Model Based Iterative Reconstruction for Bright Field Electron Tomography

Singanallur V. Venkatakrishnan[a], Lawrence F. Drummy[b], Marc De Graef[c], Jeff P. Simmons[b], and Charles A. Bouman[a]

[a]Purdue University, West Lafayette, IN
[b]Air Force Research Lab, Dayton, OH
[c]Carnegie Mellon University, Pittsburgh, PA

ABSTRACT

Bright Field (BF) electron tomography (ET) has been widely used in the life sciences to characterize biological specimens in 3D. While BF-ET is the dominant modality in the life sciences it has been generally avoided in the physical sciences due to anomalous measurements in the data due to a phenomenon called "Bragg scatter" - visible when crystalline samples are imaged. These measurements cause undesirable artifacts in the reconstruction when the typical algorithms such as Filtered Back Projection (FBP) and Simultaneous Iterative Reconstruction Technique (SIRT) are applied to the data. Model based iterative reconstruction (MBIR) provides a powerful framework for tomographic reconstruction that incorporates a model for data acquisition, noise in the measurement and a model for the object to obtain reconstructions that are qualitatively superior and quantitatively accurate. In this paper we present a novel MBIR algorithm for BF-ET which accounts for the presence of anomalous measurements from Bragg scatter in the data during the iterative reconstruction. Our method accounts for the anomalies by formulating the reconstruction as minimizing a cost function which rejects measurements that deviate significantly from the typical Beer's law model widely assumed for BF-ET. Results on simulated as well as real data show that our method can dramatically improve the reconstructions compared to FBP and MBIR without anomaly rejection, suppressing the artifacts due to the Bragg anomalies.

1. INTRODUCTION

Bright Field (BF) electron tomography (ET) has been widely used in the life sciences to characterize biological specimens in in 3D.[1] BF-ET typically involves acquiring microscope images (of transmitted electrons) corresponding to various tilts of a sample, and using an algorithm on the acquired "tilt-series" to reconstruct the attenuation coefficient of 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 tomography modality.

While BF-ET is the dominant modality in the life sciences it has been generally avoided in the physical sciences due to contrast reversals[2] from Bragg scatter in crystalline samples. 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 leading to an anomalous measurement uncharacteristic of attenuation due to thickness alone. The presence of Bragg anomalies in the data can result in artifacts since typical tomographic reconstruction algorithms (like FBP and SIRT[3]) do not account for these effects.

Model based iterative reconstruction (MBIR) provides a powerful framework for tomographic reconstruction that incorporates a model for data acquisition, measurement noise and for the object to obtain reconstructions that are qualitatively superior and quantitatively accurate.[4–6] While Levine[7] has applied MBIR to BF-ET in

---

Further author information:
S.V. Venkatakrishnan: svenkata@purdue.edu
C.A. Bouman: bouman@purdue.edu
J.P. Simmons: Jeff.Simmons@wpafb.af.mil
L.F. Drummy: Lawrence.Drummy@wpafb.af.mil
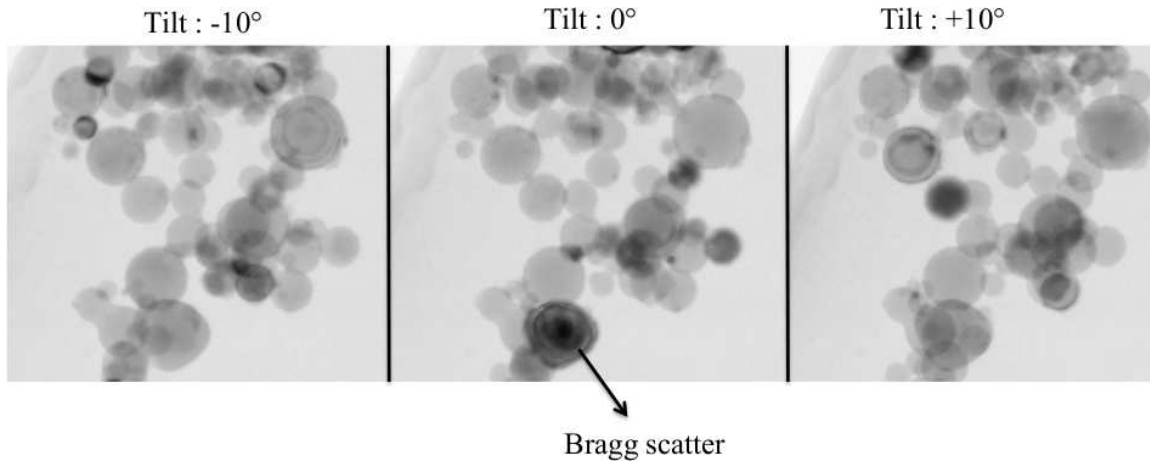M. De Graef: mdg@andrew.cmu.edu

Figure 1. Illustration of the effect of Bragg scatter on a real 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 the orientation of the crystals, electrons are scattered away from the BF detector leading to fewer electrons being collected.

the case of thick specimens and shown that it can improve the quality of the reconstructions, his work deals with cases where there is no anomalous Bragg scatter in the measurement.

In this paper, we present a MBIR algorithm for BF-ET which identifies and rejects the anomalous measurements in the data during the reconstruction. We use a Beer's law forward model and combine it with a prior model for the material to formulate the reconstruction as minimizing a cost function. The cost function is designed so that it rejects the measurements that deviate significantly from the assumed forward model due to Bragg scatter. We then develop a fast algorithm to minimize the cost function. Our algorithm, which is based on the iterative coordinate descent (ICD), works by constructing a substitute to the original cost[4] at every point, and minimizing this new function. This operation significantly lowers the computational complexity of the optimization and speeds up the overall convergence of the algorithm. Intuitively, our method starts with a initial reconstruction, forward projects and compares it with the true measurements and in case the deviation is large, classifies those measurements as anomalous. Using the new set of non-anomalous measurements, the sample is reconstructed and we repeat the process. Thus as the reconstruction progresses, measurements are constantly being monitored and the anomalous ones are rejected.

We apply our method to simulated data sets with Bragg scatter like anomalies. Results show that our method can significantly improve the reconstructions compared to FBP and MBIR without anomaly rejection, suppressing the artifacts that arise due to the anomalous measurements. We also apply our method to a real BF data set and demonstrate that it can reduce the artifacts compared to FBP and MBIR without anomaly rejection.

The organization of the rest of the paper is as follows. In section 2 we introduce a statistical model for the measurement, combine it with a prior model for the material and formulate the MBIR cost function which accounts for anomalous Bragg measurements. In section 3 we develop an efficient algorithm to minimize the cost function. In section 4 we present results from a simulated data set, followed by results from a real data set. Finally in section 5 we draw our conclusions.

2. MEASUREMENT MODEL AND COST FORMULATION

The goal of BF-ET is to reconstruct the attenuation coefficient at every point in the sample. An electron beam is focused on the material and the electrons that are transmitted through the sample 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 which can be used for tomographic reconstruction of the attenuation coefficients.

The reconstruction in the MBIR framework is typically given by the joint-MAP[8] estimate

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

where g is the vector of measurements, f is the vector of unknown voxels (attenuation coefficients), ϕ is a vector of unknown parameters, $p(g|f, \phi)$ is the likelihood and $p(f)$ is the prior probability for the unknown voxels. Next we derive the above cost function for the case of BF-ET accounting for Bragg scatter in the measurements.

First we present an expression for the likelihood term, assuming there is no Bragg scatter in the measurement. Let λ_i be the electron counts corresponding to the ith measurement and λ_D be the counts that would be measured in the absence of the sample. We model the attenuation of the beam through the material using Beer's law. Thus the projection integral corresponding to the ith measurement is given by $\log\left(\frac{\lambda_D}{\lambda_i}\right)$. There can be cases in which the dosage λ_D is not measured and we can include it as a unknown nuisance parameter in the MBIR framework. If g is a $M \times 1$ vector with $g_i = -\log(\lambda_i)$, f is a $N \times 1$ vector of unknown attenuation coefficients of the material, $d = -\log(\lambda_D)$ is an unknown offset, then using a Taylor series approximation to the likelihood function, formed by assuming λ_i's are independent Poisson random variables,[9] gives

$$
\begin{aligned}
-\log p(g|f, d) &\approx \frac{1}{2}\|g - Af - d\,\mathbb{1}\|_\Lambda^2 + h(g) \\
&= \frac{1}{2}\sum_{i=1}^{M}(g_i - A_{i,*}f - d)^2\,\Lambda_{ii} + h(g)
\end{aligned}
\tag{2}
$$

where A is a $M \times N$ forward projection matrix, Λ is a diagonal matrix with entries Λ_{ii} set to be inversely proportional to the variance of the measurement and $h(.)$ is some function of the data. Ignoring the effect of electronic noise, we set $\Lambda_{ii} = \lambda_i$.[4] We note that our formulation can account for more sophisticated models as introduced in,[10] but in this paper we focus on using the Beer's law as it has been found to be accurate for a certain class of thin ($\leq 1\mu$m) samples.

The above likelihood can be directly used to formulate the MBIR cost function for BF-ET when measurements are consistent with our model. However due to Bragg scatter, some of the measurements are anomalous, i.e. they deviate significantly from the model. Bragg scatter typically results in electrons being scattered away from the direct path resulting in fewer counts than would be expected. Fig. 1 shows an example of three tilts from a BF tilt series with regions having significant Bragg scatter (indicated using an arrow). A precise way of accounting for this would be to model the mechanisms that causes Bragg scatter, but this can be very complicated; and so in this work we account for Bragg anomalies by rejecting those measurements. We use the penalty function

$$
\beta_T(x) = \begin{cases} x^2 & |x| < T \\ T^2 & |x| \geq T \end{cases}
$$

to limit the influence of anomalous measurements. This function (Fig. 2) plays a similar role to the weak-spring potential[11] used to model image priors, where it is used to limit the influence of pixels across an edge. Thus the new likelihood is given by

$$
-\log p(g|f, d) = \frac{1}{2}\sum_{i=1}^{M}\beta_T\left((g_i - A_{i,*}f - d)\left(\sqrt{\Lambda_{ii}}\right)\right) + h(g) \tag{3}
$$

where $A_{i,*}$ is the ith row of the forward projection matrix A, $h(g)$ is a constant independent of f and d. The threshold T can be set so that for the Bragg scattered measurements the ratio of the data fit error, $(g_i - A_{i,*}f - d)$, to the noise standard deviation in the measurement, $(\frac{1}{\sqrt{\Lambda_{ii}}})$, is greater than T. T can be set as a user input. Note that the above likelihood does not correspond to a proper density function since the area under the corresponding density function is unbounded. However this formulation can still be used to compute the joint MAP estimate.
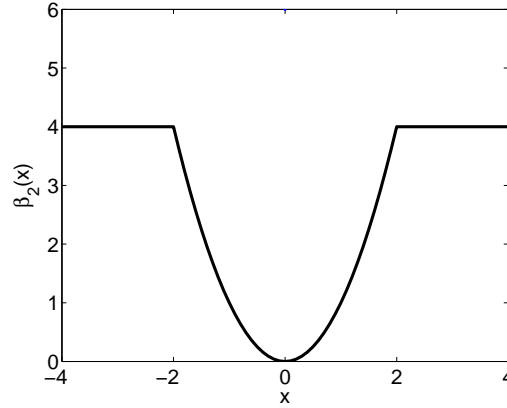
Figure 2. Illustration of the penalty function β_T used for the likelihood term with $T = 2$. Large model mismatch errors are penalized by restricting their influence on the overall cost function.

To model the prior $p(f)$ we use a q-generalized Gaussian Markov random field (qGGMRF)[12] of the form

$$
\begin{aligned}
p(f) &= \frac{1}{Z_f} \exp\{-s(f)\} \\
s(f) &= \sum_{\{j,k\}\in\mathcal{N}} w_{jk}\rho(f_j - f_k) \\
\rho(f_j - f_k) &= \frac{\left|\frac{f_j - f_k}{\sigma_f}\right|^q}{c + \left|\frac{f_j - f_k}{\sigma_f}\right|^{q-p}}
\end{aligned}
\tag{4}
$$

where Z_f is a normalizing constant, $\mathcal{N}$ is the set of pairs of neighboring voxels (e.g. a 26 point neighborhood), and p, q, 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. Typically $1 \le p \le q \le 2$ is used to ensure convexity of the function $\rho(.)$, thereby simplifying the subsequent MAP optimization. We fix $q = 2$ and $c = 0.001$ so that the prior behaves similar to a GGMRF,[13] and it has a bounded second derivative which is a useful property for the subsequent optimization.

Substituting (3) and (4) into (1), the reconstruction is obtained by minimizing the cost

$$
c(f, d) = \frac{1}{2} \sum_{i=1}^{M} \beta_T\left((g_i - A_{i,*}f - d)\sqrt{\Lambda_{ii}}\right) + s(f).
\tag{5}
$$

Alternately we can define

$$
\beta_{T,i}(f, d) = \begin{cases} (g_i - A_{i,*}f - d)^2 \Lambda_{ii} & |(g_i - A_{i,*}f - d)\sqrt{\Lambda_{ii}}| < T \\ T^2 & |(g_i - A_{i,*}f - d)\sqrt{\Lambda_{ii}}| \ge T \end{cases}
$$

so the cost function can be written as

$$
c(f, d) = \frac{1}{2} \sum_{i=1}^{M} \beta_{T,i}(f, d) + s(f)
\tag{6}
$$

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

$$
\left(\hat{f}, \hat{d}\right) \leftarrow \underset{f \ge 0, d}{\operatorname{argmin}} c(f, d)
$$

3. OPTIMIZATION ALGORITHM

The cost function (6) is non-convex, and thus finding the global minimum is difficult. Here we attempt to find a desirable local minimum of the cost. We minimize the cost function in (6) using the ICD[14] algorithm. In ICD we start with a initial value for the variables, and then they are updated one at a time; so that each update lowers the value of the cost function. Minimizing (6) with respect to each variable can be computationally expensive due to complicated form of the likelihood and prior terms, so we can instead construct a substitute to the original function and minimize this new function. A substitute function to the original function is constructed so that it bounds the original function from above, and so that minimizing the substitute function results in a lower value of the original cost.

Our goal is to find a substitute function to the original cost so that it can be easily minimized with respect to each voxel or the unknown parameter d. We find a substitute function for the likelihood terms and the prior terms separately and then sum them to form a single substitute function for the original cost.

We design substitute functions $Q_{T,i}(f, d; f', d')$ for each of $\beta_{T,i}(f, d)$ in (6) at a given point (f', d'). In particular

$$Q_{T,i}(f, d; f', d') = \begin{cases} (g_i - A_{i,*}f - d)^2 \Lambda_{ii} & |(g_i - A_{i,*}f' - d')\sqrt{\Lambda_{ii}}| < T \\ T^2 & |(g_i - A_{i,*}f' - d')\sqrt{\Lambda_{ii}}| \geq T \end{cases} \tag{7}$$

is a substitute function for $\beta_{T,i}(.)$ as shown in Appendix A.

We can find a substitute function for each potential function $\rho(f_j - f_k)$ at the point f' of the form

$$\rho(f_j - f_k; f_j' - f_k') = \frac{a_{jk}}{2}(f_j - f_k)^2 + b_{jk}. \tag{8}$$

Using such a form results in a simple closed form updates for the voxels during the optimization. The values of a_{jk} and b_{jk} can be derived as shown in Appendix B and are given by

$$a_{jk} = \begin{cases} \frac{\rho'(f_j' - f_k')}{(f_j' - f_k')} & f_j' \neq f_k' \\ \rho''(0) & f_j' = f_k' \end{cases} \tag{9}$$

$$b_{jk} = \rho(f_j' - f_k') - \frac{a_{jk}}{2}(f_j' - f_k')^2 \tag{10}$$

Thus a substitute function to $s(f)$ at $f = f'$ is

$$s(f; f') = \sum_{\{j,k\} \in \mathcal{N}} w_{jk}\rho(f_j - f_k; f_j' - f_k'). \tag{11}$$

3.1 Algorithm

Based on the present value of (f, d) which we denote (f', d') we define the following indicator variable:

$$\tilde{b}_i = \begin{cases} 1 & |(g_i - A_{i,*}f' - d')\sqrt{\Lambda_{ii}}| < T \\ 0 & |(g_i - A_{i,*}f' - d')\sqrt{\Lambda_{ii}}| \geq T \end{cases} \tag{12}$$

Intuitively $\tilde{b}_i$ indicates which measurements are classified as anomalous and which are not, based on the current state of the reconstruction. Using (7) and (11) a substitute function to the original cost (6) at (f', d') is

$$Q_T(f, d; f', d') = \frac{1}{2}\sum_{i=1}^M Q_{T,i}(f, d; f', d') + s(f; f')$$

$$= \frac{1}{2}\sum_{i=1}^M (g_i - A_{i,*}f - d)^2 \Lambda_{ii}\tilde{b}_i + \sum_{\{j,k\} \in \mathcal{N}} w_{jk}\rho(f_j - f_k; f_j' - f_k') + \frac{1}{2}\sum_{i=1}^M (1 - \tilde{b}_i)T^2 \tag{13}$$

function $[\hat{f}, \hat{d}] \leftarrow$ Reconstruct$(g,\ f',\ d',\ R)$

    %**Inputs**: Measurements g, Initial reconstruction f', Initial dosage d', Fraction of entries to reject R

    %**Outputs**: Reconstruction $\hat{f}$ and dosage parameter $\hat{d}$

    $e' = g - Af' - d'\mathbb{1}$

    $r = 0$ ▷ Initial fraction to reject

    $T \leftarrow \infty,\ \tilde{b} \leftarrow \mathbb{1}$ ▷ Initially do not reject any measurement.

    while Stopping criteria is not met **do**

        for each voxel j in random order **do**   ▷ Iterate over all voxels.

$$\tilde{\theta}_2 \leftarrow \sum_{i=1}^{M} (A_{i,j})^2 \Lambda_{ii} \tilde{b}_i$$

$$\tilde{\theta}_1 \leftarrow -\sum_{i=1}^{M} e_i' \Lambda_{ii} \tilde{b}_i (A_{i,j})$$

            for $k \in \mathcal{N}_j$ **do**

                Compute substitute function parameter a_{jk} using (9)

            end for

$$u^* \leftarrow \frac{\displaystyle\sum_{k \in \mathcal{N}_j} w_{jk} a_{jk} f_k' + \tilde{\theta}_2 f_j' - \tilde{\theta}_1}{\displaystyle\sum_{k \in \mathcal{N}_j} w_{jk} a_{jk} + \tilde{\theta}_2}$$

            $f_j \leftarrow \max(u^*, 0)$

            $e' \leftarrow e' - (f_j - f_j') A_{*,j}$

            $f_j' \leftarrow f_j$

            Update $\tilde{b}$ using (12)   ▷ Computing the new Q_T function

        end for

        $d' \leftarrow$ Update d using (16)   ▷ Dosage parameter update

        Update e'

        Update $\tilde{b}$ using (12)

        If $r < R$, then $r = r + R/10$ ▷ Increment the rejection threshold

        Compute new T ▷ Sort the array of $e_i * \sqrt{\Lambda_{ii}}$ and set $T = r^{\text{th}}$ percentile

        Update $\tilde{b}$ using (12)

    end while

    $\hat{f} \leftarrow f,\ \hat{d} \leftarrow d$

end function

Figure 3. MBIR algorithm for BF data with Bragg scatter. The algorithm works by constructing a substitute to the original function based on the current values of the voxels and dosage parameter and minimizing this substitute function with respect to a single variable. The process is then repeated. The algorithm can be efficiently implemented by keeping track of the error sinogram.[15] Further the rejection ratio r is progressively increased till it reaches the target value R to prevent the algorithm from getting stuck in undesirable local minima.

3.1.1 Voxel Update

The voxels are updated one at a time in random order similar to[4] in order to speed up the overall convergence of the algorithm. To minimize with respect to voxel j, we can fix $f_k = f'_k \ \forall k \in \{1, \cdots, M\} \setminus \{j\}$ and $d = d'$ in (13). The cost function we need to minimize is

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

where $\tilde{\Lambda}_{ii} = \Lambda_{ii} \tilde{b}_i$, $\tilde{\theta}_1 = -(e')^t \tilde{\Lambda}(A_{*,j})$, $\tilde{\theta}_2 = (A_{*,j})^t \tilde{\Lambda}(A_{*,j})$, $A_{*,j}$ is the jth column of the forward projection matrix A, $e' = g - Af' - d'\mathbb{1}$, f'_j is the present value of voxel j, and $\mathcal{N}_j$ is the set of all neighbors of voxel j. 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

$$u^* = \frac{\displaystyle\sum_{k \in \mathcal{N}_j} w_{jk} a_{jk} f'_k + \tilde{\theta}_2 f'_j - \tilde{\theta}_1}{\displaystyle\sum_{k \in \mathcal{N}_j} w_{jk} a_{jk} + \tilde{\theta}_2}. \tag{14}$$

Enforcing the positivity constraint, the update for the voxel is

$$\hat{f}_j \leftarrow \max\left(u^*, 0\right) \tag{15}$$

3.1.2 Nuisance Parameter Update

In order to minimize the substitute function with respect to the dosage parameter d, we take the derivative of the substitute function (13) $Q_T(f', d; f', d')$ with respect to d and set it to zero. This gives the optimal update for d as

$$\hat{\alpha} \leftarrow \frac{(e')^t \tilde{\Lambda} \mathbb{1}}{\mathbb{1}^t \tilde{\Lambda} \mathbb{1}}$$
$$\hat{d} \leftarrow d + \hat{\alpha} \tag{16}$$

3.1.3 Multi-resolution Initialization

The optimization can be further sped up using a multi-resolution initialization.[16] In multi-resolution initialization, we perform a reconstruction at a coarser resolution (larger voxel sizes) and use its output to initialize a finer resolution reconstruction. This transfers the computational load to the coarser scale where the optimization can be done quickly due to to the reduced dimensionality of the problem. Since our prior behaves similar to a GGMRF,[13] we adapt the scaling parameter σ_f according to Eq.28 in.[17]

We set the rejection threshold parameter T in the algorithm indirectly, via an user input R, the approximate fraction of the total measurements that the are affected by Bragg scatter. We refer to this as the target rejection rate. Given a value of R, T can be chosen so that RM measurements are not used in the cost function.

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. Further, in order to prevent the algorithm from getting stuck in undesirable local minima, the rejection percentage is gradually increased to the desired target rejection rate (and therefore T is gradually decreased). The MBIR BF-ET algorithm for a single resolution is summarized in Fig. 3.

4. RESULTS

In this section we compare three algorithms for BF-ET - FBP, MBIR without Bragg anomaly correction and MBIR with anomaly correction. While BF-ET has been avoided in the physical sciences because algorithms like FBP are know to be unreliable for this type of data, we include it to indicate how an algorithm designed for the application (MBIR) can outperform a standard algorithm (FBP). FBP is performed in Matlab using the *iradon*
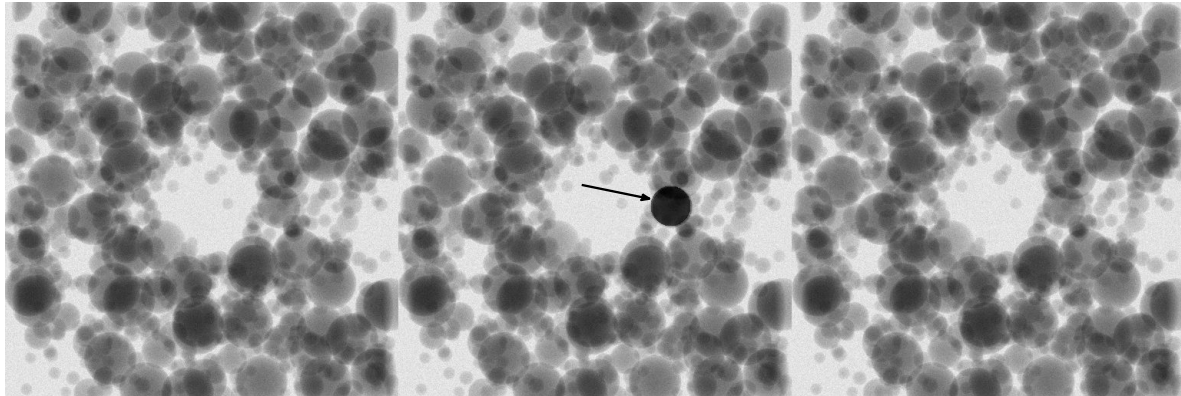
Figure 4. Simulated BF data corresponding to a phantom of spheres for three successive tilts. The figure in the center shows the simulated Bragg scatter obtained by lowering the counts by about 50%. This is an anomaly in the projection data, and if not accounted for, can cause artifacts in the reconstruction.


$x - z$ slice from phantom

Figure 5. A single $x - z$ slice from the phantom data used for qualitative comparison against various reconstruction algorithms.

(a) FBP   (b) MBIR without Bragg correction   (c) MBIR with Bragg correction

Figure 6. Comparison of BF reconstructions on a data set with few tilts having anomalous measurements. (a) shows a single $x - z$ slice from a FBP reconstruction. (b) shows the MBIR reconstruction without Bragg correction. The reconstruction is comparable to the phantom because the fraction of anomalous measurements is relatively low. The strong streaking artifacts are significantly suppressed compared to FBP. (c) shows the reconstruction with Bragg anomaly correction with the rejection threshold set to 5% . The reconstruction is superior to the case in which we apply no correction as well as FBP. All images are scaled in the range of $0 - 7.45 \times 10^{-3}$ nm^{-1}.
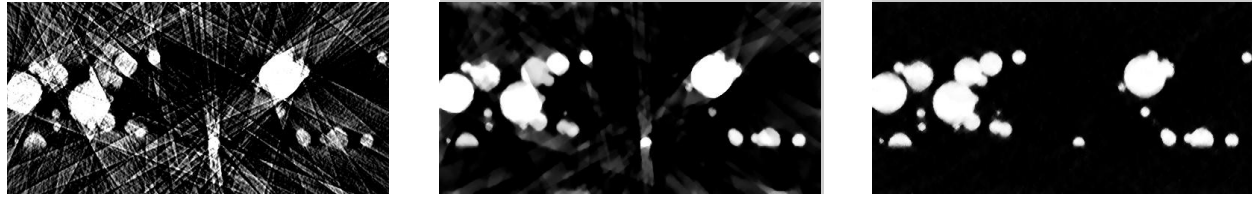
(a) FBP   (b) MBIR without Bragg correction   (c) MBIR with Bragg correction

Figure 7. Comparison of BF reconstructions for a data set with a high percentage of the tilts having Bragg scatter. (a) shows a single $x - z$ slice from a FBP reconstruction. (b) shows the MBIR reconstruction without Bragg correction. The reconstruction has streaks because of the Bragg scatter but much lesser compared to FBP. (c) shows the reconstruction with Bragg anomaly correction with the rejection threshold set to 10%. The method effectively suppresses the artifacts in (a) and (b), and produces a more accurate reconstruction. All images are scaled in the range of $0 - 7.45 \times 10^{-3}$ nm^{-1}.

command. The output is clipped to be positive. For the MBIR reconstructions, σ_f and R are chosen to obtain the best visual quality of reconstruction. The offset d is initialized by taking the mean value of the signal in a void region of the sample.

For the first experiment we use a 3-D phantom containing spheres with an attenuation coefficient of 7.45×10^{-3} nm^{-1}. The sample has a dimensions of 256 nm $\times$ 512 nm $\times$ 512 nm ($z - x - y$ respectively). The phantom is forward projected at 141 tilts in the range of $-70°$ to $+70°$ in steps of $1°$ with a dosage $\lambda_D = 1850$ counts using the Beer's law model with Poisson noise. At certain tilts ($\approx 40\%$) some of the measurements are decreased by 50% to simulate Bragg scatter like effects (Fig. 4). While accurate simulation of Bragg scatter is very complicated, we attempt to demonstrate some artifacts that can occur even in this very simple case and show how our algorithm can handle it. We reconstruct a 3D volume consisting of 12 ($x - z$) slices.

Fig. 5 shows a single $x - z$ slice from the original phantom. Fig. 6 shows the reconstructions of the corresponding $x - z$ slice. The FBP reconstruction (Fig. 6 (a)) has streaking artifacts due to Bragg scatter as well as the absence of a prior model for the object. The MBIR reconstructions (Fig. 6 (b) and (c)) significantly suppress the streaking artifacts compared to FBP. We observe that the MBIR result without Bragg rejection (Fig. 6(b)) is qualitatively comparable to the MBIR with the Bragg rejection (Fig. 6(c)) though it has some streaking artifacts. This suggests that when the fraction of anomalous measurements is low, the algorithms with and without the Bragg rejection produce qualitatively comparable results. However Table 1 shows that MBIR with the Bragg anomaly correction produces a quantitatively more accurate reconstruction.

Fig. 7 shows the $x - z$ slice reconstructed using the different algorithms when we use only a subset of 47 tilts from the phantom data set. Most of these tilts have Bragg anomalies and hence the fraction of anomalous measurements is much higher in this data set. The FBP reconstruction (Fig. 7(a)) has strong streaking artifacts indicating why it has not been used for BF reconstructions in the physical sciences. The MBIR without Bragg anomaly correction (Fig. 7 (b)) shows prominent streaking artifacts in the reconstruction even though it is much lesser than in FBP. However MBIR with anomaly correction (Fig. 7(c)) produces a reconstruction which effectively suppresses these artifacts. Table 1 shows that MBIR with the Bragg anomaly correction significantly improves the quantitative accuracy of the reconstruction compared to FBP as well as MBIR without anomaly correction.

Fig. 8 shows a $x - z$ slice (≈ 581 nm $\times$ 900 nm) reconstructed from a real sample of spherical Aluminum nanoparticles. The BF-TEM data consists of 15 tilts in the range of $-70°$ to $+70°$ in steps of $10°$. The FBP reconstruction (Fig. 8 (a)), has strong streaking artifacts. The reconstruction using the MBIR algorithm with no anomaly correction (Fig. 8 (b)), also has streaking artifacts similar to those in the simulated data set of Fig. 7 but much lesser than in the case of FBP. Fig. 8 (c) shows that using the Bragg anomaly correction can result in reconstructions which have significantly reduced streaking artifacts compared to MBIR without Bragg anomaly rejection (Fig. 8 (b)) as well as FBP (Fig. 8 (a)). The Bragg rejection threshold was set to reject 10% of the data.

Table 1. Comparison of the Root Mean Square Error of the reconstruction with respect to the original phantom for various scenarios. MBIR with Bragg anomaly correction produces quantitatively more accurate reconstructions.

Data Set	Algorithm	Bragg Correction	RMSE (nm^{-1})
Limited Bragg (Fig. 6)	FBP	No	13×10^{-4}
Limited Bragg (Fig. 6)	MBIR	No	5.7×10^{-4}
Limited Bragg (Fig. 6)	MBIR	Yes	4.87×10^{-4}
More Bragg (Fig. 7)	FBP	No	23×10^{-4}
More Bragg (Fig. 7)	MBIR	No	11.95×10^{-4}
More Bragg (Fig. 7)	MBIR	Yes	6.52×10^{-4}
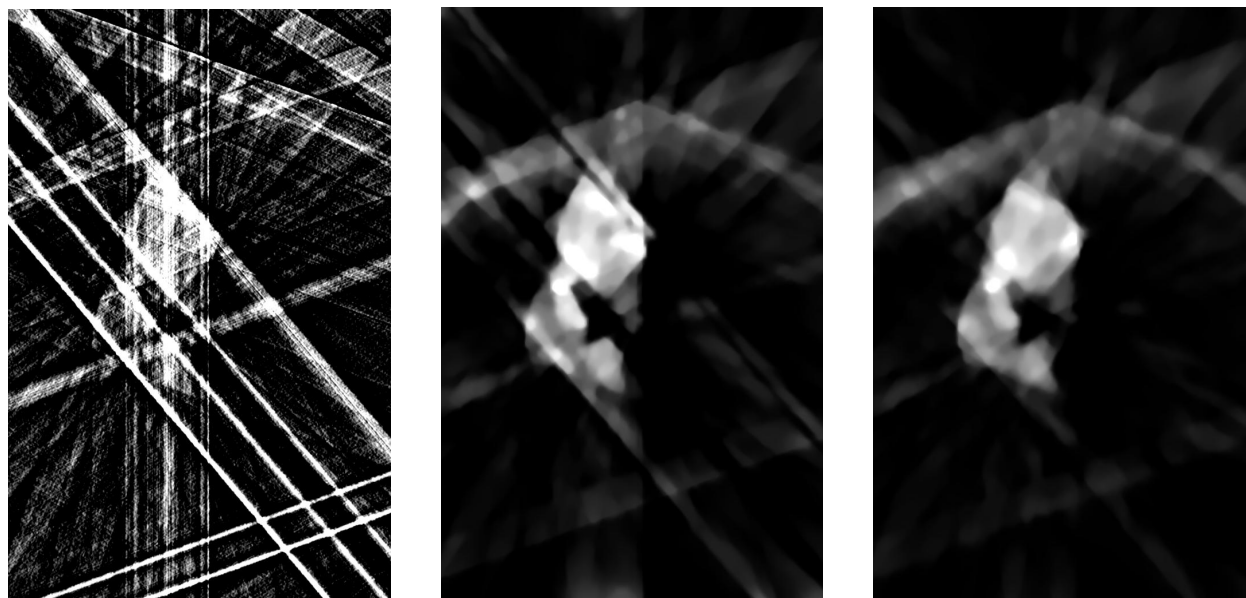
(a) FBP (b) MBIR without Bragg correction (c) MBIR with Bragg correction

Figure 8. A single $x - z$ slice reconstructed from a BF-TEM data set of Aluminum sphere nanoparticles. The FBP reconstruction (a) has very strong streaking artifacts, suggesting why it has been avoided for BF-ET. The MBIR algorithm with the anomaly rejection (c) is superior to the case in which we apply no correction, suppressing the streaking artifacts seen in (b). 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 $0 - 4.0 \times 10^{-3}$ nm^{-1} and the rejection threshold is set to 10%.

5. CONCLUSION

In this work we presented a MBIR algorithm for BF-ET which can significantly decrease the artifacts in the reconstruction due to anomalous Bragg scatter. Our method works by modeling the image formation and the sample being imaged to formulate a cost function which rejects measurements that do not fit the model accurately as a part of the reconstruction. Results on simulated and real data sets demonstrate that our method can effectively suppress the artifacts due to Bragg scatter, producing qualitatively and quantitatively accurate reconstructions.

APPENDIX A. SUBSTITUTE FUNCTION FOR LIKELIHOOD

THEOREM A.1. *Each* $Q_{T,i}(f, d; f', d')$ *is a substitute function to* $\beta_{T,i}(f, d)$.

Proof. Let us define the following functions. $Q : \mathbb{R} \to \mathbb{R}$, $\beta : \mathbb{R} \to \mathbb{R}$ such that

$$\beta(x) = \begin{cases} x^2 & |x| < T \\ T^2 & |x| \geq T \end{cases}$$

$$Q(x; x') = \begin{cases} x^2 & |x'| < T \\ T^2 & |x'| \geq T \end{cases}$$

Then $Q(x; x')$ is a substitute function to $\beta(x)$. If we define a function $h_i : \mathbb{R}^{N+1} \to \mathbb{R}$ such that $h_i(f, d) = (g_i - A_{i,*}f - d)\sqrt{\Lambda_{ii}}$. Then $Q(h_i(f, d); h_i(f', d'))$ is a substitute function to $\beta(h_i(f, d))$ by Property 7.6 in.[18] Thus the theorem is proved by recognizing $Q_{T,i}(f, d; f', d') = Q(h_i(f, d); h_i(f', d'))$ and $\beta_{T,i}(f, d) = \beta(h_i(f, d))$. □

APPENDIX B. SUBSTITUTE FUNCTION FOR PRIOR

In order to find a suitable substitute function to the prior $s(f)$ each of the potential functions $\rho(f_j - f_k)$ can be replaced by a function $\rho(f_j - f_k; f'_j - f'_k)$ which satisfy the following properties[4]

$$\rho(f_j - f_k; f'_j - f'_k) \geq \rho(f_j - f_k) \quad \forall f_j \in \mathbb{R} \tag{17}$$

$$\rho'(f'_j - f'_k; f'_j - f'_k) = \rho'(f'_j - f'_k) \tag{18}$$

where f' is the point of approximation. Intuitively (17) ensures that the substitute function upper bounds the original potential function and (18) 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

$$\rho(f_j - f_k; f'_j - f'_k) = \frac{a_{jk}}{2}(f_j - f_k)^2 + b_{jk} \tag{19}$$

because it results in a simple closed form update for a given voxel. Thus we need to find the values of a_{jk} and b_{jk} which satisfies (17) and (18). Taking the derivative of the substitute function (19) and matching it to the derivative of the original potential function we get

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

To choose b_{jk} 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

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

ACKNOWLEDGMENTS

REFERENCES

[1] Bárcena, M. and Koster, A. J., "Electron tomography in life science," *Seminars in Cell & Developmental Biology* **20**(8), 920 – 930 (2009).

[2] Midgley, P. and Weyland, M., "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy* **96**(34), 413 – 431 (2003).

[3] Kak, A. C. and Slaney, M., [*Principles of Computerized Tomographic Imaging*], Society for Industrial and Applied Mathematics, Philadephia, PA (2001).

[4] Yu, Z., Thibault, J., Bouman, C., Sauer, K., and Hsieh, J., "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," *IEEE Trans. on Image Processing* **20**, 161 –175 (Jan. 2011).

[5] Venkatakrishnan, S., Drummy, L., Jackson, M., De Graef, M., Simmons, J., and Bouman, C., "Bayesian Tomographic Reconstruction for High Angle Annular Dark Field (HAADF) Scanning Transmission Electron Microscopy (STEM)," in [*2012 IEEE Statistical Signal Processing Workshop (SSP) (SSP'12)*], (Aug. 2012).

[6] Fessler, J., "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. on Medical Imaging* **13**, 290 –300 (June 1994).

[7] Levine, Z. H., Kearsley, A. J., and Hagedorn, J. G., "Bayesian tomography for projections with an arbitrary transmission function with an application in electron microscopy," *Journal of Research of the National Institute of Standards and Technology* **111**, 411 –417 (Nov. 2006).

[8] Mohammad-Djafari, A., "Joint estimation of parameters and hyperparameters in a Bayesian approach of solving inverse problems," in [*Image Processing, 1996. Proceedings., International Conference on*], **1**, 473 –476 vol.2 (Sept. 1996).

[9] Bouman, C. and Sauer, K., "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Processing* **5**, 480 –492 (Mar. 1996).

[10] Levine, Z. H., "Theory of bright-field scanning transmission electron microscopy for tomography," *Journal of Applied Physics* **97**(3), 033101 (2005).

[11] Blake, A., "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 2–12 (Jan. 1989).

[12] Thibault, J.-B., Sauer, K. D., Bouman, C. A., and Hsieh, J., "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Medical Physics* **34**(11), 4526–4544 (2007).

[13] Bouman, C. and Sauer, K., "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. on Image Processing* **2**, 296 –310 (July 1993).

[14] Sauer, K. and Bouman, C. A., "A local update strategy for iterative reconstruction from projections," *IEEE Trans. on Signal Processing* **41**, 534–548 (Feb. 1993).

[15] Sauer, K. and Bouman, C., "Bayesian Estimation of Transmission Tomograms Using Segmentation Based Optimization," *IEEE Trans. on Nuclear Science* **39**, 1144–1152 (1992).

[16] Kamasak, M., Bouman, C., Morris, E., and Sauer, K., "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. on Medical Imaging* **24**, 636 –650 (May 2005).

[17] Oh, S., Milstein, A., Bouman, C., and Webb, K., "A general framework for nonlinear multigrid inversion," *IEEE Trans. on Image Processing* **14**, 125 –140 (Jan. 2005).

[18] Bouman, C. A., [*Model Based Image And Signal Processing*] (2012).