a Gaussian random variable $g_{k,i}$, with mean

$$E[g_{k,i}] = \alpha_k \int \int G_k(u, v)h_i(u, v)\, dudv + d_k \qquad (3)$$

where $h_i(u, v)$ is a kernel which averages the electron flux over the area corresponding to the ith measurement. Let f be a discretized version of $f(x, y, z)$ organized as a $N \times 1$ vector, where N is the total number of voxels, A_k is a

$M \times N$ projection matrix for tilt k, and M is the number of measurements per tilt. Then substituting (1) in (3) we get

$$\begin{aligned} E[g_{k,i}] &\approx I_k \int \int \left(\int_{w_{min}}^{w_{max}} f_k(u, v, w)dw\right) h_i(u, v)dudv + d_k \\ &= I_k \left[A_k f\right]_i + d_k \qquad (4) \end{aligned}$$

where $[A_k f]_i$ is the ith entry of the vector $A_k f$, $I_k = I_0 \alpha_k$ is the product of the source electron dose (counts) and the PMT gain at tilt k. For simplicity we will call I_k the gain associated with the measurements at tilt k. The variance of each measurement is modeled by

$$\text{Var}[g_{k,i}] = \sigma_k^2 E[g_{k,i}] \qquad (5)$$

where σ_k^2 is a parameter used to model the noise variance at tilt k and $E[g_{k,i}]$ accounts for the Poisson characteristics of the measurement. We assume that all the measurements are conditionally independent. If $E[g_{k,i}] \approx g_{k,i}$, $\Lambda_k = \text{diag}\left(\frac{1}{g_{k,1}}, \ldots, \frac{1}{g_{k,M}}\right)$, $g_k = [g_{k,1}, \ldots, g_{k,M}]^t$, $g = [g_1^t, \ldots, g_K^t]^t$, $I = [I_1, \ldots, I_K]$, $d = [d_1, \ldots, d_K]$ and $\sigma^2 = [\sigma_1^2, \ldots \sigma_K^2]$ then using (4) and (5) we get

$$\begin{aligned} p(g|f, I, d, \sigma^2) &= \left(\prod_{k=1}^{K} \frac{1}{\left(2\pi\sigma_k^2\right)^{\frac{M}{2}} |\Lambda_k|^{-\frac{1}{2}}}\right) \\ &\quad \exp\left\{-\frac{1}{2}\sum_{k=1}^{K}\frac{1}{\sigma_k^2} \|g_k - I_k A_k f - d_k 1\|_{\Lambda_k}^2\right\} \qquad (6) \end{aligned}$$

where K is the total number of tilts.

III. PRIOR MODEL

We use a q-GGMRF [35] model for the probability density of f. The parameters of this model can be adjusted to account for sharp or diffuse interfaces between materials. Moreover since it has a bounded second derivative, the substitute function method of [12] can be used, significantly speeding up the subsequent MAP estimation algorithm. If $\mathcal{N}$ is the set of all pairs of neighboring voxels (e.g. a 26 point neighborhood in 3D), w_{ij} is a weighting kernel which is inversely proportional to the distance between voxel i and voxel j, normalized so that $\sum_{j \in \mathcal{N}_i} w_{ij} = 1$, $\mathcal{N}_i$ is the set of all neighbors of voxel i, then the density function corresponding to the q-GGMRF prior is given by

$$\begin{aligned} p(f) &= \frac{1}{Z} \exp\left\{-\sum_{\{i,j\}\in\mathcal{N}} w_{ij}\rho(f_i - f_j)\right\} \\ \rho(f_i - f_j) &= \frac{\left|\frac{f_i - f_j}{\sigma_f}\right|^q}{c + \left|\frac{f_i - f_j}{\sigma_f}\right|^{q-p}} \qquad (7) \end{aligned}$$

where Z is a normalizing constant and p, q, c and σ_f are q-GGMRF parameters. Typically $1 \leq p \leq q \leq 2$ is used to ensure convexity of the function $\rho(.)$, thereby simplifying the subsequent MAP optimization. The value of σ_f is typically set to achieve a balance between resolution and noise in the reconstructions. In this paper we fix $q = 2$. When $p = 1$ the prior model corresponds to strong edge

preserving reconstructions while $p = 2$ corresponds to smooth reconstructions (bearing some similarities to the Cahn-Hilliard phase field model [42]). We note that when c is zero, the $p = 1$ case of the q-GGMRF corresponds to the total variation prior [13]. Thus the q-GGMRF provides a flexible prior model framework enabling us to model a range of possible materials from those with very sharp interfaces to those with smooth interfaces.

IV. MAP ESTIMATION AND MBIR ALGORITHM

We use the MAP estimate to reconstruct the values of the HAADF scatter coefficients. Since the gains (I), offsets (d) and variance parameters (σ^2) are typically not measured we treat them as nuisance parameters in the MAP estimation framework. The MAP estimate of the parameters is given by

$$(\hat{f}, \hat{I}, \hat{d}, \hat{\sigma}^2)$$
$$= \operatorname{argmin}_{f \geq 0, I \in \Omega, d, \sigma^2} \left\{ -\log p(g|f, I, d, \sigma^2) - \log p(f) \right\}$$

where $\Omega = \left\{ I \in \mathbb{R}^K : \frac{1}{K} \sum_{k=1}^K I_k = \bar{I} \right\}$, $p(I, d, \sigma^2)$ is uniformly distributed, f is conditionally independent of (I, d, σ^2). We impose positivity constraints on the voxels ($f \geq 0$) as it is physically meaningful to have positive values of the HAADF scatter coefficients. The constraint Ω forces the average value of the gains to be equal to a value $\bar{I}$ to prevent our algorithm from diverging to unreasonable values of the HAADF scatter coefficients. The choice of $\bar{I}$ is arbitrary but affects the scaling of f and hence the choice of the prior model parameter, σ_f. Hence if $\bar{I}$ is set to the product of detector gain and source electron dose, then the reconstructions will be quantitative.

Using (6) and (7) we obtain the MAP estimate by minimizing the cost

$$\begin{aligned} \mathrm{c}(f, I, d, \sigma^2) &= \frac{1}{2} \sum_{k=1}^K \frac{1}{\sigma_k^2} \|g_k - I_k A_k f - d_k 1\|_{\Lambda_k}^2 \\ &+ \frac{1}{2} \sum_{k=1}^K \log \left(\left(2\pi \sigma_k^2 \right)^M |\Lambda_k|^{-1} \right) \\ &+ \sum_{\{i,j\} \in \mathcal{N}} w_{ij} \rho(f_i - f_j). \end{aligned} \quad (8)$$

In general the cost function $\mathrm{c}(f, I, d, \sigma^2)$ is convex in f but not jointly convex in (f, I, d, σ^2). We adapt the ICD algorithm [14] to minimize the cost function (8). In ICD the parameters are updated one at a time such that each update results in a lower value of the cost function. The basic structure of our algorithm is to repeatedly perform the following steps until convergence is achieved.

(i) For each voxel j, $\hat{f}_j \leftarrow \operatorname*{argmin}_{f_j \geq 0} \mathrm{c}(f, I, d, \sigma^2)$.

(ii) $(\hat{I}, \hat{d}) \leftarrow \operatorname*{argmin}_{I \in \Omega, d} \mathrm{c}(\hat{f}, I, d, \sigma^2)$.

(iii) $\hat{\sigma}^2 \leftarrow \operatorname*{argmin}_{\sigma^2} \mathrm{c}(\hat{f}, \hat{I}, \hat{d}, \sigma^2)$.

In Step (i) we update each voxel to lower the original cost function using a substitute function approach [12]. The form of the substitute function is chosen so that it results in simple closed form updates for the voxels, speeding up the implementation of our method. It has also been shown in [12] that the substitute function approach speeds up the overall convergence of the algorithm to the minimum. Minimizing the substitute function does not minimize the original cost but it results in voxel updates that lower the original cost. In Step (ii) we find the minimum of the cost function with respect to the gain I and offset parameters d by turning the constrained optimization to an unconstrained problem by using a Lagrange multiplier. Finally in (iii) we minimize the cost function with respect to σ^2. Thus each of the above updates lowers the value of the original cost function. The algorithm is terminated if the relative change in the total magnitude of the reconstruction is less than a preset threshold. Next, we derive the update equations for the three steps of our algorithm.

A. Voxel Update

We adopt a strategy similar to [12] based on a substitute function approach and random order update of the voxels to minimize the cost with respect to each voxel. Sets of voxels having the same (x, z) coordinate constitute a voxel line. Voxels in the same voxel line share geometry computation and therefore it is computationally efficient to update a single line of voxels together. Updating the voxel lines in random order and using a substitute function approach has also been shown to speed up the convergence of the ICD algorithm [12]. Further the voxel updates are implemented in parallel (on multicore machines) by allocating sets of contiguous $x - z$ slices to different cores similar to [43].

In order to find the optimal update for the jth voxel we begin by rewriting the cost function (8) by ignoring terms not involving voxel j as

$$\tilde{c}(u) = \tilde{\theta}_1 u + \frac{\tilde{\theta}_2}{2} \left(u - \tilde{f}_j \right)^2 + \sum_{i \in \mathcal{N}_j} w_{ji} \rho(u - f_i) \quad (9)$$

where $\tilde{\Lambda}_k = \frac{1}{\sigma_k^2} \Lambda_k$, $\tilde{\theta}_1 = -\sum_{k=1}^K e_k^t \tilde{\Lambda}_k (I_k A_{k,*,j})$, $\tilde{\theta}_2 = \sum_{k=1}^K (I_k A_{k,*,j})^t \tilde{\Lambda}_k (I_k A_{k,*,j})$, $A_{k,*,j}$ is the jth column of the forward projection matrix A_k at tilt k, $e_k = g_k - I_k A_k f - d_k 1$, $\tilde{f}_j$ is the present value of voxel j, and $\mathcal{N}_j$ is the set of all neighbors of voxel j. We must minimize this function with respect to u to find the optimal update for voxel j. Taking the derivative with respect to u and setting it to zero does not give a closed form update for u due to the complicated form of the potential function $\rho(\cdot)$. Therefore we use a substitute function approach in which we find a function which bounds (9) from above, such that minimizing the new function results in a lower value of the original cost function. Typically the substitute function is chosen so that it can be minimized in a computationally efficient manner. We find a substitute function for each $\rho(u - f_i)$ of the form

$$\rho(u - f_i; \tilde{f}_j - f_i) = \frac{a_{ji}}{2}(u - f_i)^2 + b_{ji}. \quad (10)$$

Using such a form results in a simple closed form update for a given voxel. The values of a_{ji} and b_{ji} can be derived as

```
UpdateVoxel(j,e) {
```
$$\tilde{\theta}_2 \leftarrow \sum_{k=1}^{K}(I_k A_{k,*,j})^t \tilde{\Lambda}_k (I_k A_{k,*,j})$$
$$\tilde{\theta}_1 \leftarrow -\sum_{k=1}^{K} e_k^t \tilde{\Lambda}_k (I_k A_{k,*,j})$$

for each $i \in \mathcal{N}_j$ { Compute substitute function parameter a_{ji} using (11)}

$$u^* \leftarrow \frac{\sum_{i \in \mathcal{N}_j} w_{ji} a_{ji} f_{i+\tilde{\theta}_2 \tilde{f}_j - \tilde{\theta}_1}}{\sum_{i \in \mathcal{N}_j} w_{ji} a_{ji} + \tilde{\theta}_2}$$

$$f_j \leftarrow \max(u^*, 0)$$
$$e_k \leftarrow e_k - (f_j - \tilde{f}_j) I_k A_{k,*,j}, k = 1, 2, \cdots K$$
```
}
```

Fig. 2. Pseudocode for updating a single voxel j. The parameters for the 1D optimization problem are computed based on a substitute function approach and a new value for the voxel is evaluated. The error sinogram e is then updated based on this new voxel value.

shown in Appendix I and are given by

$$a_{ji} = \begin{cases} \frac{\rho'(\tilde{f}_j - f_i)}{(\tilde{f}_j - f_i)} & \tilde{f}_j \neq f_i \\ \rho''(0) & \tilde{f}_j = f_i \end{cases} \quad (11)$$

$$b_{ji} = \rho(\tilde{f}_j - f_i) - \frac{a_{ji}}{2}(\tilde{f}_j - f_i)^2. \quad (12)$$

Given these values, the new cost function we need to minimize is

$$\tilde{c}_{\text{sub}}(u) = \tilde{\theta}_1 u + \frac{\tilde{\theta}_2}{2}\left(u - \tilde{f}_j\right)^2 + \sum_{i \in \mathcal{N}_j} w_{ji} \rho(u - f_i; \tilde{f}_j - f_i).$$

Since $\rho(u - f_i; \tilde{f}_j - f_i)$ is quadratic in u, the minimum of $\tilde{c}_{\text{sub}}(u)$ has a closed form and is given by

$$u^* = \frac{\sum_{i \in \mathcal{N}_j} w_{ji} a_{ji} f_i + \tilde{\theta}_2 \tilde{f}_j - \tilde{\theta}_1}{\sum_{i \in \mathcal{N}_j} w_{ji} a_{ji} + \tilde{\theta}_2}. \quad (13)$$

Enforcing the positivity constraint, the final update for voxel j is given by

$$\hat{f}_j \quad \leftarrow \quad \max(u^*, 0). \quad (14)$$

The pseudocode for updating a single voxel is given in Fig. 2.

B. Gain and Offset Update

To minimize the function (8) with respect to the gains and offsets subject to the constraint, we turn the constrained optimization problem into an unconstrained one by using a Lagrange multiplier λ. If

$$Q_k = \begin{bmatrix} (A_k \hat{f})^t \tilde{\Lambda}_k (A_k \hat{f}) & (A_k \hat{f})^t \tilde{\Lambda}_k 1 \\ (A_k \hat{f})^t \tilde{\Lambda}_k 1 & 1^t \tilde{\Lambda}_k 1 \end{bmatrix},$$

$$b_k = \begin{bmatrix} b_{k,1} \\ b_{k,2} \end{bmatrix} = \begin{bmatrix} g_k^t \tilde{\Lambda}_k A_k \hat{f} \\ g_k^t \tilde{\Lambda}_k 1 \end{bmatrix}, \quad \tilde{\Lambda}_k = \frac{1}{\sigma_k^2}\Lambda_k,$$

$$I = [I_1, \ldots, I_K], \quad Q_k^{-1} = \begin{bmatrix} \tilde{q}_{k,11} & \tilde{q}_{k,12} \\ \tilde{q}_{k,21} & \tilde{q}_{k,22} \end{bmatrix},$$

```
UpdateGainOffset(f, e) {
    for each tilt k {
```
$$A_k f \leftarrow \frac{g_k - d_k 1 - e_k}{I_k}$$

Compute Q_k, b_k
```
    }
```
Compute $\hat{\lambda}$ using (15)

for each tilt k {

Compute $\hat{I}_k$ and $\hat{d}_k$ using (16)
$$e_k \leftarrow g_k - \hat{I}_k(A_k f) - \hat{d}_k 1$$
```
    }
}
```

Fig. 3. Pseudocode for updating the gain and offset parameters. The parameters for the constrained optimization problem are computed efficiently using the present error sinogram and used to evaluate the optimal gains and offsets. Based on these new values, the error sinogram is updated.

the optimal update can be derived as shown in Appendix II and is given by

$$\hat{\lambda} = \frac{\sum_{k=1}^{K}(\tilde{q}_{k,11} b_{k,1} + \tilde{q}_{k,12} b_{k,2}) - K\bar{I}}{\frac{1}{K}\sum_{k=1}^{K} \tilde{q}_{k,11}} \quad (15)$$

and

$$\begin{bmatrix} \hat{I}_k \\ \hat{d}_k \end{bmatrix} = Q_k^{-1}\left(b_k - \frac{1}{K}\begin{bmatrix} \hat{\lambda} \\ 0 \end{bmatrix}\right). \quad (16)$$

The pseudocode for updating the gains and offsets efficiently is given in Fig. 3.

C. Variance Parameter Update

To find the optimal value of the variance parameter, we can take the gradient of (8) with respect to σ^2 and set it to zero. This gives us the optimal update for each σ_k^2 as

$$\hat{\sigma}_k^2 \leftarrow \frac{\|e_k\|_{\Lambda_k}^2}{M} \quad (17)$$

where $e_k = g_k - \hat{I}_k A_k \hat{f} - \hat{d}_k 1$ and M is the number of measurements for each tilt.

D. Multiresolution Initialization

In model based X-Ray CT it is common to initialize the reconstruction using FBP. In HAADF-STEM tomography initializing with FBP is challenging because there are several nuisance parameters for which we have no initial estimates. Moreover since the data sets are large and the optimization tend to be computationally intensive, the overall convergence can be sped up by using a multiresolution algorithm to initialize the parameters. Multiresolution methods have proved to be useful in initializing parameters in other tomographic modalities [19], [24], [45]. These methods transfer the computational load to the coarser scales where the optimization algorithm is faster because the dimensionality

```
UpdateParameters(f, I, d, σ², N_InnerIter, T){
    Λ̃_k ← (1/σ²_k)Λ_k, k = 1,···, K
    OuterIter ← 0  //holds the number of iterations
    do {
        e_k ← g_k − I_k(A_k f) − d_k 𝟙, k = 1,···, K
        for Iter=1 to N_InnerIter  //inner loop
            for each voxel j in random order
                (f̂_j, e) ← UpdateVoxel(j, e)
        (Î, d̂, e) ← UpdateGainOffset(f̂, e)
        for each tilt k
            σ̂²_k ← ‖e_k‖²_Λ_k / M
            Λ̃_k ← (1/σ̂²_k)Λ_k
        if (OuterIter == 0)
            N_InnerIter ← 1
        OuterIter ← OuterIter+1
    } while (Stopping Threshold ≥ T or OuterIter ≤ 1)
}
```

Fig. 4. Pseudocode to update all the parameters and find the optimal value of the voxels. First, all the voxels are updated in random order, followed by the gains, offsets and noise variance. The value $N_{\text{InnerIter}}$ determines the number of inner iterations over the voxels. At the very first outer iteration at the coarsest resolution this parameter is set to a fixed value greater than 1. In all subsequent iterations this value is set to 1. The algorithm is terminated if the relative change in the magnitude of the reconstruction is less than a threshold T and the number of outer iterations is greater than one. Keeping track of the error sinogram (e_k) at each tilt k results in a computationally efficient algorithm.

```
MultiResolutionReconstruction(g, σ_f, Ī, N_Res, N_Inner, T){
    Set d^(0) according to (18) and (19)
    I^(0) = Ī𝟙, σ^(2(0)) = 𝟙, f^(0) = 0
    for k = 1 to N_Res
    if(k == 1)
        (f^(k), I^(k), d^(k), σ^(2(k))) ←
    UpdateParameters(f^(k−1), I^(k−1), d^(k−1), σ^(2(k−1)), N_Inner, T)
    else
        f^(k−1) ← UpSample(f^(k−1))
        σ_f^(k−1) ← Update σ_f using Eq. 28 in [44]
        (f^(k), I^(k), d^(k), σ^(2(k))) ←
    UpdateParameters(f^(k−1), I^(k−1), d^(k−1), σ^(2(k−1)), 1, T)
}
```

Fig. 5. Pseudocode to perform multiresolution initialized reconstruction. N_{Res} is the number of resolutions and N_{Inner} is the number of inner iterations at the coarsest resolution. The superscript attached to each unknown parameter indicates the resolution index. A superscript of 0 corresponds to the coarsest resolution with higher values corresponding to finer resolutions. At the coarsest resolution the parameters are initialized as shown. At subsequent resolutions the algorithm is initialized using the output of the previous resolution. The function Upsample produces a finer resolution object from a coarser resolution object by using a suitable up-sampling technique.
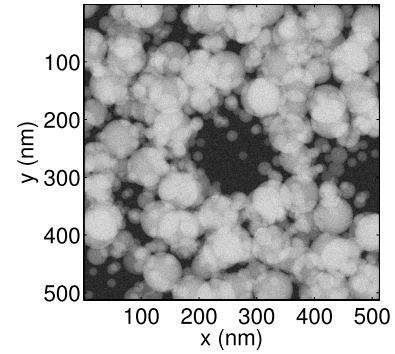

Fig. 6. Simulated HAADF microscope image of aluminum spheres in vacuum when the object is not tilted.

of the problem is effectively reduced. In our application we set the parameter c in the prior model to be small, and $q = 2$ so our prior behaves similar to a GGMRF [46]. Therefore we adapt the scaling parameter of the prior model, σ_f, for different resolutions using Eq. 28 in [44]. Next we detail a method to initialize the parameters at the coarsest resolution.

1) Parameter Initialization at Coarsest Scale: In HAADF-STEM tomography, we observe that the signal measured at each view is offset by an additive constant due to the detector. In order to estimate these offsets in the optimization framework we require a reasonable initial estimate. To do so we assume that the material is homogenous. Thus the measured signal at each view k will be proportional to $\frac{1}{\cos(\theta_k)}$ plus an unknown offset. At the coarsest resolution the offset parameter vector d is initialized by a least squares fit of the average count in each view k to $\frac{1}{\cos(\theta_k)}$. More specifically, if G is a $K \times 1$ vector containing the average counts of the data at each tilt, and

$$D = \begin{bmatrix} \frac{1}{\cos(\theta_1)} & \cdots & \frac{1}{\cos(\theta_K)} \\ 1 & \cdots & 1 \end{bmatrix}^t$$

then the least squares estimate $\hat{\phi} = [\hat{\phi}_1 \ \hat{\phi}_2]^t$ is given by

$$\hat{\phi} = (D^t D)^{-1}(D^t G) \tag{18}$$

and initial value of the offset is set to

$$d = \hat{\phi}_2 \mathbb{1}. \tag{19}$$

The initial value of the voxels (f) is set to 0. Since the source electron dose and detector gains are typically unknown, the gains I, are set to $\bar{I}\mathbb{1}$. Finally the noise variance parameter σ^2 is set to $\mathbb{1}$.

2) Iterations Over Voxels at the Coarsest Scale: At the very first iteration at the coarsest scale, given the initial values of the parameters, we update all the voxels iteratively multiple times ($N_{\text{InnerIter}}$ in Fig. 4) so that we do not converge to a local minimum with unreasonable values for the HAADF scatter coefficients. Beyond this stage the parameters are all updated once per iteration.

The algorithm can be efficiently implemented by keeping track of the error vector in the course of the iterations as shown in Fig. 4. We note that the stopping criteria for our method is that the relative change in the magnitude of the reconstruction is less than a preset threshold and the number of iterations is greater than one. The pseudocode for the full multiresolution initialization based reconstruction is shown in Fig. 5.
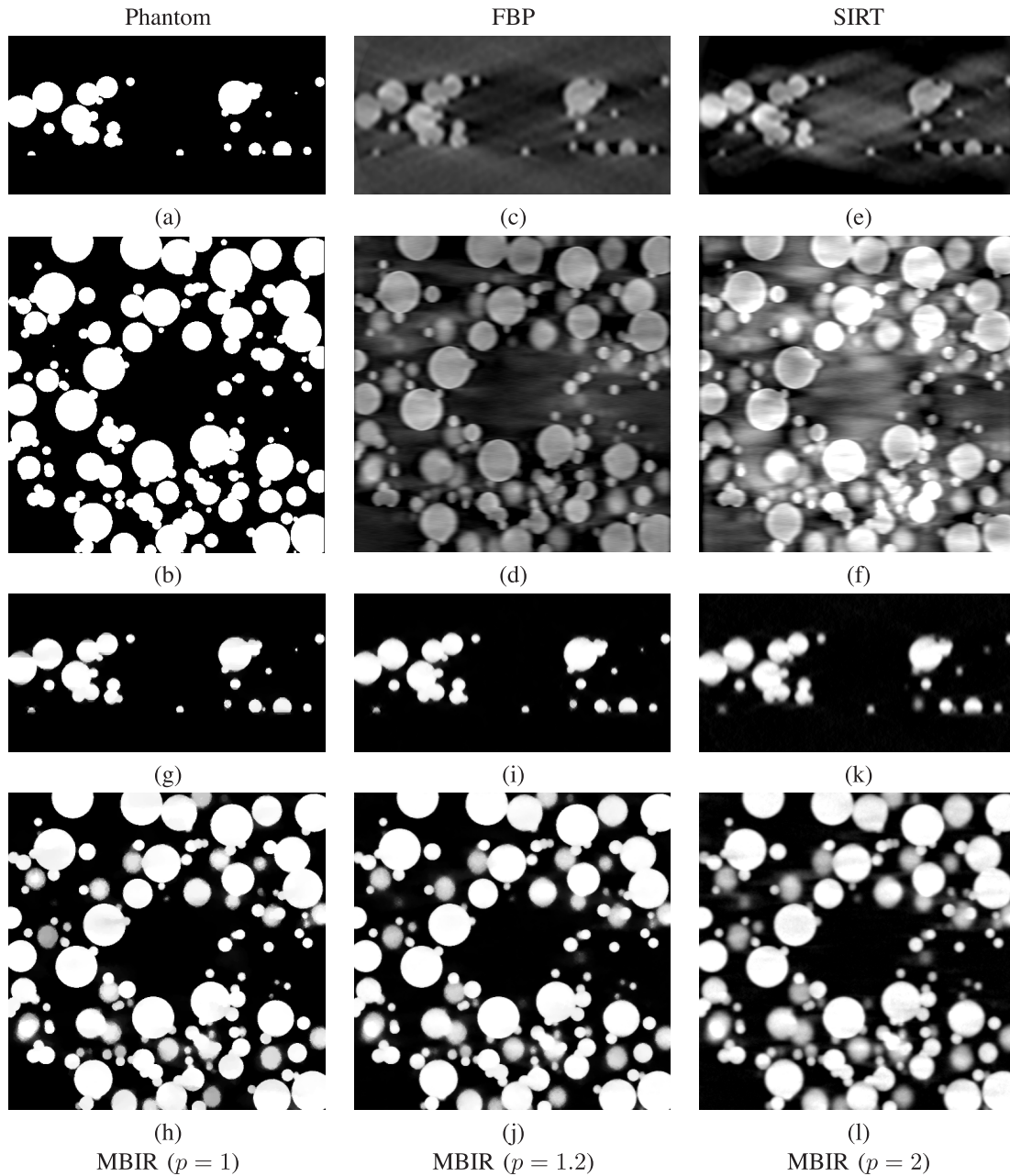
Fig. 7.   Reconstructions using FBP, SIRT, and MBIR with different prior models. (a), (c), (e), (g), (i), and (k) show a $x - z$ slice and (b), (d), (f), (h), (j), and (l) show a $x - y$ slice. (a) and (b) show ground truth corresponding to a single slice. (c) and (d) show reconstructions from FBP, (e) and (f) show SIRT reconstructions using IMOD. (g)–(l) show MBIR reconstructions with different values of the prior model parameter p. FBP and SIRT reconstructions are noisy and have streaking artifacts in the $x - z$ plane while MBIR significantly suppresses these artifacts and can produce sharper reconstructions. Increasing the value of the prior model parameter p from 1 to 2 produces smoother MBIR reconstructions. Thus the value of p can be chosen to best match the type of material being imaged.

V. EXPERIMENTAL RESULTS

A. Simulated Data Set

We begin by studying the performance of the MBIR algorithm on a simulated data set produced from a phantom consisting of spheres in vacuum. Our objective is to compare the results of our method to the most widely used algorithms for HAADF-STEM tomography, FBP and SIRT. The FBP and SIRT reconstructions are performed using a popular electron microscopy package, IMOD [47]. Since the scaling is arbitrary and no positivity is enforced in IMOD [47], we clip the reconstructions to be positive and perform a least squares

fit to scale the reconstructions to a similar range as the true phantom. We use the visual quality as well as the root mean square error (RMSE) between the reconstruction and the original phantom to evaluate the reconstructions.

Fig. 6 show the simulated HAADF data corresponding to the phantom at zero tilt. The images have been displayed by scaling them using the minimum and maximum count in the data. The spheres in the phantom have a scattering coefficient of 4.132×10^{-4} nm^{-1} corresponding to Aluminum at 300 kV with detection angles 50–250 mrad (value obtained using the Monte-Carlo simulator, CASINO [48]). The sphere diameters vary (up to 100 nm) and the sample thickness is about 128 nm.
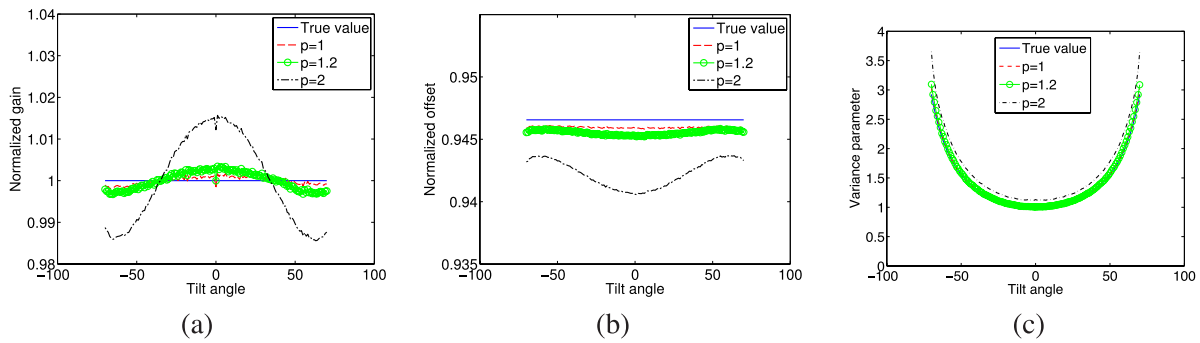
Fig. 8. Final value of the gains (a), offset (b), and variance parameters (c) upon termination of the algorithm. The gains are normalized by the dosage used and the offset are normalized to the mean value of the counts in the data set. We observe that the MBIR algorithm can estimate these unknown parameters and hence they need not be explicitly measured.
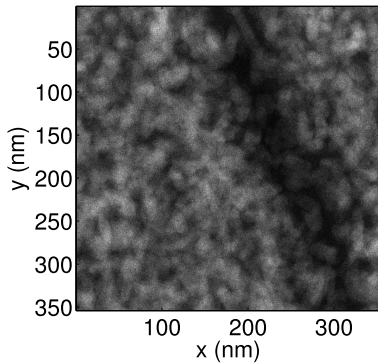

Fig. 9. Acquired HAADF data for a Titanium dioxide nanoparticle data set at zero tilt. The dark regions represent a void in the material.

TABLE I
COMPARISON OF THE ROOT MEAN SQUARE ERROR OF THE RECONSTRUCTION WITH RESPECT TO THE ORIGINAL PHANTOM FOR VARIOUS SCENARIOS. SIRT AND FBP HAVE HIGHER RMSE THAN MBIR INDICATING THAT MBIR CAN PRODUCE QUANTITATIVELY ACCURATE RECONSTRUCTIONS

Method	RMSE (nm^{-1})
FBP	1.278×10^{-4}
SIRT	0.988×10^{-4}
MBIR ($p = 1$)	0.284×10^{-4}
MBIR ($p = 1.2$)	0.324×10^{-4}
MBIR ($p = 2$)	0.496×10^{-4}

The object is tilted from 70° to –70° in steps of 1° and the projection images are obtained using (1) with an electron flux of 50 000 counts per nm^2 for every measurement and the detector gain set to 1. The offset d_k is set to 9000 counts for each view. To each HAADF projection measurement Gaussian noise is added with variance parameter σ^2 set so that the noise variance increases with tilt. The values of σ_k^2 corresponds to a minimum SNR $\left(\min_{k,i} 10 \log \left(\frac{g_{k,i}}{\sigma_k^2} \right) \right)$ of 34.471 dB. The projection images are acquired at a pixel size of 1 nm × 1 nm. All the reconstructions are performed with $\sigma_f = 4.1 \times 10^{-5}$ nm^{-1}, $q = 2$, and $c = 0.01$ using a 3 stage multiresolution initialization. At the coarsest resolution the initial value of the image is set to $f = 0$, the variance parameter is set to 1 and the number of inner iterations, $N_{\text{InnerIter}}$ (see Fig. 5), is set to 10. The interpolation between resolutions is performed using pixel replication. The value of $\bar{I}$ is set to 50 000, the value of the source electron dose - detector gain product, so that the reconstructions are quantitative. The stopping threshold is set to 0.1%.

Fig. 7 (a) and (b) show a single $x - z$ and $x - y$ slice from the original phantom. Fig. 7 (c) and (d) shows the corresponding slices from the reconstruction obtained using the FBP algorithm implemented in IMOD [47]. The algorithm results in blurry reconstructions with significant streaking artifacts in the $x - z$ plane. In the $x - y$ plane the reconstructions are noisy. Fig. 7 (e) and (f) show the reconstructions using the SIRT algorithm. While some of the noise appears suppressed

compared to FBP we still observe the streaking artifacts in the $x - y$ plane and noise in the $x - y$ plane. Moreover in regions where the phantom has no material, we observe that the reconstructions still contain material with non zero scatter coefficients. The RMSE of the SIRT reconstructions are lower than those of the FBP reconstructions as shown in Table I.

Fig. 7 (g)–(l) shows the reconstructions when we apply MBIR algorithm to the simulated data set with different values of the q-GGMRF parameter p. This phantom has discontinuous boundaries so as we would expect the total-variation prior [$p = 1$ in Fig. 7 (g) and (h)] is well matched to its behavior and produces the lowest RMSE (see Table I) reconstruction. Fig. 7 (i) and (j) shows the MBIR reconstruction when we set $p = 1.2$. This produces results with slightly more smooth edges than the $p = 1$ case. This prior been found to be useful [35] since it provides a good balance between preserving edges and modeling the smooth regions in the reconstruction. Finally Fig. 7 (k) and (l) show the MBIR reconstruction when $p = 2$. In this case the edges are most diffuse and hence the RMSE is higher than the $p = 1$ and $p = 1.2$ case for this phantom. In all cases the streaking artifacts in the $x - z$ plane are significantly suppressed and noise in the $x - y$ slices is effectively reduced compared to FBP and SIRT. The RMSE of the MBIR reconstructions (see Table I) are lower than those of the SIRT and FBP suggesting that the MBIR reconstructions are quantitatively more accurate than FBP and SIRT. Thus this experiment illustrates that MBIR is superior to FBP and SIRT and furthermore the parameters of the algorithm can be chosen to model a range of interfaces from very sharp to diffuse.
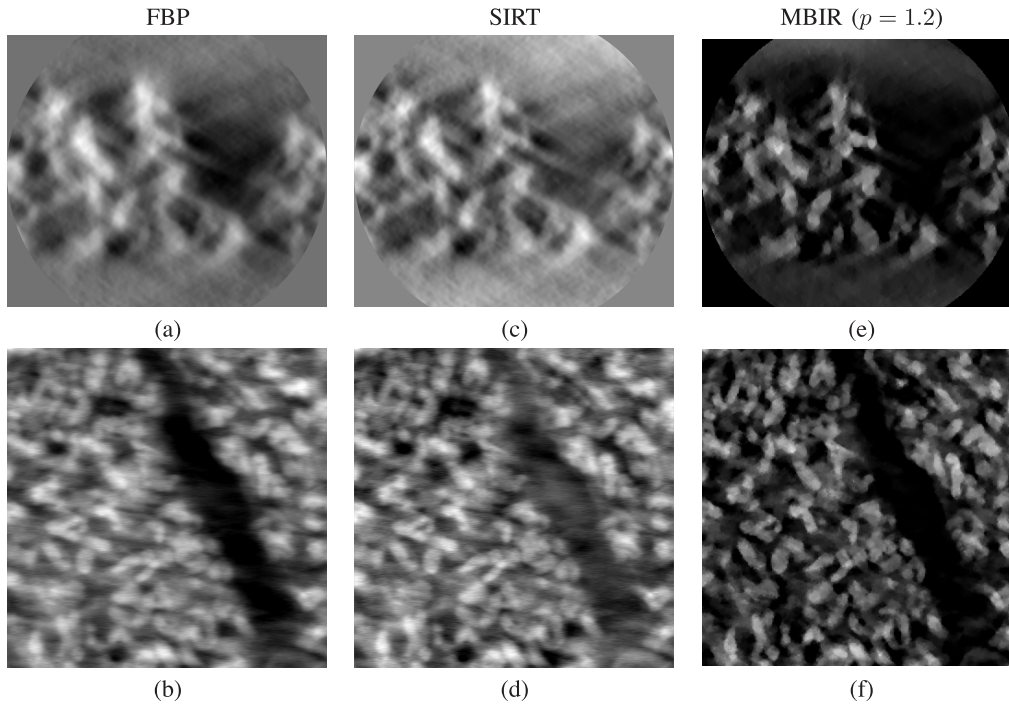
FBP　　　　　　　　　SIRT　　　　　　　　MBIR ($p = 1.2$)

(a)　　　　　　　　　　(c)　　　　　　　　　　(e)

(b)　　　　　　　　　　(d)　　　　　　　　　　(f)

Fig. 10. Comparison of MBIR with FBP and SIRT on a real HAADF-STEM data set. The top row shows a $x - z$ reconstructed slice and the bottom shows a $x - y$ reconstructed slice. (a) and (b) FBP reconstruction, (c) and (d) SIRT with 20 iterations, (e) and (f) MBIR with $p = 1.2$. MBIR produces images with no streaks in the $x - z$ plane and significantly suppresses noise in the $x - y$ plane.

MBIR ($p = 1$)　　　　　　　　MBIR ($p = 2$)
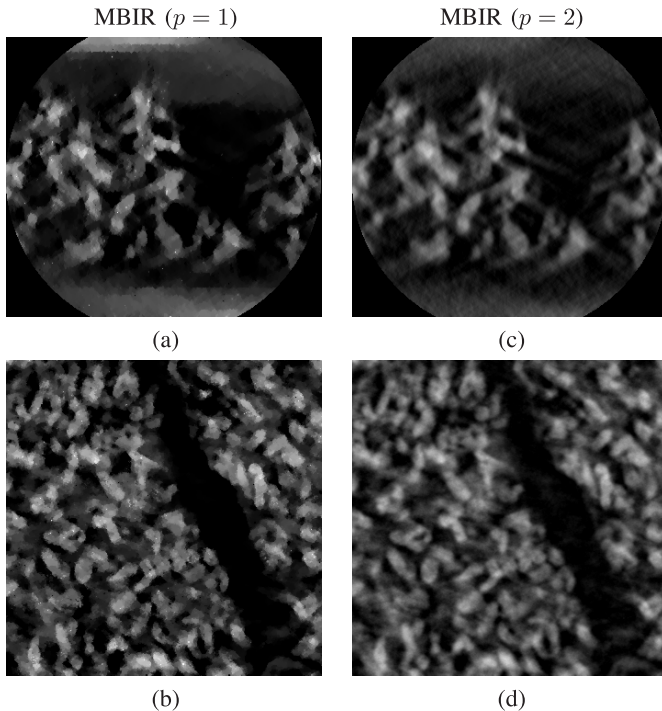

(a)　　　　　　　　　　　　　(c)

(b)　　　　　　　　　　　　　(d)

Fig. 11. Effect of varying q-GGMRF shape parameter p on the reconstruction. The top row shows a $x - z$ reconstructed slice and the bottom row shows a $x - y$ reconstructed slice. (a) and (b) $p = 1$, (c) and (d) $p = 2$. As the value of p increases the edges in the reconstructions become smoother.

Fig. 8 (a)–(c) show the final estimated values of the normalized gain, $\frac{I_k}{I}$, normalized offset, $\frac{d_k}{\bar{g}}$, (where $\bar{g}$ is the average count in the data set) and variance parameters, σ_k^2. We observe that the parameters are accurately estimated demonstrating

that the MBIR method can accurately account for missing calibration parameters without explicitly measuring them.

B. Experimental Data Set

In order to evaluate our approach on real data, we compare our algorithm with FBP and SIRT from IMOD [47]. The data acquired is of a $\approx$150 nm thick sample of polystyrene functionalized Titanium dioxide nano particle assembly [49]. The TEM used was a FEI Titan operating in STEM mode with 300 kV accelerating voltage, spot size 7, $\pm70°$ with $2°$ increments for $\pm54°$ and $1°$ increments for $54°$ to $70°$ and $-54°$ to $-70°$. The exposure time was 12.6 seconds, magnification was set to 225 kX, the frame size set to 2048×2048, with 0.34 nm pixel size, and STEM dynamic focus activated. A Fischione model 3000 HAADF photomultiplier tube detector was used at camera length of 130 mm. We use a $\approx$350 nm $\times$ 350 nm section of the projection images for reconstruction. Fig. 9 shows a single projection image acquired when the object is at zero tilt, displayed by scaling it to the range of the data.

The FBP and SIRT reconstructions are performed with voxels of size 0.343nm $\times$ 0.343nm $\times$ 0.343nm. The filter parameters for FBP are chosen to produce the most visually appealing results. The particles of interest in this data set are approximately cylindrical with a diameter of 18 nm and a height of 40 nm [49]. Thus in order to reduce computation, MBIR is performed with voxels of size (3×0.343) nm $\times (3 \times 0.343)$ nm $\times (3 \times 0.343)$ nm. The parameter σ_f is chosen for the best visual quality of reconstruction. The value of c is set to 0.01. We use a three stage multiresolution initialization for the reconstruction. The interpolation between resolutions is performed using pixel replication. At the coarsest resolution
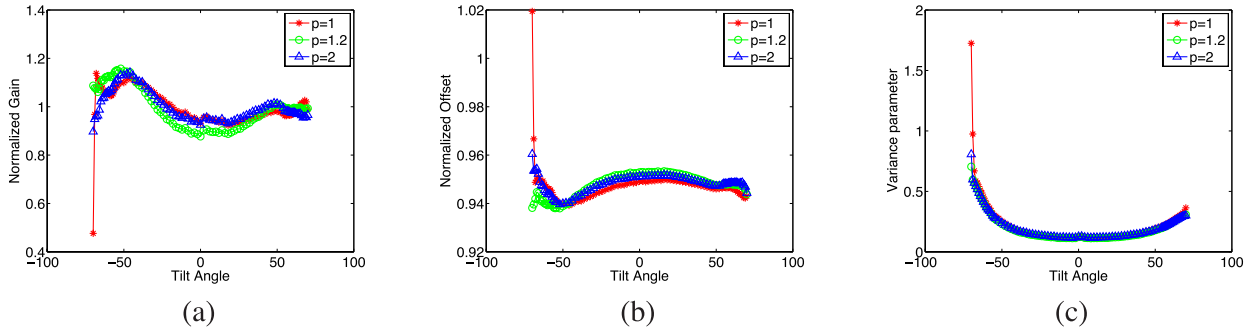
Fig. 12. Final estimates for the normalized gain $- \frac{I_k}{\bar{I}}$ (a), normalized offset $- \frac{d_k}{g}$ (b) and variance parameter $- \sigma_k^2$ (c) with respect to tilt θ_k for $p = 1$, $p = 1.2$ and $p = 2$.

the number of inner iterations ($N_{\text{InnerIter}}$ in Fig. 5) is set to 10. Since the source electron dose and detector gains are unknown in this case, we set $\bar{I} = 20\,000$ and jointly estimate the gains along with the offsets and variance parameters as a part of the reconstruction. The stopping threshold is set to 0.9%. The dimensions of the reconstructed volume are set so as to account for all the voxels contributing to the projection data. In presenting the results we only show voxels that can be reliably reconstructed from the projection data i.e. at every tilt there is a projection measurement corresponding to those voxels. Additionally while displaying the results, we use a scaling window ranging from the minimum to the maximum value in the reconstructed volume.

Fig. 10 shows a single $x - z$ and $x - y$ slice from FBP, SIRT and the MBIR reconstruction. We observe that in SIRT and FBP there are streaking artifacts in the $x - z$ plane of reconstruction while MBIR significantly suppresses these artifacts. In the $x - y$ plane the effects of noise are effectively suppressed in MBIR clearly showing the Titanium dioxide nano particles against the background support material. This demonstrates the effectiveness of the method even for this particularly limited tilt data set.

Finally we study the effect of varying the parameter p of the q-GGMRF prior on the quality of reconstruction. Fig. 11 shows a reconstructed $x - z$ and $x - y$ slice for $p = 1$ and $p = 2$. Values of p close to 1 represent strong edge preserving reconstructions values close to 2 result in smoother reconstructions. This is an additional flexibility which our model based approach offers compared to SIRT and FBP i.e. the ability to incorporate our prior knowledge of the nature of the interfaces of the material. Knowledge of smooth interfaces can be incorporated by setting $p = 2$ and sharp interfaces can be represented by the setting p close to 1. Fig. 12 shows the final values of the gains, offsets and variance parameters upon termination of the algorithm for the case when $p = 1$, $p = 1.2$ and $p = 2$. However, we do not have knowledge of the true value of the parameters to compare with those estimated by our algorithm.

VI. CONCLUSION

In this article, we presented a model based iterative reconstruction algorithm for HAADF-STEM tomography. We used a model for the measurement process and combined it with a

prior model of the 3D volume to formulate the reconstruction as a MAP estimation problem. We presented a fast algorithm for quantitative reconstruction of the local HAADF scatter rate per unit distance which also accounted for unknown calibration parameters. Furthermore, the algorithm is flexible to incorporate knowledge of the type of interfaces present in the material via the prior model. We showed that our algorithm produces superior reconstructions compared to FBP and SIRT on a simulated and a real data set.

APPENDIX I
SINGLE VOXEL UPDATE USING SUBSTITUTE FUNCTION

In order to find a suitable substitute function to the original cost (9), each of the potential functions $\rho(u - f_i)$ can be replaced by a function $\rho(u - f_i; f_j' - f_i)$ which satisfy the following properties [12]

$$\rho(u - f_i; f_j' - f_i) \geq \rho(u - f_i) \quad \forall u \in \mathbb{R} \quad (20)$$
$$\rho'(f_j' - f_i; f_j' - f_i) = \rho'(f_j' - f_i) \quad (21)$$

where f_j' is the point of approximation. Intuitively (20) ensures that the substitute function upper bounds the original potential function and (21) ensures that the derivatives of the original function and the substitute function are matched at the point of approximation. We use a substitute function of the form

$$\rho(u - f_i; f_j' - f_i) = \frac{a_{ji}}{2}(u - f_i)^2 + b_{ji} \quad (22)$$

because it results in a simple closed form update for a given voxel. Thus we need to find the values of a_{ji} and b_{ji} which satisfies (20) and (21). Taking the derivative of the substitute function (22) and matching it to the derivative of the original potential function we get

$$a_{ji} = \begin{cases} \dfrac{\rho'(f_j' - f_i)}{(f_j' - f_i)} & f_j' \neq f_i \\ \rho''(0) & f_j' = f_i. \end{cases}$$

To choose b_{ji} we set the value of the original potential function and substitute function to be the same at the point of approximation f_j'. This gives

$$b_{ji} = \rho(f_j' - f_i) - \frac{a_{ji}}{2}(f_j' - f_i)^2$$

APPENDIX II
GAIN AND OFFSET UPDATE WITH UNKNOWN TRANSMITTED ATTENUATION

Rewriting the first summation of the cost function in (8) as a sum of quadratics in $[I_k, d_k]$ and dropping terms which do not involve I_k, d_k gives us a new cost function corresponding to the unconstrained optimization problem

$$\tilde{c}(I, d, \lambda) = \frac{1}{2} \sum_{k=1}^{K} \left([I_k \, d_k] \, Q_k \begin{bmatrix} I_k \\ d_k \end{bmatrix} - 2 \, [I_k \, d_k] \, b_k \right) + \lambda \left(\frac{1}{K} \sum_{k=1}^{K} I_k - \bar{I} \right)$$

where Q_k and b_k are defined as in section IV-B. To find the minimum of the cost function, we take gradient with respect to each $[I_k \, d_k]$ and set it to zero. This gives us

$$Q_k \begin{bmatrix} I_k \\ d_k \end{bmatrix} - b_k + \begin{bmatrix} \frac{\lambda}{K} \\ 0 \end{bmatrix} = 0 \qquad (23)$$

and

$$\begin{bmatrix} I_k \\ d_k \end{bmatrix} = Q_k^{-1} \left(b_k - \begin{bmatrix} \frac{\lambda}{K} \\ 0 \end{bmatrix} \right). \qquad (24)$$

If $\tilde{Q}_k = Q_k^{-1}$ then $I_k = \tilde{q}_{k,11}(b_{k,1} - \frac{\lambda}{K}) + \tilde{q}_{k,12} b_{k,2}$. Using this in the constraint $\left(\frac{1}{K} \sum_{k=1}^{K} I_k - \bar{I} \right)$ and solving for λ we get

$$\hat{\lambda} = \frac{\sum_{k=1}^{K} (\tilde{q}_{k,11} b_{k,1} + \tilde{q}_{k,12} b_{k,2}) - K \bar{I}}{\frac{1}{K} \sum_{k=1}^{K} \tilde{q}_{k,11}}. \qquad (25)$$

Finally we can use this value of $\hat{\lambda}$ in (24) to solve for the optimal values of $[I_k \, d_k]$.

REFERENCES

[1] M. Bárcena and A. J. Koster, "Electron tomography in life science," *Seminars Cell & Develop. Biol.*, vol. 20, no. 8, pp. 920–930, Oct. 2009.

[2] P. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: The development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, nos. 3–4, pp. 413–431, Sep. 2003.

[3] H. Li, H. L. Xin, D. A. Muller, and L. A. Estroff, "Visualizing the 3D internal structure of calcite single crystals grown in agarose hydrogels," *Science*, vol. 326, no. 5957, pp. 1244–1247, Nov. 2009.

[4] P. R. Buseck, R. E. Dunin-Borkowski, B. Devouard, R. B. Frankel, M. R. McCartney, P. A. Midgley, M. Pósfai, and M. Weyland, "Magnetite morphology and life on Mars," *Proc. Nat. Acad. Sci. United States Amer.*, vol. 98, no. 24, pp. 13490–13495, Jul. 2001.

[5] H. Wei, Z. Wang, J. Zhang, S. House, Y.-G. Gao, L. Yang, H. Robinson, L. H. Tan, H. Xing, C. Hou, I. M. Robertson, J.-M. Zuo, and Y. Lu, "Time-dependent, protein-directed growth of gold nanoparticles within a single crystal of lysozyme," *Nature Nanotechnol.*, vol. 6, no. 2, pp. 93–97, Jan. 2011.

[6] A. E. Porter, M. Gass, K. Muller, J. N. Skepper, P. A. Midgley, and M. Welland, "Direct imaging of single-walled carbon nanotubes in cells," *Nature Nanotechnol.*, vol. 2, no. 11, pp. 713–717, Oct. 2007.

[7] Z. Saghi, D. J. Holland, R. Leary, A. Falqui, G. Bertoni, A. J. Sederman, L. F. Gladden, and P. A. Midgley, "Three-dimensional morphology of iron oxide nanoparticles with reactive concave surfaces. A compressed sensing-electron tomography (CS-ET) approach," *Nano Lett.*, vol. 11, no. 11, pp. 4666–4673, Sep. 2011.

[8] B. Goris, W. Van den Broek, K. J. Batenburg, H. H. Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, Feb. 2012.

[9] P. A. Midgley and R. E. Dunin-Borkowski, "Electron tomography and holography in materials science," *Nature Mater.*, vol. 8, no. 4, pp. 271–280, 2009.

[10] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Philadephia, PA, USA: Soc. Ind. Appl. Math., 2001.

[11] K. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. Midgley, J. Hernandez, U. Kaiser, E. Encina, E. Coronado, and G. Van Tendeloo, "3D imaging of nanomaterials by discrete tomography," *Ultramicroscopy*, vol. 109, no. 6, pp. 730–740, May 2009.

[12] Z. Yu, J. Thibault, C. Bouman, K. Sauer, and J. Hsieh, "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 161–175, Jan. 2011.

[13] K. Sauer and C. Bouman, "Bayesian estimation of transmission tomograms using segmentation based optimization," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 4, pp. 1144–1152, Aug. 1992.

[14] K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 534–548, Feb. 1993.

[15] J.-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–4544, Oct. 2007.

[16] I. Stojanovic, H. Pien, S. Do, and W. C. Karl, "Low-dose X-ray CT reconstruction based on joint sinogram smoothing and learned dictionary-based representation," in *Proc. 9th IEEE ISBI*, May 2012, pp. 1012–1015.

[17] S. Ramani and J. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imaging*, vol. 31, no. 3, pp. 677–688, Mar. 2012.

[18] J. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 13, no. 2, pp. 290–300, Jun. 1994.

[19] M. Kamasak, C. Bouman, E. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. Med. Imaging*, vol. 24, no. 5, pp. 636–650, May 2005.

[20] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, Aug. 2006.

[21] G. Wang and J. Qi, "Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization," *IEEE Trans. Med. Imaging*, vol. 31, no. 12, pp. 2194–2204, Dec. 2012.

[22] Z. H. Levine, A. J. Kearsley, and J. G. Hagedorn, "Bayesian tomography for projections with an arbitrary transmission function with an application in electron microscopy," *J. Res. Nat. Inst. Stand. Technol.*, vol. 111, no. 6, pp. 411–417, Nov. 2006.

[23] D. Boas, D. Brooks, E. Miller, C. DiMarzio, M. Kilmer, R. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," *IEEE Signal Proc. Mag.*, vol. 18, no. 6, pp. 57–75, Nov. 2001.

[24] J. C. Ye, C. Bouman, K. Webb, and R. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 909–922, Jun. 2001.

[25] W. Van den Broek, S. Van Aert, and D. Van Dyck, "A model based atomic resolution tomographic algorithm," *Ultramicroscopy*, vol. 109, no. 12, pp. 1485–1490, Nov. 2009.

[26] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, no. 2, pp. 113–122, Oct. 1982.

[27] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, Dec. 1994.

[28] H. Erdogan, G. Gualtieri, and J. A. Fessler, "An ordered subsets algorithm for transmission tomography," *Med. Biol.*, vol. 44, no. 11, pp. 2835–2851, Nov. 1998.

[29] E. Mumcuoglu, R. Leahy, S. Cherry, and Z. Zhou, "Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 687–701, Dec. 1994.

[30] J. Fessler and S. Booth, "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction," *IEEE Trans. Image Process.*, vol. 8, no. 5, pp. 688–699, May 1999.

[31] A. Tuysuzoglu and W. C. Karl, "A graph-cut method for tomographic reconstruction," in *Proc. IEEE SSP Workshop*, Aug. 2012, pp. 1–4.

[32] C. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 480–492, Mar. 1996.

[33] B. De Man, S. Basu, J.-B. Thibault, J. Hsieh, J. Fessler, C. Bouman, and K. Sauer, "A study of four minimization approaches for iterative reconstruction in X-ray CT," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5. Oct. 2005, pp. 2708–2710.

[34] S. Venkatakrishnan, L. Drummy, M. Jackson, M. De Graef, J. Simmons, and C. Bouman, "Bayesian tomographic reconstruction for high angle annular dark field (HAADF) scanning transmission electron microscopy (STEM)," in *Proc. IEEE SSP Workshop*, Aug. 2012, pp. 680–683.

[35] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–4544, Oct. 2007.

[36] S. J. Pennycook, S. D. Berger, and R. J. Culbertson, "Elemental mapping with elastically scattered electrons," *J. Microscopy*, vol. 144, no. 3, pp. 229–249, Dec. 1986.

[37] Z. Saghi, X. Xu, Y. Peng, B. Inkson, and G. Mobus, "Three-dimensional chemical analysis of tungsten probes by energy dispersive X-ray nano-tomography," *Appl. Phys. Lett.*, vol. 91, no. 25, pp. 251906-1–251906-3, Dec. 2007.

[38] W. Van den Broek, A. Rosenauer, B. Goris, G. Martinez, S. Bals, S. Van Aert, and D. Van Dyck, "Correction of non-linear thickness effects in HAADF STEM electron tomography," *Ultramicroscopy*, vol. 116, pp. 8–12, May 2012.

[39] P. Ercius, M. Weyland, D. A. Muller, and L. M. Gignac, "Three-dimensional imaging of nanovoids in copper interconnects using in-coherent bright field tomography," *Appl. Phys. Lett.*, vol. 88, no. 24, pp. 243116-1–243116-3, Jun. 2006.

[40] J. M. LeBeau and S. Stemmer, "Experimental quantification of annular dark-field images in scanning transmission electron microscopy," *Ultramicroscopy*, vol. 108, no. 12, pp. 1653–1658, Nov. 2008.

[41] A. Rosenauer, K. Gries, K. Müller, A. Pretorius, M. Schowalter, A. Avramescu, K. Engl, and S. Lutgen, "Measurement of specimen thickness and composition in $Al_xGa_{1-x}N/GaN$ using high-angle annular dark field images," *Ultramicroscopy*, vol. 109, no. 9, pp. 1171–1182, Aug. 2009.

[42] J. W. Cahn and J. E. Hilliard, "Free energy of a nonuniform system. I. Interfacial free energy," *J. Chem. Phys.*, vol. 28, no. 2, pp. 258–267, 1958.

[43] P. Jin, E. Haneda, K. Sauer, and C. A. Bouman, "A model-based 3D multi-slice helical CT reconstruction algorithm for transportation security application," in *Proc. 2nd Int. Conf. Image Inf. X-Ray Comput. Tomography*, Jun. 2012, pp. 1–4.

[44] S. Oh, A. Milstein, C. Bouman, and K. Webb, "A general framework for nonlinear multigrid inversion," *IEEE Trans. Image Process.*, vol. 14, no. 1, pp. 125–140, Jan. 2005.

[45] S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector calibration in three-dimensional Bayesian optical diffusion tomography," *J. Opt. Soc. Amer. A*, vol. 19, no. 10, pp. 1983–1993, Oct. 2002.

[46] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.

[47] K. J. D. Mastronarde and J. McIntosh, "Computer visualization of three-dimensional image data using IMOD," *J. Struct. Biol.*, vol. 116, pp. 71–76, Aug. 1996.

[48] H. Demers, N. Poirier-Demers, A. R. Couture, D. Joly, M. Guilmain, N. de Jonge, and D. Drouin, "Three-dimensional electron microscopy simulation with the CASINO Monte Carlo software," *Scanning*, vol. 33, no. 3, pp. 135–146, May/Jun. 2011.

[49] M. N. Tchoul, S. P. Fillery, H. Koerner, L. F. Drummy, F. T. Oyerokun, P. A. Mirau, M. F. Durstock, and R. A. Vaia, "Assemblies of titanium dioxide-polystyrene hybrid nanoparticles for dielectric applications," *Chem. Mater.*, vol. 22, no. 5, pp. 1749–1759, Jan. 2010.

S. V. Venkatakrishnan (S'12) received the B.Tech. degree in electronics and communication engineering from the National Institute of Technology, Tiruchirappalli, India, in 2007, and the M.S. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2009, and is currently pursuing the Ph.D. degree with the School of Electrical and Computer Engineering, Purdue University. From 2009 to 2010, he was a Research and Development Engineer with Baker Hughes, Inc., Houston, TX, USA, on logging-while-drilling imaging magnetometers. His current research interests include statistical information processing, inverse problems, tomography, and electron microscopy.

Lawrence F. Drummy is a Materials Engineer with the Soft Matter Materials Branch, Functional Materials Division, Materials and Manufacturing Directorate, Air Force Research Laboratory, Dayton, OH, USA. He received the B.S. degree in physics from Rensselaer Polytechnic Institute, Troy, NY, USA, while performing research on scanning tunneling microscopy and image processing of silicon growth on surfaces. In 2003, he received the Ph.D. degree from the Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI, USA, on imaging of defect structures in organic molecular semiconductor thin films using electron microscopy. From 2006 to 2012, he was a Materials Scientist with UES Inc., Beavercreek, OH, USA, and from 2003 to 2006, he was a National Research Council Post-Doctoral Associate with AFRL. His current research interests include electron tomography of biological, polymeric and nanostructured materials, the structure of materials at interfaces, and quantitative morphology characterization of nanoparticles and nanocomposites. He has authored over 50 peer reviewed publications and one issued patent, and has organized workshops and conferences on image processing, tomography, low voltage electron microscopy, materials chemistry, and nanocomposite materials. He has been a Guest Editor for *Polymer Reviews*, *Clays and Clay Minerals*, and *Materials* special issues on topics ranging from polymer microscopy to nanocomposites. He has been a recipient of the Arthur K. Doolittle Award from the American Chemical Society Polymer Materials Science and Engineering Division, the Air Force Office of Scientific Research Star Team, the Rensselaer Medal Award; he was also a Padden Award Finalist in the Division of Polymer Physics of the American Physical Society.

Michael A. Jackson received the B.S. degree in materials science and engineering from Wright State University, Dayton, OH, USA, in 1997. From 1997 to 1999, he was with Pratt & Whitney Aircraft Engines as a Materials Engineer, where he started applying computer methods to small materials science problems. Since 1999, he has successfully designed and implemented a wide range of software systems from the collection and organization of engine inspection data to molecular dynamics simulations and, most recently, designing desktop software that allows an engineer easy access to brand new algorithms developed in academia and government institutions. He is currently the owner of BlueQuartz Software, Springboro, OH, USA.

Marc De Graef received the B.S. and M.S. degrees in physics from the University of Antwerp, Antwerp, Belgium, in 1983, and the Ph.D. degree in physics from the Catholic University of Leuven, Leuven, Belgium, in 1989, with a thesis on copper-based shape memory alloys. He was a Post-Doctoral Researcher with the Materials Department, University of California at Santa Barbara, Santa Barbara, CA, USA, before joining Carnegie Mellon in 1993. He is currently a Professor and Co-Director with the J. Earle and Mary Roberts Materials Characterization Laboratory. His current research interests include materials characterization by means of electron microscopy and X-ray tomography techniques. He is a fellow of the Microscopy Society of America and received the 2012 Educator Award from The Minerals, Metals, and Materials Society.

Jeff Simmons (M'96) is a Research Scientist with the Structural Materials Division, Materials and Manufacturing Directorate, Air Force Research Laboratory. He is currently developing mathematical algorithms for data analysis of emerging large digital datasets produced by advances in microscope characterization capabilities. Efforts center on developing reduced supervision segmentation, tomographic reconstruction, data fusion, and feature extraction algorithms that use materials science specific information for modeling image formation and for regularizing constraints. He also works on developing metric space representations of textures relevant to materials structures. He has worked with physics-based models of developing structures in metals (i.e. phase field) and of deformation of metals at the atomic scale (i.e. embedded atom method). He has advised and served on Ph.D. defense committees of materials science and of signal processing graduate students. He has worked extensively with computational methods and information technologies, having previously established and led a group that initiated development of software tools in the Materials Directorate that was focused on image processing and visualization for large empirical and computational datasets. His leadership experience includes a developing a multiuniversity/industry collaborative effort toward developing advanced algorithms for analysis of digital data as well as one directed toward physics modeling of structural development in materials. He has managed numerous external contracts that involved technologies now referred to as Integrated Computational Materials Engineering. He received the B.S. degree in metallurgical engineering from the New Mexico Institute of Mining and Technology, Socorro, NM, USA, in 1983, the M.Eng. degree in metallurgical engineering and materials science from Carnegie Mellon University, Pittsburgh, PA, USA, in 1985, and the Ph.D. degree in materials science and engineering from Carnegie Mellon in 1992.

Charles A. Bouman (S'86–M'89–SM'97–F'01) received the B.S.E.E. degree from the University of Pennsylvania, Philadelphia, PA, USA, in 1981, and the M.S. degree from the University of California at Berkeley, Berkeley, CA, USA, in 1982. From 1982 to 1985, he was a full staff member with the MIT Lincoln Laboratory and he received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 1989. He joined the faculty of Purdue University, West Lafayette, IN, USA, in 1989, where he is currently the Michael J. and Katherine R. Birck Professor of electrical and computer engineering. He holds a courtesy appointment with the School of Biomedical Engineering and is the Co-Director of Purdue's Magnetic Resonance Imaging Facility, Purdue Research Park, West Lafayette. His current research interests include statistical image models, multiscale techniques, and fast algorithms in applications, including tomographic reconstruction, medical imaging, and document rendering and acquisition. He is a fellow of the American Institute for Medical and Biological Engineering, the Society for Imaging Science and Technology (IS&T), and the SPIE Professional Society. He is a recipient of IS&Ts Raymond C. Bowman Award for outstanding contributions to digital imaging education and research, has been a Purdue University Faculty Scholar, and received the College of Engineering Engagement/Service Award, and Team Award. He was previously the Editor-in-Chief for the IEEE TRANSACTIONS ON IMAGE PROCESSING and a Distinguished Lecturer for the IEEE Signal Processing Society, and he is currently a member of the Board of Governors. He has been an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He has been a Co-Chair of the 2006 SPIE/IS&T Symposium on Electronic Imaging, Co-Chair of the SPIE/IS&T conferences on Visual Communications and Image Processing in 2000, and a Vice President of Publications and a member of the Board of Directors for the IS&T Society, and he is the founder and Co-Chair of the SPIE/IS&T Conference on Computational Imaging.