Model-Based Iterative Reconstruction for Bright-Field Electron Tomography

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

***Abstract*—Bright-Field (BF) electron tomography (ET) has been widely used in the life sciences for 3-D imaging of biological specimens. However, while BF-ET is popular in the life sciences, 3-D BF-ET imaging has been avoided in the physical sciences due to measurement anomalies from crystalline samples caused by dynamical diffraction effects such as Bragg scatter. In practice, these measurement anomalies cause undesirable artifacts in 3-D reconstructions computed using filtered back-projection (FBP). Alternatively, model-based iterative reconstruction (MBIR) is a powerful framework for tomographic reconstruction that combines a forward model for the measurement system and a prior model for the object to obtain reconstructions by minimizing a single cost function. In this paper, we present an MBIR algorithm for BF-ET reconstruction from crystalline materials that can account for the presence of anomalous measurements. We propose a new forward model for the acquisition system which accounts for the presence of anomalous measurements and combine it with a prior model for the object to obtain the MBIR cost function. We then propose a fast algorithm based on majorization–minimization to find a minimum of the corresponding cost function. Results on simulated as well as real data show that our method can dramatically improve reconstruction quality as compared to FBP and conventional MBIR without anomaly modeling.**

***Index Terms*—Model-based reconstruction, tomography, bright-field, electron microscopy.**

I. Introduction

BRIGHT-FIELD (BF) electron tomography (ET) has been widely used in the life sciences to characterize biological specimens in 3D [1] using either a transmission electron microscope (TEM) or a scanning transmission electron microscope (STEM) [2]. BF-ET typically involves focusing an electron beam on a sample, acquiring images of transmitted electrons corresponding to various sample tilts, and using an algorithm on the acquired "tilt-series" to reconstruct the object. In most cases due to the geometry of the acquisition and mechanical limitations of the tilting stages, BF-ET is a limited angle parallel beam transmission tomography modality. Further details of the ET acquisition and preprocessing are discussed in [3].

While there are a few instances where BF-ET has been used in the physical sciences [4]–[6], it has generally been avoided [7], [8], due to the occurrence of contrast reversals [9] from dynamical diffraction effects such as Bragg scatter [10]. Bragg scatter occurs when the crystal lattice is oriented in such a manner that the incident electrons are elastically scattered away from the direct path [10] leading to a measurement uncharacteristic of attenuation due to thickness alone. We refer to measurements which are not well modeled by attenuation due to thickness alone as anomalous measurements. These anomalies make interpretation of the individual BF images complicated and result in strong artifacts if the BF tilt-series is used for tomographic reconstruction using standard reconstruction algorithms such as FBP [8]. Thus BF-ET has generally been avoided in the physical sciences due to the complicated nature of the data and the inability of the standard reconstruction algorithms like FBP to handle such data.

Model-based iterative reconstruction (MBIR) provides a powerful framework for tomographic reconstruction that uses a probabilistic model for the measurement (forward model) and a probabilistic model for the object (prior model) to obtain reconstructions that are qualitatively superior and quantitatively accurate for a variety of applications [11]–[16]. Typically MBIR involves the design and minimization of a cost function corresponding to the maximum *a posteriori* probability (MAP) estimate with two sets of terms - one corresponding to a likelihood for the data and the other corresponding to a prior model for the object. While most efforts for BF-ET have used the FBP algorithm [2], [4]–[6], [8] Levine [17] has developed a MBIR algorithm for BF-ET in the case of thick specimens. However his work deals with amorphous samples, for which there are no anomalies due to dynamical diffraction in the measurement.

In this paper, we present an MBIR algorithm for accurate reconstruction of BF-ET data [18] containing anomalous measurements that typically result from crystalline samples. Our approach is based on a novel generalized Huber function that is used in the forward model (i.e., log likelihood) to account for the anomalous measurements due to Bragg or other errors. The generalized Huber function is parameterized so that it can

Manuscript received May 03, 2014; July 29, 2014 and September 27, 2014; accepted October 29, 2014. Date of publication December 04, 2014; date of current version June 26, 2015. This work was supported by AFOSR/MURI Grant FA9550-12-1-0458, by UES, Inc., under the Broad Spectrum Engineered Materials contract, and by the Electronic Imaging component of the ICMD program of the Materials and Manufacturing Directorate of the Air Force Research Laboratory, A. Rosenberger, program manager. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jeffrey A. Fessler.

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, Wright Patterson AFB, OH 45433 USA (e-mail: Lawrence.Drummy.1@us.af.mil; Jeff.Simmons.3@us.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 15233 USA (e-mail: degraef@cmu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2014.2371751

model the heavy tailed distribution of the errors that are present in anomalous measurements. Using this forward model, we derive an MBIR cost function which allows for joint estimation of both the unknown image, f, and a key parameter of the generalized Huber function. This approach allows for adaptive parameter estimation in the reconstruction process, which is important in practical applications.

We also propose an optimization algorithm that is based on majorization-minimization using a surrogate function [19], [20] together with a fast multiresolution iterative algorithm to find a minima of the MBIR cost function. The surrogate function that we propose bounds the MBIR cost as a function of both the unknown image, f, and the unknown model parameters of the generalized Huber function. Consequently, minimization of the surrogate function allows for efficient estimation of model parameters during the MBIR reconstruction process.

Importantly, the surrogate function requires the introduction of a binary auxiliary variable for each measurement. This auxiliary variable has the interpretation of forming a binary mask that classifies each measurement as either anomalous or normal. During the reconstruction process, the anomalous measurements are then attenuated relatively to normal measurements. The resulting binary mask also contains potentially important information about the Bragg scatter and consequently the crystal orientation for each particle in the 3D reconstruction. Therefore, we also propose an algorithm that extracts a Bragg feature vector for each particle in the volume by correlating the projection of each segmented particle with the binary anomaly mask at each view angle. We conjecture that the extracted Bragg feature vector for each particle can provide useful information in applications.

We apply our method to simulated data containing Bragg scatter like anomalies as well as real TEM data from crystalline particles. Results from the simulated as well as the real data set show that MBIR with anomaly modeling can significantly improve the reconstruction quality compared to FBP and conventional MBIR, suppressing the artifacts that arise due to the anomalous measurements. We also use our new method to extract a Bragg feature vector for each particle and demonstrate how this feature vector can potentially provide useful information about the crystal orientation for each particle in the 3D volume. The source code along with a GUI application implementing our method is publicly available at the website - *www.openmbir.org*.

The organization of the rest of the paper is as follows. In section II we introduce a new statistical model for the measurement system and formulate the MBIR cost function. In section III we propose an efficient algorithm to minimize the cost function. In section IV we present results from a simulated data set, followed by results from a real data set. Finally, in section V we draw our conclusions.

II. Statistical Model and Cost Formulation

The goal of BF-ET is to reconstruct an attenuation coefficient at every point in the sample. The attenuation coefficient is related to the ability of the material to scatter the incident beam away from the direct path which is dependent on the differential cross section, geometry of the detector, density of the material and incident electron energy. An electron beam is focused on the material and the electrons that are scattered by the sample through small angles are captured by a BF detector to obtain a single image. The sample is then tilted along a fixed axis and the process is repeated. Thus, at the end of the acquisition, we obtain a collection of BF images that can be used for tomographic reconstruction of the attenuation coefficients.

In order to reconstruct the attenuation coefficients associated with the sample, we use an MBIR framework. The reconstruction in the MBIR framework is typically given by the joint-MAP [21] estimate

$$(\hat{f}, \hat{\phi}) = \underset{f,\phi}{\operatorname{argmin}} \{-\log p(g|f,\phi) - \log p(f)\} \tag{1}$$

where g is the vector of measurements, f is the vector of unknown voxels (attenuation coefficients), ϕ is a vector of unknown calibration parameters (nuisance parameters), $p(g|f,\phi)$ is the likelihood (forward model) and $p(f)$ is the prior probability for the unknown voxels.

In order to develop a forward model for BF-ET that accounts for the anomalous effects, we start with the simple case when there are no anomalies. Let $\lambda_{k,i}$ be the electron counts corresponding to the i^{th} measurement at the k^{th} tilt and $\lambda_{D,k}$ be the counts that would be measured in the absence of the sample at that tilt (blank scan). We model the attenuation of the beam through the material using Beer's law. Thus, an estimate of the projection integral corresponding to the i^{th} measurement at the k^{th} tilt is given by $\log\left(\frac{\lambda_{D,k}}{\lambda_{k,i}}\right)$. Notice that unlike in high-angle annular dark-field electron microscopy [12], the BF case requires a $\log$ operation to be applied to a normalized version of the measurement. There can be cases in which the blank scan, $\lambda_{D,k}$, is not measured, and we can include it as an unknown parameter in the MBIR framework and estimate it as a part of the reconstruction. If g_k is a $M \times 1$ vector with $g_{k,i} = -\log(\lambda_{k,i})$, f is an $N \times 1$ vector of unknown attenuation coefficients of the material and, $d_k = -\log(\lambda_{D,k})$, then, assuming $\lambda_{k,i}$'s are conditionally independent Poisson random variables it has been shown that [22] the conditional mean of $g_{k,i}$ can be approximated by $A_{k,i,*}f + d_k$ and the conditional variance is proportional to $\frac{1}{\mathbb{E}[\lambda_{k,i}]}$, where A_k is a $M \times N$ forward projection matrix and $A_{k,i,*}$ is the i^{th} row of A_k. Modeling the conditional density of the measurements as a Gaussian distribution [23] results in a probability density function (pdf),

$$p(g|f,d,\sigma) = \frac{1}{Z_l} \exp\left\{-\frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M}(g_{k,i} - A_{k,i,*}f - d_k)^2 \frac{\Lambda_{k,ii}}{\sigma^2}\right\} \tag{2}$$

where K is the total number of tilts, $g = [g_1^t \cdots g_K^t]^t$ is a $KM \times 1$ data vector, Λ_k is a diagonal matrix with entries set so that $\frac{\sigma^2}{\Lambda_{k,ii}}$ is the variance of the measurement $g_{k,i}$ with σ^2 being a proportionality constant, $d = [d_1 \cdots d_K]$ is a vector containing the offset parameters, and Z_l is a normalizing constant. For such a transmission tomography model it has been shown that $\Lambda_{k,ii} = \mathbb{E}[\lambda_{k,i}] \approx \lambda_{k,i}$ [24]. We note that our formulation can account for more sophisticated physics models as introduced in [25], but in this paper we focus on using Beer's law as it has been found to be accurate for a class of materials and thickness typically studied using BF-ET [25].

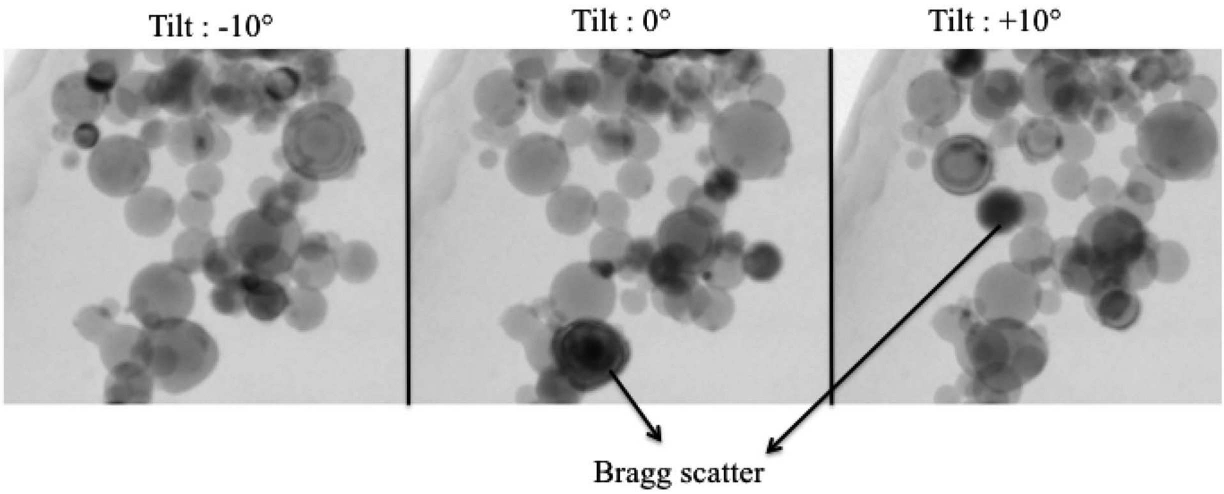

Fig. 1. Illustration of the "anomalies" present in a real BF-TEM data set of Aluminum nanoparticles. The figure shows BF images corresponding to three different tilts of the specimen. Note that certain spheres turn dark (fewer counts) and then again turn bright due to Bragg scatter (contrast reversal). These effects make it challenging to directly apply standard analytic tomographic reconstruction algorithms to the data.

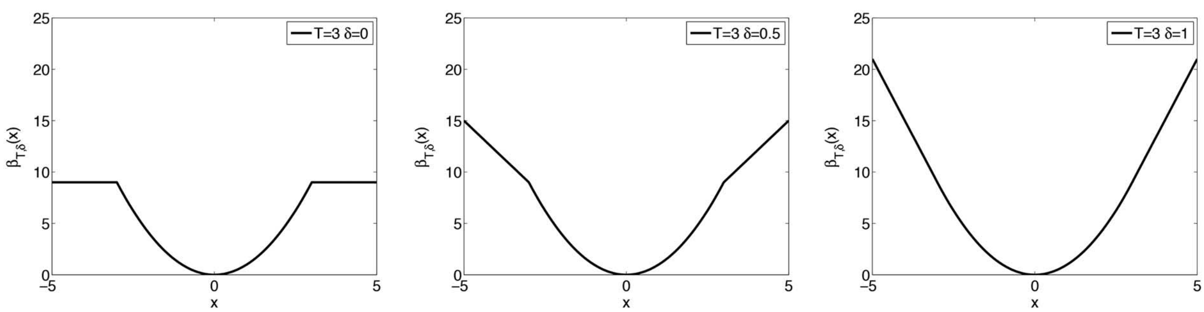

Fig. 2. Illustration of the generalized Huber function $\beta_{T,\delta}$ used for the likelihood term with $T = 3$ and $\delta = 0, 0.5$ and 1. When $\delta = 1$ the function reduces to the Huber function. Large model mismatch errors are penalized by restricting their influence on the overall cost function.

A. Generalized Huber Functions for Anomaly Modeling

Bragg scatter from crystalline material can cause the BF-ET measurements to vary substantially from the model of equation (2). Fig. 1 shows an example of three tilts from a BF tilt series with regions having significant anomalies.

A precise way of accounting for these anomalies would require identifying 3D regions of the object that consist of a single crystal, and modeling the associated crystal structure. While possible, this would be a highly ill-posed inverse problem to recover from a single 2D tilt series due to the unknown 3D orientation of the crystals. Furthermore, modeling other classes of anomalies such as Fresnel fringes [26] and extinction contours involve more complex physics making the data more difficult to invert. Therefore, instead of modeling the complicated physics of dynamical diffraction that leads to anomalies, we will use an alternate approach.

In order to account for anomalous effects like Bragg scatter, we propose a modified likelihood function that models the anomalies as outliers of a pdf

$$p(g|f,d,\sigma) = \frac{1}{Z}\exp\left\{-\frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M}\beta_{T,\delta}\left((g_{k,i} - A_{k,i,*}f - d_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma}\right)\right\} \tag{3}$$

where $\beta_{T,\delta} : \mathbb{R} \to \mathbb{R}$ such that

$$\beta_{T,\delta}(x) = \begin{cases} x^2 & |x| < T \\ 2\delta T|x| + T^2(1-2\delta) & |x| \geq T \end{cases}$$

and Z is a normalizing constant. We call $\beta_{T,\delta}$ the generalized Huber function. Fig. 2 shows the generalized Huber function for three different values of δ. Notice that δ controls the tail behavior of the density function while T controls the threshold beyond which a measurement is considered anomalous. When $\delta = 0$, $\beta_{T,\delta}$ corresponds to the weak-spring potential [27] used for image modeling and results in a function with the heaviest tails among the three cases. However, when $\delta = 0$ we cannot jointly estimate the calibration parameters because the likelihood is not a valid density function since it does not integrate to 1. When $\delta = 1$, $\beta_{T,\delta}$ reduces to the Huber function [28] which is a convex function and corresponds to a pdf with the lightest tail among the three cases. When T is very large then $\beta_{T,\delta}$ is effectively a quadratic function and the likelihood reduces to the standard transmission tomography model in (2). Thus the generalized Huber function can be adjusted to have heavier tails than the density function in (2) to account for the various anomalies in the data set.

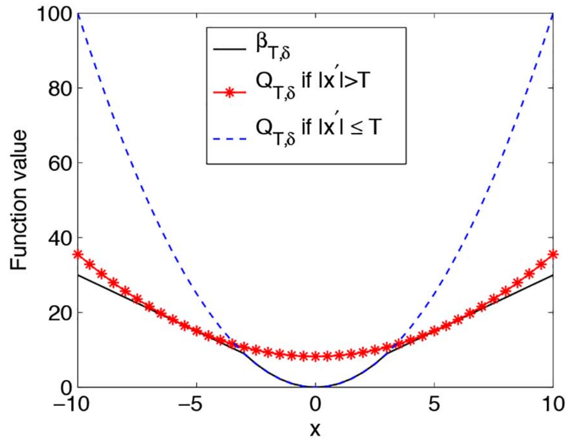

Fig. 3. Illustration of the surrogate function, $Q_{T,\delta}$, to the generalized Huber function with $T = 3$ and $\delta = \frac{1}{2}$. The surrogate function is plotted for two cases : one when $x' = 2$ and the other when $x' = 5$. The surrogate function in each case is a quadratic function thereby simplifying the subsequent optimization.

Restricting $0 < \delta \le 1$ and using the fact that

$$\int p(g|f,d,\sigma)dg = 1$$

we can show that the normalizing constant has the form

$$Z = \sigma^{MK} \times \text{Constants}.$$

Hence, the modified log-likelihood function for the BF-ET case is given by

$$\begin{aligned} &-\log p(g|f,d,\sigma) \\ &= \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M}\beta_{T,\delta}\left((g_{k,i} - A_{k,i,*}f - d_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma}\right) \\ &\quad + MK\log(\sigma) + \text{Constants}. \end{aligned} \tag{4}$$

Each term in the summation corresponds to a penalty on the ratio of the data mismatch error $(g_{k,i} - A_{k,i,*}f - d_k)$ to the noise standard deviation $\left(\frac{\sigma}{\sqrt{\Lambda_{k,ii}}}\right)$. Thus T has the interpretation that if the data fit error is greater than T times the noise standard deviation then that measurement is likely to be an anomaly. Notice that typically σ is not known and hence we will jointly estimate it as a part of the reconstruction.

B. MBIR Cost Formulation

Combining the statistical model in (4) with a prior model of the form

$$p(f) = \frac{1}{Z_f}\exp\{-s(f)\} \tag{5}$$

where Z_f is a normalizing constant, the reconstruction is obtained by minimizing the cost

$$\begin{aligned} &c(f,d,\sigma) \\ &= \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M}\beta_{T,\delta}\left((g_{k,i} - A_{k,i,*}f - d_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma}\right) \\ &\quad + MK\log(\sigma) + s(f). \end{aligned} \tag{6}$$

function $[\hat{f}, \hat{d}, \hat{\sigma}] \leftarrow$ RECONSTRUCT(g, f', d',σ')
  %**Inputs**: Measurements g, Initial reconstruction f', Initial offset d', Initial variance parameter σ'
  %**Outputs**: Reconstruction $\hat{f}$ and nuisance parameters $(\hat{d}$,$\hat{\sigma})$
  $e' = g - Af' - d'\mathbb{1}$ ▷ Initialize error vector
  while Stopping criteria is not met **do**
    for each voxel j **do** ▷ Voxel updates
      Compute θ_1 and θ_2 using (17)
      for $k \in \mathcal{N}_j$ **do**
        Compute surrogate function parameter a_{jk} using (11)
      end for
      Compute u^* using (18)
      $f_j \leftarrow \max(u^*, 0)$
      $e' \leftarrow e' - (f_j - f'_j)A_{*,j}$
      $f'_j \leftarrow f_j$
      Update b' using (15)
    end for
    $d' \leftarrow$ Update d using (20) ▷ Offset parameter update
    Update e'
    Update b' using (15)
    $\sigma' \leftarrow$ Update σ using (21) ▷ Variance parameter update
    Update b' using (15)
  end while
  $\left(\hat{f}, \hat{d}, \hat{\sigma}\right) \leftarrow (f', d', \sigma')$
end function

Fig. 4. MBIR algorithm for BF data with anomalies. The algorithm works by constructing a surrogate to the original function based on the current values of the voxels and nuisance parameter and minimizing this surrogate function with respect to each variable. The process is then repeated. The algorithm can be efficiently implemented by keeping track of the error sinogram, e', and the anomaly classifier vector, b'.

Alternately, we can define $h_{k,i} : \mathbb{R}^{N+K+1} \to \mathbb{R}$ to be a function such that

$$h_{k,i}(f,d,\sigma) = (g_{k,i} - A_{k,i,*}f - d_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma}.$$

The cost function can then be written as

$$\begin{aligned} c(f,d,\sigma) = \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M}\beta_{T,\delta}(h_{k,i}(f,d,\sigma)) \\ + MK\log(\sigma) + s(f). \end{aligned} \tag{7}$$

Additionally, we will constrain $f \ge 0$, as it is physically meaningful to have positive values of the attenuation coefficients. Thus, the MBIR BF-ET reconstruction is given by

$$\left(\hat{f}, \hat{d}, \hat{\sigma}\right) \leftarrow \underset{f\ge 0,d,\sigma}{\arg\min}\; c(f,d,\sigma).$$

The cost function (7) is non-convex in general, and thus finding the global minimum is computationally expensive. Therefore we will present an algorithm to find a local minimum of the cost. Furthermore, since the likelihood term of (7), is

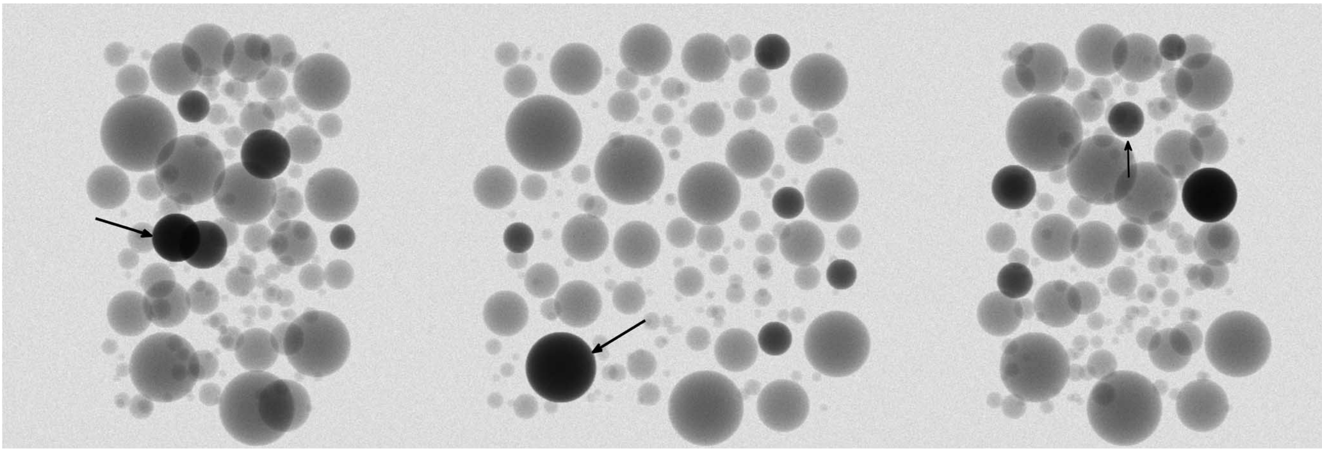

Fig. 5. Simulated BF data corresponding to a 3D phantom of spheres for three successive tilts. The arrows in the figure show example locations with the simulated Bragg scatter obtained by increasing the attenuation coefficient of a few spheres in the phantom. We model these as anomalies in the projection data as they can cause artifacts in the reconstructions produced using the standard reconstruction techniques.

not differentiable, gradient-based methods can not be directly applied. Hence, we use a majorization-minimization strategy [19], [20] combined with alternating minimization to find a minimum of the cost.

III. Optimization Algorithm

Our optimization approach is based on the repeated minimization of a differentiable surrogate function. The function $q(z; z')$ is a surrogate function for the function $t(z)$ at the point z' if the following two conditions hold.

$$\begin{aligned} q(z; z') &\geq t(z) \\ q(z'; z') &= t(z'). \end{aligned} \tag{8}$$

If $Q(f, d, \sigma; f', d', \sigma')$ is a surrogate function to $c(f, d, \sigma)$ at the point (f', d', σ'), our algorithm consists of repeating the following steps until convergence

1) For each voxel j

$$f'_j \leftarrow \underset{f_j \geq 0, f_k = f'_k \forall k \neq j}{\operatorname{argmin}} Q(f, d', \sigma'; f', d', \sigma')$$

2) $d' \leftarrow \underset{d}{\operatorname{argmin}}\, Q(f', d, \sigma'; f', d', \sigma')$

3) $\sigma' \leftarrow \underset{\sigma}{\operatorname{argmin}}\, Q(f', d', \sigma; f', d', \sigma')$

The algorithm is terminated if the ratio of the average change in the magnitude of the reconstruction to the average magnitude of the reconstruction is less than a preset threshold. In addition we use a multiresolution initialization [29] to speed up the convergence of the algorithm and prevent the method from getting stuck in undesirable local minima. In multi-resolution initialization, we perform a reconstruction at a coarser resolution (larger voxel sizes) and use this result to initialize a finer resolution reconstruction. This transfers the computational load to the coarser scale where the optimization can be done quickly due to the reduced dimensionality of the problem.

Note that the surrogate function approach ensures monotonic decrease of the original cost function (7) with each update; so the sequence of costs must be convergent. In addition we have empirically observed that the reconstructions also converge. While theoretical convergence proofs exist for majorization techniques with alternating minimization [30], [31], we have no formal proof of convergence in this case due to the complicated nature of the cost function.

To derive the exact updates for the above algorithm we will first design a surrogate function to the original cost assuming any general prior model $s(f)$. Next we will present a specific $s(f)$ and derive a surrogate for this case and use it to derive the update equations for each iteration.

A. Construction of Surrogate Function

We design surrogate functions for each function $\beta_{T,\delta}(h_{k,i}(f, d, \sigma))$ in (7) at a given point (f', d', σ') and sum them up to form a surrogate to the overall cost function. In order to design a surrogate function, note that each term in the summation in (7) is a composition between the generalized Huber function $\beta_{T,\delta}$ and the function $h_{k,i}$. Therefore, we first design a surrogate function to the generalized Huber function, $\beta_{T,\delta}$, and then use this function along with a composition property to design a surrogate function to the composition $\beta_{T,\delta} \circ h_{k,i}$.

In particular

$$Q_{T,\delta}(x; x') = \begin{cases} x^2 & |x'| < T \\ \frac{\delta T}{|x'|} x^2 + \delta T |x'| + T^2(1 - 2\delta) & |x'| \geq T \end{cases}$$

is a surrogate function to $\beta_{T,\delta}(x)$. Fig. 3 shows the construction of a surrogate function to the generalized Huber function for the case when $T = 3$ and $\delta = 0.5$. Notice that while the generalized Huber function is non-differentiable, the surrogate function is quadratic and hence differentiable in x.

Next, we will use the composition property of surrogate functions to design a surrogate function for each $\beta_{T,\delta} \circ h_{k,i}$ in (7). The composition property of surrogate function states that if $q(z; z')$ is a surrogate function to $t(z)$ at z' then $q(h(z); h(z'))$ is a surrogate function to $t(h(z))$ at z' (proof in Appendix A). Using the composition property of surrogate functions, the composition of $Q_{T,\delta}$ with $h_{k,i}$, $Q_{T,\delta}(h_{k,i}(f, d, \sigma); h_{k,i}(f', d', \sigma'))$ is a surrogate function to $\beta_{T,\delta}(h_{k,i}(f, d, \sigma))$ in (7).

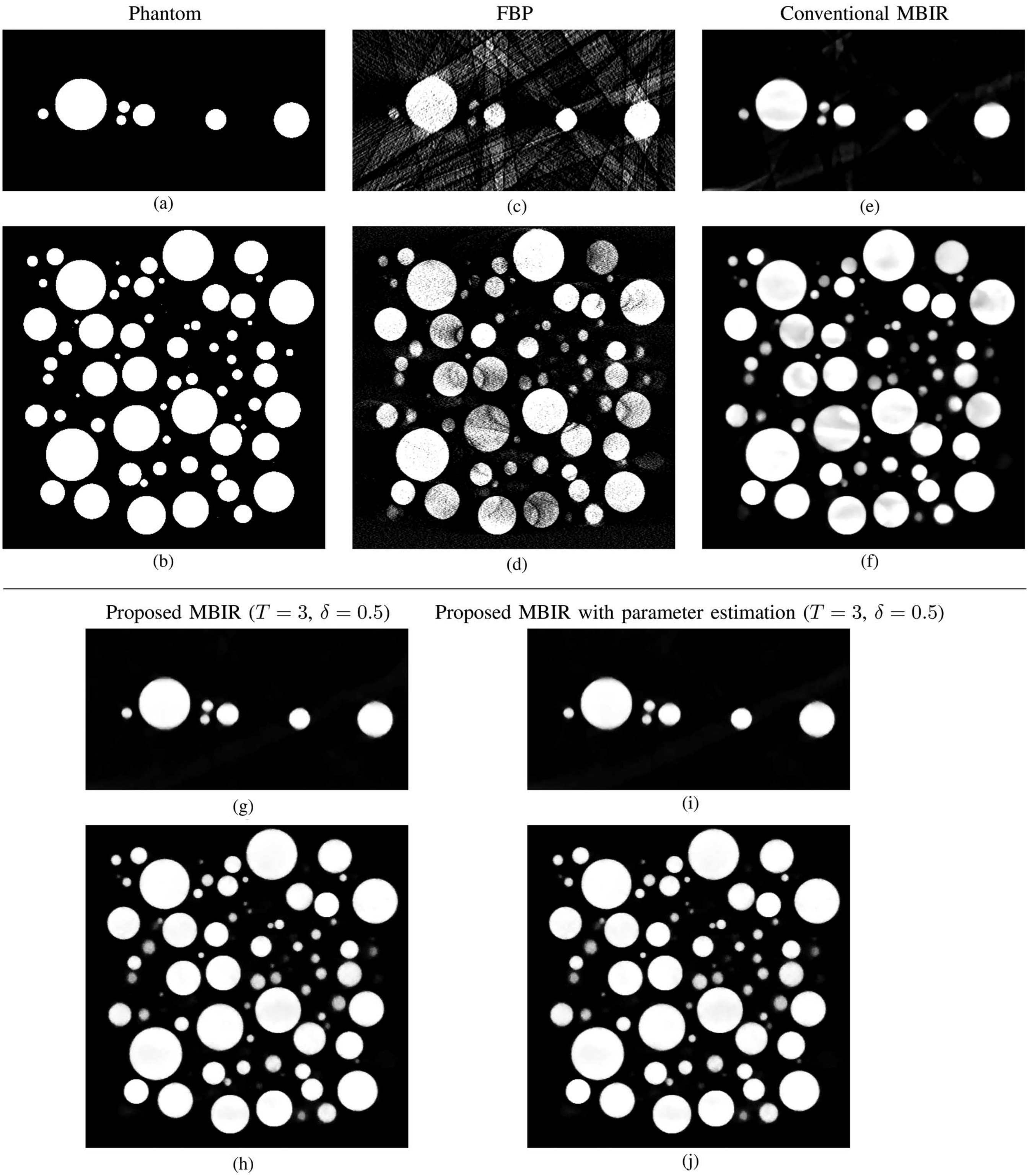

Fig. 6. Comparison of BF reconstructions for a data set with Bragg scatter like anomalies. (a) and (b) show a single $x - z$ and $x - y$ cross-section from the phantom used. The horizontal direction represents the x axis. (c) and (d) show the corresponding cross sections from a FBP reconstruction. (e) and (f) show the conventional MBIR reconstruction. The reconstruction has streaks because of Bragg scatter but much lesser compared to FBP. (g) and (h) show the cross-section from MBIR with anomaly modeling ($T = 3$ and $\delta = 0.5$). The method effectively suppresses the artifacts in (c) - (f), and produces a more accurate reconstruction. Finally (i) and (j) show the reconstruction using MBIR with anomaly modeling and nuisance parameter estimation. The reconstructions are comparable to the MBIR-AM case showing that the algorithm can work well despite of the unknown parameters. All images are scaled in the range of $0 - 7.45 \times 10^{-3}\ \mathrm{nm}^{-1}$.

TABLE I
COMPARISON OF THE ROOT MEAN SQUARE ERROR OF THE RECONSTRUCTION WITH RESPECT TO THE ORIGINAL PHANTOM FOR VARIOUS SCENARIOS. MBIR WITH ANOMALY MODELING PRODUCES QUANTITATIVELY MORE ACCURATE RECONSTRUCTIONS

Algorithm	RMSE ($\times 10^{-4}$ nm^{-1})
FBP	13.90
MBIR	4.95
MBIR-AM ($T = 3, \delta = 0.5$)	4.30
MBIR-AM-PE ($T = 3, \delta = 0.5$)	4.31

Using the surrogate function for each $\beta_{T,\delta}(h_{k,i}(f, d, \sigma))$,

$$\begin{aligned}\tilde{Q}&(f, d, \sigma; f', d', \sigma')\\ &= \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M} Q_{T,\delta}(h_{k,i}(f, d, \sigma); h_{k,i}(f', d', \sigma'))\\ &\quad + MK\log(\sigma) + s(f) \end{aligned} \tag{9}$$

is a surrogate function to the original cost (7). Hence, even though the terms corresponding to the generalized Huber function in the original cost function may be non-differentiable, we have constructed a surrogate function which overcomes this difficulty and makes the optimization tractable.

B. Prior Model and Surrogate Function

We use a special case of the q-generalized Gaussian Markov random field (qGGMRF) [32] for the prior, resulting in

$$s(f) = \sum_{\{j,k\}\in\mathcal{N}} w_{jk}\rho(f_j - f_k)$$

$$\rho(f_j - f_k) = \frac{\left|\frac{f_j - f_k}{\sigma_f}\right|^2}{c + \left|\frac{f_j - f_k}{\sigma_f}\right|^{2-p}}$$

$\mathcal{N}$ is the set of pairs of neighboring voxels (e.g. a 26 point neighborhood), and p, c and σ_f are qGGMRF parameters. The weights w_{jk} are inversely proportional to the distance between voxels j and k, normalized to 1. We fix $c = 0.001$ and restrict $1 \le p \le 2$ similarly to [12].

In order to simplify the optimization, we also introduce a surrogate function for the prior with the form

$$\rho(f_j - f_k; f'_j - f'_k) = \frac{a_{jk}}{2}(f_j - f_k)^2 + b_{jk}. \tag{10}$$

The values of a_{jk} and b_{jk} can be derived as shown in [12] and are given by

$$a_{jk} = \begin{cases} \frac{\rho'(f'_j - f'_k)}{(f'_j - f'_k)} & f'_j \neq f'_k \\ \rho''(0) & f'_j = f'_k \end{cases} \tag{11}$$

$$b_{jk} = \rho(f'_j - f'_k) - \frac{a_{jk}}{2}(f'_j - f'_k)^2. \tag{12}$$

Thus a surrogate function to $s(f)$ at $f = f'$ is given by

$$s(f; f') = \sum_{\{j,k\}\in\mathcal{N}} w_{jk}\rho(f_j - f_k; f'_j - f'_k). \tag{13}$$

Substituting (13) into (9) results in the final surrogate function given by

$$\begin{aligned}Q&(f, d, \sigma; f', d', \sigma')\\ &= \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M} Q_{T,\delta}(h_{k,i}(f, d, \sigma); h_{k,i}(f', d', \sigma'))\\ &\quad + MK\log(\sigma) + \sum_{\{j,k\}\in\mathcal{N}} w_{jk}\rho(f_j - f_k; f'_j - f'_k). \end{aligned} \tag{14}$$

In order to simplify the subsequent updates, we define the following binary indicator variable,

$$b'_{k,i} = \begin{cases} 1 & |(g_{k,i} - A_{k,i,*}f' - d'_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'}| < T \\ 0 & |(g_{k,i} - A_{k,i,*}f' - d'_k)\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'}| \ge T. \end{cases} \tag{15}$$

Intuitively $b'_{k,i}$ indicates if a given measurement is classified as anomalous or not, based on the current state of the reconstruction. If we define the error $e_{k,i} = g_{k,i} - A_{k,i,*}f - d_k$ and $e'_{k,i} = g_{k,i} - A_{k,i,*}f' - d'_k$ we can rewrite (14) as

$$\begin{aligned}Q&(f, d, \sigma; f', d', \sigma')\\ &= \frac{1}{2}\sum_{k=1}^{K}\sum_{i=1}^{M} e_{k,i}^2\frac{\Lambda_{k,ii}}{\sigma^2}\left(b'_{k,i} + (1 - b'_{k,i})\frac{\delta T\sigma'}{|e'_{k,i}|\sqrt{\Lambda_{k,ii}}}\right)\\ &\quad + MK\log(\sigma) + \sum_{\{j,k\}\in\mathcal{N}} w_{jk}\rho(f_j - f_k; f'_j - f'_k)\\ &\quad + \text{Terms not dependent on}(f, d, \sigma). \end{aligned} \tag{16}$$

Thus we observe that the surrogate function can be easily constructed by maintaining an error vector, e', based on the current values of the unknowns, and the indicator variable, b'.

C. Update Equations Corresponding to the Surrogate Function

1) Voxel Update: The voxels are updated in random order similarly to [11] in order to speed up the overall convergence of the algorithm. To speed up the implementation of the algorithm the voxel updates are parallelized along the y-direction similar to [33], which also ensures a monotonic decrease of the cost function. To minimize the surrogate function with respect to voxel j, we can fix $f_k = f'_k\ \forall k \in \{1, \cdots, M\} \setminus \{j\}$, $d = d'$ and $\sigma = \sigma'$ in (16). The cost function to minimize is

$$\tilde{c}_{\text{sub}}(u) = \theta_1 u + \frac{\theta_2}{2}\left(u - f'_j\right)^2 + \sum_{k\in\mathcal{N}_j} w_{jk}\rho(u - f'_k; f'_j - f'_k)$$

where $\mathcal{N}_j$ is the set of all neighbors of voxel j and

$$\begin{aligned}\theta_1 &= -\sum_{k=1}^{K}\sum_{i=1}^{M} A_{k,i,j}\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'}\left[b'_{k,i}e'_{k,i}\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'} + (1 - b'_{k,i})\delta T\frac{e'_{k,i}}{|e'_{k,i}|}\right]\\ \theta_2 &= \sum_{k=1}^{K}\sum_{i=1}^{M} A_{k,i,j}^2\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'}\left[b'_{k,i}\frac{\sqrt{\Lambda_{k,ii}}}{\sigma'} + (1 - b'_{k,i})\frac{\delta T}{|e'_{k,i}|}\right]. \end{aligned} \tag{17}$$

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

$$u^* = \frac{\sum_{k\in\mathcal{N}_j} w_{jk}a_{jk}f'_k + \theta_2 f'_j - \theta_1}{\sum_{k\in\mathcal{N}_j} w_{jk}a_{jk} + \theta_2}. \tag{18}$$

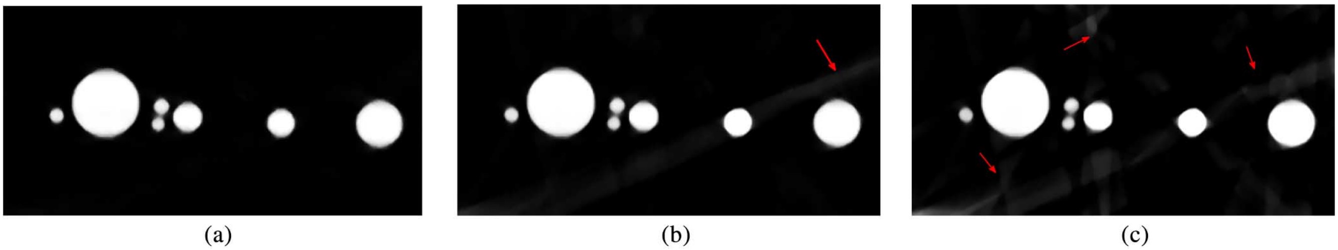

Fig. 7. Illustrates the impact of varying anomaly threshold T on the proposed MBIR reconstructions. (a) shows an $x-z$ cross section from the 3-D reconstruction when $T=1$. (b) and (c) shows the corresponding slices when $T=5$ and $T=20$. Notice that for (b) and (c) there are visible streaking artifacts. A value of $T=3$ as shown in Fig. 6 produces an accurate reconstruction.

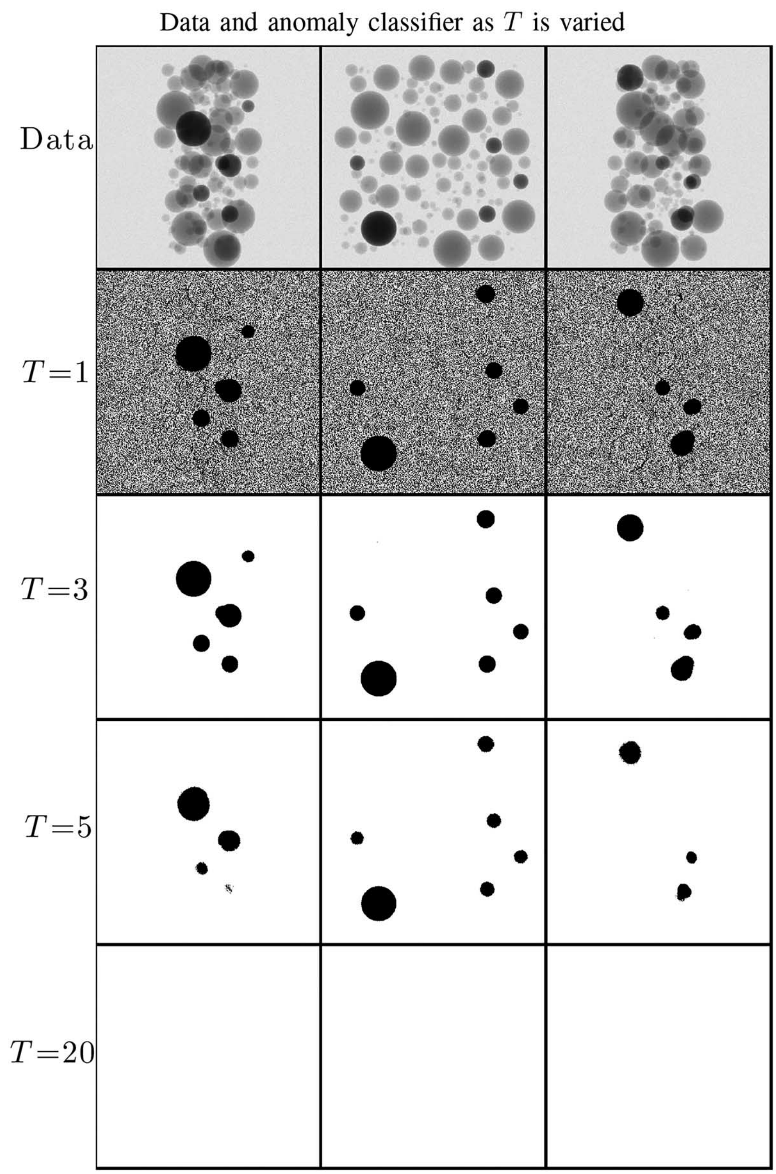

Fig. 8. Data (top row) and corresponding anomaly classifier upon termination of the algorithm for three tilts from the phantom data set corresponding to different values of T. The white regions indicate areas classified as non-anomalous and the black regions correspond to the anomalies identified by the algorithm. As the value of T increases the algorithm starts to misclassify anomalies. A value of $T=3$ provides a good trade off between the false alarms and missed detections.

Enforcing the positivity constraint, the update for the voxel is

$$f_j' \leftarrow \max\left(u^*, 0\right). \tag{19}$$

2) Offset Parameter Update: In order to minimize the surrogate function with respect to the offset parameter d, we take the gradient of the surrogate function (16)

$Q(f', d, \sigma'; f', d', \sigma')$ with respect to d and set it to zero. This gives the optimal update as

$$d'_k \leftarrow d'_k + \frac{\sum_{i=1}^{M} \sqrt{\Lambda_{k,ii}} \left[e'_{k,i} b'_{k,i} \frac{\sqrt{\Lambda_{k,ii}}}{\sigma'} + \delta T \frac{e'_{k,i}}{|e'_{k,i}|} (1 - b'_{k,i}) \right]}{\sum_{i=1}^{M} \sqrt{\Lambda_{k,ii}} \left[b'_{k,i} \frac{\sqrt{\Lambda_{k,ii}}}{\sigma'} + \frac{\delta T}{|e'_{k,i}|} (1 - b'_{k,i}) \right]}. \tag{20}$$

3) Variance Parameter Update: In order to update the variance parameter we minimize the surrogate function (16) with respect to σ setting $f = f'$ and $d = d'$. This gives the optimal update as

$$\sigma' \leftarrow \sqrt{\frac{\sum_{k=1}^{K} \sum_{i=1}^{M} e'^2_{k,i} \Lambda_{k,ii} b'_{k,i} + \sum_{k=1}^{K} \sum_{i=1}^{M} (1 - b'_{k,i}) \delta T |e'_{k,i}| \sigma' \sqrt{\Lambda_{k,ii}}}{MK}}. \tag{21}$$

The MBIR BF-ET algorithm for a single resolution is summarized in Fig. 4.

D. Initialization

Since the MBIR cost function is non-convex, initializing the algorithm with a reasonable initial estimate is important. We use a multi-resolution initial condition to prevent the algorithm from becoming stuck in undesirable local minima. We initialize the values of f to 0 nm^{-1} at the coarsest scale. The value of d and σ are initialized from a region of the image where there is no sample present. Furthermore, at the coarsest scale we perform a few iterations (typically 10) over the voxels with the value of T set to be very large in order to obtain a reasonable initial condition for the overall multiresolution algorithm. Since the size of the voxels is large at the coarse scales, this initialization is computationally inexpensive to perform.

IV. Results

In this section we compare four algorithms for BF-ET: FBP, conventional model-based iterative reconstruction (MBIR), the proposed MBIR with anomaly modeling and known parameter values (MBIR-AM), and the proposed method with anomaly modeling and parameter estimation (MBIR-AM-PE). We apply the methods to simulated data as well as real data. For the simulated data we will compare results from all four methods while in the real data, since we do not know the parameters, we will not consider the MBIR-AM case. Finally, we will present a method for using the anomalies identified by our method to associate a Bragg feature vector for each particle in the reconstructed volume.

The FBP reconstructions are performed in Matlab using the *iradon* command and the output is clipped to be positive. For the MBIR reconstructions with anomaly modeling, we set $T = 3$, $\delta = 0.5$, and $p = 1.2$. The value of σ_f is chosen to obtain the lowest root mean square error (RMSE) for the simulated data set and is chosen to obtain the best visual quality of reconstruction for the real data set. Since our prior behaves similar to a GGMRF [34], we adapt the scaling parameter σ_f according to Eq. 28 in [35] for the multi-resolution reconstructions. The offset parameter for each tilt, d_k, is initialized to the mean value of the $\log$ of the measurements from a void region in the sample. The variance parameter, σ^2, is initialized as the ratio of the mean value of the $\log$ measurements to the mean value of the measurements from a void region in the sample.

TABLE II
Comparison of the Root Mean Square Error ($\times 10^{-4}\text{nm}^{-1}$) of the Reconstruction With Respect to the Original Phantom When Varying T and δ. A Value of $T = 3$ and $\delta = 0.5$ Produces the Lowest RMSE Reconstruction

T \ δ	0.1	0.5	1
1	4.50	4.42	4.44
3	4.40	**4.31**	4.60
5	4.33	5.09	5.14
20	5.06	5.06	5.06

A. Simulated Data Set

We use a 3-D cubic phantom containing spheres of varying radii with an attenuation coefficient of $7.45 \times 10^{-3} \text{ nm}^{-1}$ to generate the simulated data set. The phantom has a dimension of 256 nm × 512 nm × 512 nm ($z - x - y$ respectively). It is projected at 36 tilts in the range of $-70°$ to $+70°$ in steps of $4°$ about the y axis with a dosage $\lambda_{D,k} = 1865$ counts using the Beer's law model with Gaussian noise having variance equal to the mean of the signal. The value of σ is set to 1. At certain tilts the attenuation of a fraction of the spheres are adjusted to simulate Bragg scatter like effects (Fig. 5) as in a real data set.

Fig. 6 (a) and (b) shows a single $x - z$ and $x - y$ cross-section from the original phantom. Fig. 6 (c)–(j) shows the corresponding cross-section from the reconstructed volume using the different algorithms. The FBP reconstruction [Fig. 6(c) and (d)] has strong streaking artifacts in the $x - z$ cross section and noise in the $x - y$ cross section. The conventional MBIR [Fig. 6(e) and (f)] shows prominent streaking artifacts in the $x - z$ cross section even though there are fewer artifacts than in FBP. Furthermore, there is some underestimation at the center of the spherical particles. However MBIR with anomaly modeling (MBIR-AM) [Fig. 6(g) and (h)] produces a reconstruction which effectively suppresses these artifacts. In the $x - y$ cross section, we notice that the MBIR reconstructions are less noisy as compared to FBP and that the anomaly modeling significantly improves the reconstruction. Next, we evaluate the performance of the proposed MBIR algorithm with parameter estimation (MBIR-AM-PE). Fig. 6 (i) and (j) show that the MBIR-AM-PE can accurately reconstruct the 3D volume suppressing the artifacts despite the unknown nuisance parameters. The value of σ upon termination of the algorithm is 1.770. We note that this value is not the final converged value of the parameter since our stopping criteria depends only on the relative change in voxel values. However we still get a good reconstruction at this termination point. In addition to the qualitative

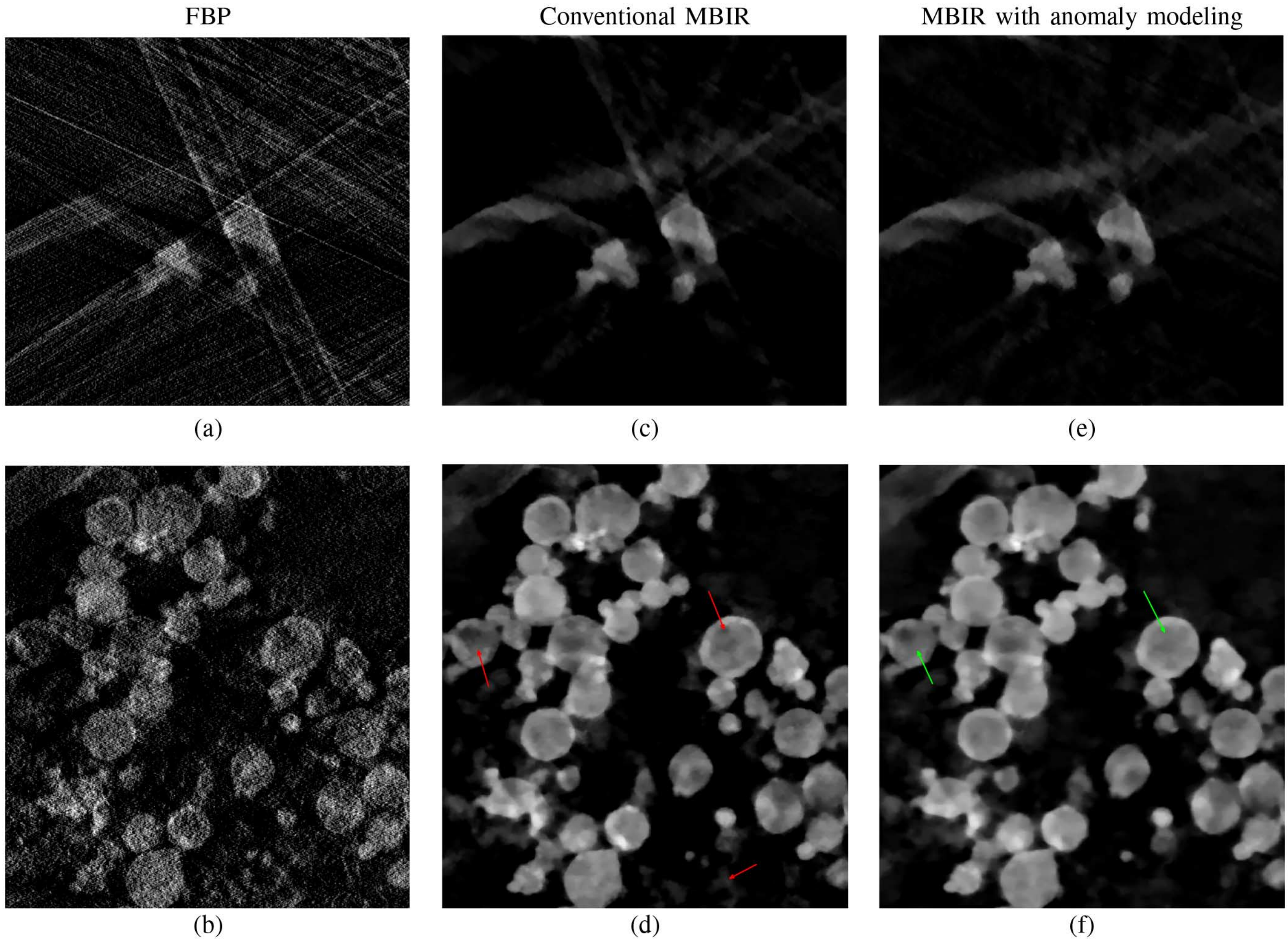

Fig. 9. A single $x-z$ and $x-y$ cross-section reconstructed using different algorithms from a BF-TEM data set of Aluminum sphere nanoparticles. The horizontal direction represents the x axis. The FBP reconstruction (a)-(b) has very strong streaking artifacts in the $x-z$ cross-section, and noise in the $x-y$ cross-section suggesting why it has been avoided for BF-ET. The MBIR algorithm with the anomaly modeling and parameter estimation ($T = 3$ and $\delta = 0.5$) (e)-(f) is superior to the conventional MBIR (c)-(d), suppressing the streaking artifacts seen in (c). In the case of MBIR, the circular cross section of the spherical particles are clearly visible compared to FBP. All images are scaled in the range of $0 - 6.0 \times 10^{-3}\ \mathrm{nm}^{-1}$.

improvements shown in Fig. 6, Table I shows that MBIR with the anomaly modeling (MBIR-AM and MBIR-AM-PE) significantly improves the quantitative accuracy of the reconstruction compared to FBP as well as conventional MBIR. The wall-clock time taken for the MBIR-AM-PE reconstruction ($256 \times 512 \times 512$ voxels) using a node with two 2.60 GHZ Intel Xeon-E5s (total of 16 cores) was approximately 11 minutes.

Finally, we study the sensitivity of the MBIR reconstructions to the choice of parameters T and δ. Fig. 7 shows an $x-z$ cross section from the reconstructions for different values of T when $\delta = 0.5$. Notice that as T increases, streak artifacts start to appear in the reconstruction. This is because some of the anomalous measurements are misclassified. Fig. 8 shows the binary classifier mask, b', corresponding to three successive tilts from simulated data upon completion of the reconstruction. This variable indicates which measurements are classified as anomalous based on the generalized Huber function used for the reconstruction. Notice that when $T = 1$, several non-anomalous measurements are classified as anomalous (false alarms). When $T = 5$ the algorithm misses certain anomalies. When $T = 20$, all the measurements are classified as non-anomalous leading to large errors in the reconstruction. A value of $T = 3$ provides a good tradeoff and is intuitively appealing because this implies that if the data fit error for a measurement is less than 3 times the standard deviation of the noise, then that measurement is non-anomalous. Thus the trade off between false positives and missed detection of anomalies can be varied via the parameter T in the algorithm. Table II shows the RMSE when we vary δ for the different values of T. The value of δ controls the influence of the anomalous measurements on the reconstruction. A value of δ close to 0 implies the anomalous measurements are weighted less in the cost function, while $\delta = 1$ implies the anomalies are weighted significantly. For this particular simulation, we get the lowest RMSE for the $T = 3$ and $\delta = 0.5$ case.

B. Real Data Set

In order to evaluate our approach on real data, we use a data set of approximately 700 nm thick crystalline aluminum

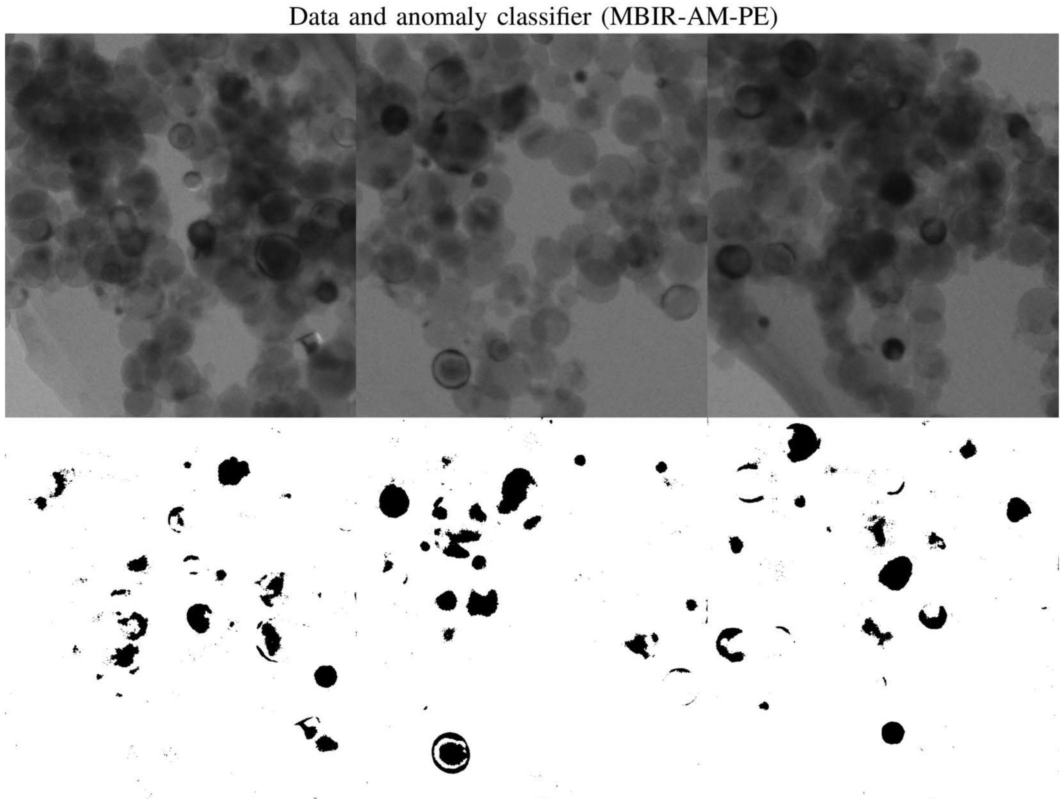

Fig. 10. Data and corresponding anomaly classifier upon termination of the algorithm for the real data set corresponding to three different tilts. The white region in the classifier correspond to non-anomalous measurements and the black regions indicate an anomaly. While the classifier selects certain non-anomalous regions notice that the regions in the data with anomalies are accurately classified by the algorithm.

nanoparticles in a carbon support. We used a FEI Titan TEM with a 300 kV accelerating voltage, and a spot size[1] of 5. The exposure time was set to 1 second, magnification was set to 100 kX, the frame size set to 2048×2048, with a pixel size of 0.83 nm $\times\, 0.83$ nm. The detector used was a CCD with a 30 μm objective aperture resulting in a detector which captures electron scattered in the $0 - 15$ mrad range. The BF-TEM data consists of 33 tilts in the range of $-70°$ to $+70°$. We use a ≈ 580 nm $\times$ 580 nm section of the projection images for reconstruction. 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 [12]. We reconstructed the data set using our algorithm (MBIR-AM-PE), FBP and conventional MBIR without anomaly modeling. All reconstructions are performed with voxels of size 0.83 nm $\times$ 0.83 nm $\times$ 0.83 nm.

Fig. 9 (a) and (b) show an $x - z$ and $x - y$ cross-section reconstructed from the data using FBP. The reconstruction has strong streaking artifacts in the $x - z$ plane and noise in the $x - y$ plane similar to the simulated data set. The reconstruction using the conventional MBIR algorithm [Fig. 9 (c) and (d)], also has streaking artifacts in the $x - z$ plane that are similar to those in the simulated data set of Fig. 6. However,

[1]The spot size is a manufacturer dependent unit-less parameter that refers to the size of the condenser aperture and controls the electron flux on the sample.

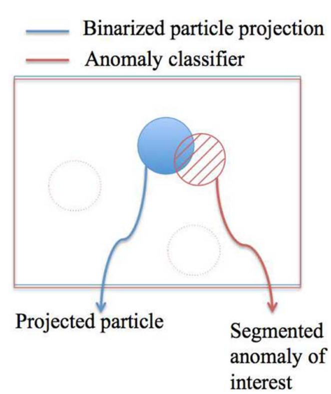

Fig. 11. Illustration of calculation of the similarity score between a certain anomaly and the projection of a given segmented particle in the 3D volume. The projection is binarized and the score is then computed as the extent of overlap between the anomaly and the projection using (22).

the conventional MBIR result also significantly reduces streaking artifacts as compared to FBP. This is likely do to the fact that MBIR reduces the weighting of the highly attenuated projections corresponding to measurements with anomalous Bragg scatter. Fig. 9 (e) shows that using the anomaly modeling and parameter estimation reduces streaking in the $x - z$ plane. The arrows in Fig. 9 (d) and (f) indicate regions where the MBIR with anomaly modeling reduces the under-estimation as well as other artifacts in the $x - y$ cross-section compared to the conventional MBIR.

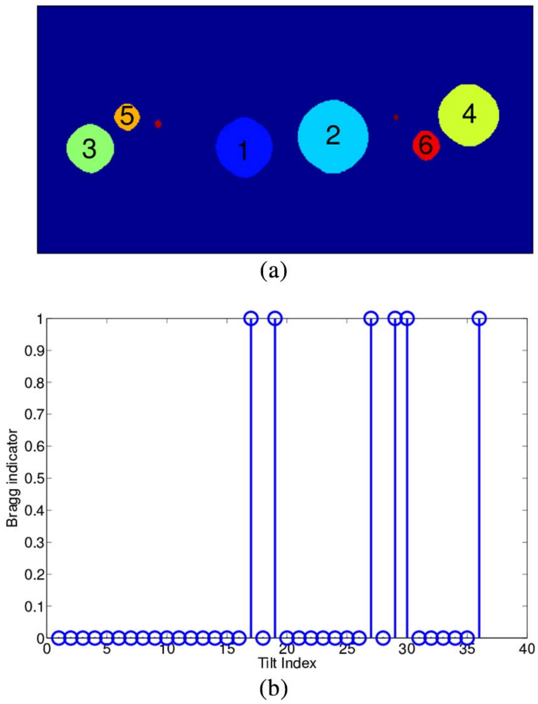

Fig. 12. Illustration of Bragg feature identification for the simulated data set. (a) shows the output of segmentation from a single slice of the reconstruction. The particle labeled 2 was simulated to be in the Bragg condition at tilts indexed by 17, 19, 27, 29, 30 and 36. (b) shows the estimated Bragg feature vector for the particle labeled 2 using the proposed algorithm. In this case the estimated Bragg feature vector matches the ground truth indicating that the Bragg condition can be accurately identified.

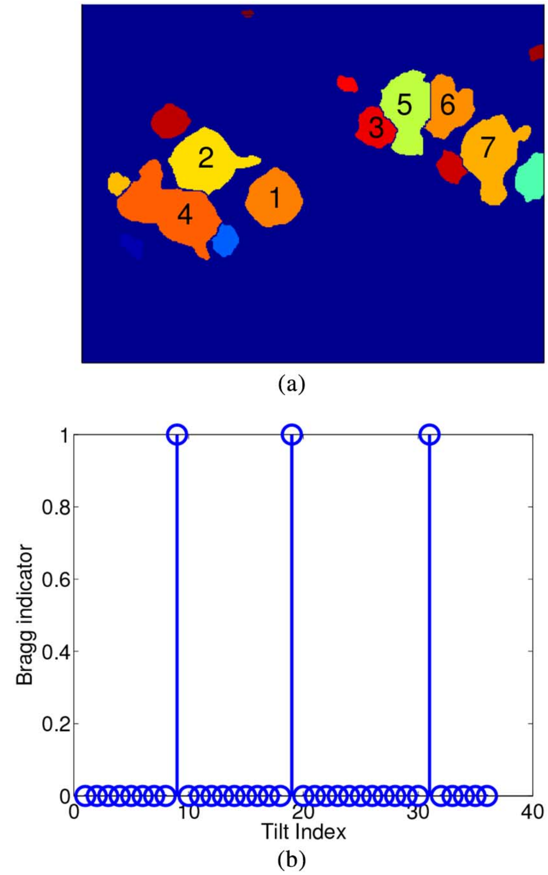

Fig. 13. Illustration of Bragg feature identification for a single particle in a real data set. (a) shows the output of segmentation from a single slice of the reconstruction. (b) shows the Bragg feature vector for the particle labeled (1). While we do not have ground truth for this, we visually observed that the Bragg feature extracted matches what can be seen in the acquired tilt series.

The wall-clock time taken for the proposed MBIR reconstruction ($844 \times 4516 \times 1008$ voxels) using a node with two 2.60 GHZ Intel Xeon-E5s (total of 16 cores) was approximately 9 hours and 40 minutes.

Fig. 10 shows the binary classifier mask b' along with 3 successive tilts from the real data upon termination of the reconstruction algorithm (MBIR-AM-PE). Notice that most of the anomalous measurements are successfully identified by the generalized Huber function at the end of the reconstruction. Similar to the simulated data set a few of the noisy measurements are also classified as anomalous but this does not effect the final quality of the reconstruction significantly.

C. Bragg Feature Extraction

While the particles that undergo Bragg diffraction in a given tilt result in anomalous measurements, the Bragg scatter event contains potentially useful information about the crystal structure and the orientation of the particles. For this reason it is advantageous to correlate the anomalous Bragg event identified in the acquired data with the particle in which it occurred, to produce a Bragg feature vector for each particle.

We will use the binary classifier mask, b', produced by our algorithm along with the reconstructed volume to associate a particle in the volume to an identified anomaly. In order to extract the Bragg feature vector for each particle we apply the following algorithm :

1) Segment the reconstructed volume into individual particles. We use a fixed threshold for segmentation followed by a watershed transformation [36] to separate the fused particles.
2) Identify the connected components (CC) of the anomaly classifier.
3) For each identified CC in the anomaly classifier, project each particle, binarize the projection and find its similarity with that CC. Fig. 11 illustrates how to find the similarity between the binarized projection of a single particle and a particular CC anomaly at a given tilt. If p_k is a binarized version of the projection of a given particle at tilt k and b_k is the binary anomaly classifier with the relevant CC segmented out, we define the similarity score as

$$S_k = 1 - \frac{\|p_k^t \bar{b}_k\|_1 + \|\bar{p}_k^t b_k\|_1}{\|p_k\|_1 + \|b_k\|_1} \tag{22}$$

where $\bar{b}_k$ and $\bar{p}_k$ refers to the binary complement operator and $\|.\|_1$ is the l_1 norm.
4) For a given CC anomaly (at tilt k), find the particle that has the maximum similarity score with it. If this score is higher than a threshold, we label that particle as being in the Bragg condition at tilt k.

In this manner we can associate a binary Bragg feature vector with each segmented particle from the MBIR reconstruction. In order to test our algorithm we apply it to the simulated data set as well as the real data presented earlier. Fig. 12 (a) shows the result of segmentation from a single reconstructed slice of the MBIR-AM-PE reconstruction in IV-A. In our simulation we had set the particle labeled 2 to be in the Bragg condition at tilts indexed by $17, 19, 27, 29, 30$ and 36. Fig. 12 (b) shows the estimated Bragg feature vector for the particle labeled 2 by using the above algorithm. We observe that this matches the ground truth, illustrating the potential of the proposed technique.

Fig. 13 (a) and (b) show a similar result from the real data set. Note that in this case segmentation of the particles are very challenging. However for the particle labeled 1 we are still able to recover the Bragg feature vector and this matches our visual observation from the tilt series data. The results of Fig. 13 show that it is possible to extract potentially useful information about when a single particle is in the Bragg scattering condition. However, the method depends on the assumptions that the volume can be accurately segmented into individual particles corresponding to single crystal orientations.

V. Conclusion

In this paper we presented a MBIR algorithm for BF-ET which can significantly decrease the artifacts in the reconstruction due to anomalies such as Bragg scatter. Our method works by modeling the image formation and the sample being imaged to formulate a cost function that lowers the influence of measurements that do not fit the model accurately. Results on simulated and real data demonstrate that our method can effectively suppress the artifacts due to the anomalies, producing qualitatively and quantitatively accurate reconstructions. We also proposed a simple method for extracting a Bragg feature vector for each particle in a volume that contains potentially useful information about crystal orientation.

Acknowledgments

This work was supported by an AFOSR/MURI grant #FA9550-12-1-0458, by UES Inc. under the Broad Spectrum Engineered Materials contract, and by the Electronic Imaging component of the ICMD program of the Materials and Manufacturing Directorate of the Air Force Research Laboratory, Andrew Rosenberger, program manager.

Appendix A
Composition Property of Surrogate Functions

Theorem 1: Composition property - Let $q(z;z')$ be a surrogate function for the minimization of $t(z)$ on $\mathcal{A} \subset \mathbb{R}^N$; and let $h : \mathcal{A} \to \mathcal{A}$. Then define

$$\tilde{t}(z) \triangleq t(h(z))$$
$$\tilde{q}(z;z') \triangleq q(h(z);h(z')).$$

Then $\tilde{q}(z;z')$ is a surrogate function for $\tilde{t}(z)$.

Proof: The theorem can be proved by verifying the sufficiency coditions for surrogate functions in (8). Notice that at $z = z'$, $\tilde{q}(z';z') = \tilde{t}(z')$ because q is a surrogate function to t. Furthermore, for any $z \subset \mathcal{A}$, $\tilde{q}(z;z') \geq \tilde{t}(z)$ by the construction of q i.e., $q(z;z') \geq t(z)$ for $z \subset \mathcal{A}$. Therefore the compostion is still a surrogate function because it satisfies the sufficiency conditions in (8). ■

References

[1] M. Bárcena and A. J. Koster, "Electron tomography in life science," *Semin. Cell Develop. Biol.*, vol. 20, no. 8, pp. 920–930, 2009.
[2] A. Sousa, M. Hohmann-Marriott, A. Azari, G. Zhang, and R. Leapman, "BF STEM tomography for improved 3D imaging of thick biological sections," *Microsc. Microanal.*, vol. 15, pp. 572–573, 2009.
[3] J. Frank, *Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. New York, NY, USA: Springer, 2008.
[4] R. J. Spontak, M. C. Williams, and D. A. Agard, "Three-dimensional study of cylindrical morphology in a styrene-butadiene-styrene block copolymer," *Polymer*, vol. 29, no. 3, pp. 387–395, 1988.
[5] A. Koster *et al.*, "Development and application of 3-dimensional transmission electron microscopy (3D-TEM) for the characterization of metal-zeolite catalyst systems," in *Proc. 12th Int. Congr. Catal. (ICC)*, 2000, vol. 130, pp. 329–334.
[6] M. Bar Sadan *et al.*, "Toward atomic-scale bright-field electron tomography for the study of fullerene-like nanostructures," *Nano Lett.*, vol. 8, no. 3, pp. 891–896, 2008.
[7] P. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: The development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, no. 34, pp. 413–431, 2003.
[8] S. Bals, C. F. Kisielowski, M. Croitoru, and G. V. Tendeloo, "Annular dark field tomography in TEM," *Microsc. Microanal.*, vol. 11, pp. 2118–2119, 2005.
[9] P. Ercius, M. Weyland, D. A. Muller, and L. M. Gignac, "Three-dimensional imaging of nanovoids in copper interconnects using incoherent bright field tomography," *Appl. Phys. Lett.*, vol. 88, no. 24, p. 243116, 2006.
[10] M. De Graef, *Introduction to Conventional Transmission Electron Microscopy*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[11] 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.
[12] S. Venkatakrishnan *et al.*, "A model based iterative reconstruction algorithm for high angle annular dark field—Scanning transmission electron microscope (HAADF-STEM) tomography," *IEEE Trans. Image Process.*, vol. 22, no. 11, Nov. 2013.
[13] J. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 13, no. 2, pp. 290–300, Jun. 1994.
[14] D. Boas *et al.*, "Imaging the body with diffuse optical tomography," *IEEE Signal Proc. Mag.*, vol. 18, no. 6, pp. 57–75, Nov. 2001.
[15] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, p. R541, 2006.
[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 Int. Symp. Biomed. Imag. (ISBI)*, May 2012, pp. 1012–1015.
[17] 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. Standards Technol.*, vol. 111, no. 6, pp. 411–417, Nov. 2006.
[18] S. V. Venkatakrishnan, L. F. Drummy, M. De Graef, J. P. Simmons, and C. A. Bouman, "Model based iterative reconstruction for bright field electron tomography," *Proc. SPIE Comput. Imag. XI,* vol. 8657, pp. 86 570A–86 570A–12, Feb. 14, 2013.
[19] J. Zheng, S. Saquib, K. Sauer, and C. Bouman, "Parallelizable Bayesian tomography algorithms with rapid, guaranteed convergence," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1745–1759, Oct. 2000.
[20] H. Erdogan and J. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 801–814, Sep. 1999.

[21] A. Mohammad-Djafari, "Joint estimation of parameters and hyperparameters in a Bayesian approach of solving inverse problems," in *Proc. Int. Conf. Image Process.*, Sep. 1996, vol. 1, pp. 473–476.

[22] 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.

[23] R. Zhang, J. Thibault, C. Bouman, K. Sauer, and J. Hsieh, "Model-based iterative reconstruction for dual-energy X-ray CT using a joint quadratic likelihood model," *IEEE Trans. Med. Imag.*, vol. 33, no. 1, pp. 117–134, Jan. 2014.

[24] K. Sauer and C. Bouman, "Bayesian estimation of transmission tomograms using segmentation based optimization," *IEEE Trans. Nucl. Sci.*, vol. 39, pp. 1144–1152, Aug. 1992.

[25] Z. H. Levine, "Theory of bright-field scanning transmission electron microscopy for tomography," *J. Appl. Phys.*, vol. 97, no. 3, p. 033101, 2005.

[26] D. Williams and C. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science. Diffraction. II.* New York, NY, USA: Springer.

[27] A. Blake, "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 1, pp. 2–12, Jan. 1989.

[28] R. L. Stevenson and E. J. Delp, "Fitting curves with discontinuities," in *Proc. 1st IEEE Int. Workshop Robust Comput. Vis.*, Oct. 1990, pp. 127–136.

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

[30] M. Jacobson and J. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2411–2422, Oct. 2007.

[31] M. Razaviyayn, M. Hong, and Z. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optimiz.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[32] 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, 2007.

[33] 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 Form. X-Ray Comput. Tomogr.*, Salt Lake City, UT, USA, Jun. 2012, pp. 297–300.

[34] 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.

[35] 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.

[36] F. Meyer, "Topographic distance and watershed lines," *Signal Process.*, vol. 38, no. 1, pp. 113–125, Jul. 1994.

S. V. Venkatakrishnan (S'12) received the BTech degree in electronics and communication engineering from the National Institute of Technology Tiruchirappalli, India, in 2007, and the MS and PhD degrees in electrical and computer engineering from Purdue University, West Lafayette, Indiana, in 2009 and 2014, respectively. From 2009 to 2010, he worked as a research and development engineer with Baker Hughes Inc., Houston, Texas, on logging while drilling–imaging magnetometers. He is currently a post-doctoral fellow with the Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, working on algorithms for X-ray imaging. His research interests include statistical information processing, inverse problems, computational imaging, and machine learning.

Lawrence F. Drummy received the BS degree in physics from Rensselaer Polytechnic Institute and the PhD degree in imaging of defect structures in organic molecular semiconductor thin films using electron microscopy from the Department of Materials Science and Engineering, University of Michigan, in 2003. He is currently a materials engineer with the Soft Matter Materials Branch, Functional Materials Division, Materials and Manufacturing Directorate, Air Force Research Laboratory, Dayton, Ohio. From 2006 to 2012, he was a materials scientist with UES Inc. and from 2003 to 2006, he was a National Research Council postdoctoral associate at AFRL. 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. His 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 was the recipient of the Arthur K. Doolittle Award from the American Chemical Society Polymer Materials Science and Engineering Division, Air Force Office of Scientific Research Star Team, the Rensselaer Medal Award, and the Padden Award from the Division of Polymer Physics of the American Physical Society.

Michael A. Jackson received the BS degree in materials science and engineering from Wright State University, in 1997. From 1997 to 1999, he worked for Pratt & Whitney Aircraft Engines as a materials engineer, where he started applying computer methods to small materials science problems. He is currently the owner of BlueQuartz Software, Springboro, Ohio.

Marc De Graef received the BS and MS degrees in physics from the University of Antwerp, Belgium, in 1983, and the Ph.D. degree in physics from the Catholic University of Leuven, Belgium, in 1989. He then spent three and a half years as a post-doctoral researcher in the Materials Department, University of California at Santa Barbara before joining Carnegie Mellon, in 1993. He is currently the professor and codirector of the J. Earle and Mary Roberts Materials Characterization Laboratory. His 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 the recipient of the 2012 Educator Award from The Minerals, Metals, and Materials Society.

Jeff Simmons (M'96) received the BS degree in metallurgical engineering from the New Mexico Institute of Mining and Technology, in 1983, and the Master's of Engineering in metallurgical engineering and materials science and PhD degree in materials science and engineering from Carnegie Mellon University, in 1985 and 1992, respectively. He is a research scientist with the Structural Materials Division, Materials and Manufacturing Directorate, Air Force Research Laboratory. He has advised and served on Ph.D. defense committees of Materials Science and of Signal Processing graduate students. He has managed numerous external contracts that involved technologies now referred to as integrated computational materials engineering (ICME).

Charles A. Bouman (S'86–M'89–SM'97–F'01) received the BSEE degree from the University of Pennsylvania, in 1981, the MS degree from the University of California at Berkeley, in 1982, and the PhD degree in electrical engineering from Princeton University, in 1989. From 1982 to 1985, he was a full staff member with MIT Lincoln Laboratory. He joined the faculty of Purdue University in 1989, where he is currently the Michael J. and Katherine R. Birck professor of Electrical and Computer Engineering. He also holds a courtesy appointment with the School of Biomedical Engineering and is a codirector of Purdues

Magnetic Resonance Imaging Facility, Purdues Research Park. His research interests include the use of 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 (AIMBE), the Society for Imaging Science and Technology (IS&T), and the SPIE professional society. 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 also been cochair of the 2006 SPIE/IS&T Symposium on Electronic Imaging, cochair of the SPIE/IS&T conferences on Visual Communications and Image Processing 2000 (VCIP), a vice president of Publications and a member of the Board of Directors for the IS&T Society, and he is the founder and cochair of the SPIE/IS&T conference on Computational Imaging He is the recipient of the IS&Ts Raymond C. Bowman Award for outstanding contributions to digital imaging education and research, has been a Purdue University Faculty Scholar, and the recipient of the College of Engineering Engagement/Service Award and Team Award.