WAVELET FRAME BASED ALGORITHM FOR 3D RECONSTRUCTION IN ELECTRON MICROSCOPY

MING LI[†‡], ZHITAO FAN[‡], HUI JI[‡], AND ZUOWEI SHEN[‡]

Abstract. In electron microscopy, 3D reconstruction is one key component in many computerized techniques for solving 3D structures of large protein assemblies using electron microscopy images of particles. Main challenges in 3D reconstruction include very low signal-to-noise ratio and very large scale of data sets involved in the computation. Motivated by the recent advances of sparsity-based regularization in wavelet frame domain for solving various linear inverse problems in imaging science, we proposed a wavelet tight frame based 3D reconstruction approach that exploits the sparsity of the 3D density map in wavelet tight frame system. The proposed approach not only runs efficiently in terms of CPU time, but also requires a much lower memory footprint than existing framelet-based regularization methods. The convergence of the proposed iterative scheme and the functional it minimizes is also examined, together with the connection to existing wavelet frame based regularizations. The numerical experiments showed good performance of the proposed method when being used in two electron microscopy techniques: single particle method and electron tomography.

Key words. wavelet frame, 3D reconstruction, electron microscopy, sparse approximation

AMS subject classifications. 68U10, 65J22, 65T60, 94A08, 90C90

1. Introduction. In the past few decades, with the advance in specimen preparation methods and image processing techniques, electron microscopy (EM) techniques [1, 2, 3, 4, 5, 6] have become indispensable tools for determining the 3D structures of macromolecules, macromolecular complexes and cells. Among all computerized techniques used in EM, the single particle method (SPM) and the electron tomography (ET) are two of the most popular ones. The specimen preparation methods mainly include the negative staining method in which specimens are stained in heavy metal salts and the frozen-hydrated method in which specimens are embedded in vitreous ice. The EM technique using ice-embedded specimens is often called cryo-EM. The SPM includes a set of image processing techniques for building up the 3D structure of a particle by using 2D digitized EM images (projections) of many identical particles. The main modules in the SPM include image alignment and classification; image filtering and contrast transfer function (CTF) correction; and 3D reconstruction (see [4, 5, 7, 6] for more details). These modules are integrated in an iterative framework to iteratively refine the 3D result using the projection matching methods [8]. The ET is a tomography technique to reconstruct a detailed 3D structure of the macromolecular complex or cell from the 2D images collected from different tilted stages of the object. The main image processing techniques in the ET include CTF correction and micrograph alignment; 3D reconstruction; segmentation and 3D classification (see [3, 9, 10] for more details).

The 3D reconstruction problem plays an important role in the EM techniques including both the SPM and the ET. The 3D reconstruction is about reconstructing the 3D density map of the object from multiple EM images. Similar to computerized tomography using parallel beams, both the SPM and the ET reconstruct the 3D density map f of the object from multiple images $\{g_i\}$ with different projection directions $\{\mathrm{d}_i\}$ and different electron beams. The 3D density map and the EM images

[†]LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

[‡]Department of Mathematics, National University of Singapore, Singapore 19076

are related by the following linear system:

$$g = Au + \varepsilon, \tag{1.1}$$

where $g \in \mathbb{R}^M$ denotes the concatenation of multiple EM images, $u \in \mathbb{R}^N$ denotes the 3D density map of the object, $A \in \mathbb{R}^{M\times N}$ denotes the measuring matrix, and ε denotes the noise. The 3D reconstructions in the SPM and the ET have the same model (1.1). The main difference between the SPM and the ET lies in how EM images are collected: the SPM is for imaging many identical biological particles with varying orientations, while the ET is for imaging large biological object such as cells, molecular machines which do not have multiple identical instances, by tilting around one axis practically within $[-70°, +70°]$.

The matrix A is the matrix form of the so-called 3D ray transform after discretization. During the formation of an EM image in a transmission electron microscope, the phase of the electron wave when passing through the object is shifted by $\Phi(r)$ which can be described as the integral of the Coulomb potential distribution $C(r, z)$ within the object along the z-direction of the electron wave's propagation

$$\Phi(r) = \int C(r, z)dz,$$

where $r = [x, y]^\top$. Using the weak-phase object approximation and the contrast transfer theory (see e.g. [4] for more details), the contrast of the EM image is linearly proportional to the projected Coulomb potential $\Phi(r)$. The physical density distribution u retrieved with EM is the Coulomb potential. Thus, an EM image can be approximated by the projection of the density map of the object along some direction, which can be expressed by the 3D ray transform of a function $f \in L_2(\mathbb{R}^3)$

$$Pf(\langle x, \mathrm{d}^\perp\rangle \mathrm{d}^\perp) = \int_{\mathbb{R}^1} f(x + t\mathrm{d})dt, \quad x \in \mathbb{R}^3, \tag{1.2}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in $\mathbb{R}^3$, $\mathrm{d} \in \mathbb{R}^3$ denotes the projection direction associated with the corresponding EM image, and $\mathrm{d}^\perp$ denotes the plane perpendicular to d. The linear system (1.1) is the discretized version of (1.2) where the matrix A in (1.1) is the linear operator obtained by discretizing the integeral in (1.2).

Due to the extremely low signal to noise ratio (SNR) of EM images, reconstructing a 3D density map of the object from the EM images is a very challenging task. An effective 3D reconstruction method requires very sophisticated techniques to suppress the noise during the reconstruction. Another main challenge lies in the computational efficiency of the method. The amount of input data in EM is often enormous in practice. For example, the cryo-EM 3D reconstruction of 50S ribosomal subunit published in [11] uses 82575 images with size 100×100 to obtain a cubic density map with size $100 \times 100 \times 100$. Thus, a desired reconstruction method should be very computational efficient with the following properties: highly parallelized; small memory footprint; and efficient CPU time usage. In this paper, we aim at developing a powerful and efficient 3D reconstruction algorithm for the SPM and the ET. The developed method is not only very robust to noise, but also is very efficient on both memory usage and CPU time usage.

1.1. Motivations. The 3D reconstruction method proposed in this paper is motivated by the recent development of the sparse based methods for image restorations that find sparse solutions by minimizing ℓ_1-norm of wavelet tight frame coefficients.

The tight frame based methods assume that the data for recovery is likely to have a sparse approximation under the wavelet tight frames. For a given signal $u \in \mathbb{R}^N$, let $W \in \mathbb{R}^{L\times N}$ (usually $L \gg N$) denote the analysis operator associated with some wavelet tight frame system. The corresponding synthesis operator is then its transpose $W^\top$. There are three types of wavelet tight frame based approaches: *synthesis* approach, *analysis* approach, and *balanced* approach. The minimization model of the synthesis approach is defined as

$$u^* = W^\top \alpha^*; \quad \alpha^* = \underset{\alpha\in\mathbb{R}^L}{\arg\min}\ \frac{1}{2}\|AW^\top\alpha - g\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1, \tag{1.3}$$

where $\lambda \in \mathbb{R}^L_+$ is some given regularization vector. The analysis approach directly recovers the data u by solving the following minimization:

$$u^* = \underset{u\in\mathbb{R}^N}{\arg\min}\ \frac{1}{2}\|Au - g\|_2^2 + \|\mathrm{diag}(\lambda)Wu\|_1. \tag{1.4}$$

The balanced approach recovers the data by solving the following minimization model

$$u^* = W^\top \alpha^*; \quad \alpha^* = \underset{\alpha\in\mathbb{R}^L}{\arg\min}\ \frac{1}{2}\|AW^\top\alpha - g\|_2^2 + \frac{\kappa}{2}\|(I - WW^\top)\alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1, \tag{1.5}$$

for some parameter $0 \le \kappa \le \infty$. All these three approaches becomes equivalent when using an orthonormal wavelet basis, but they are different with different outcomes when using a redundant wavelet frame. Interested readers are referred to [12, 13] for more details.

For many general image restoration problems, the balanced approach (1.5) and the analysis approach (1.4) are often more preferred than the synthesis approach (1.3), as the former two empirically tend to yield more visually appealing results with less artifacts. The two models (1.5) and (1.4) have been used for solving many image restoration problems (see e.g. [14, 15, 16, 17]), the reconstruction problems in X-ray CT imaging [18], the 3D surface reconstruction problems in medical imaging [19] and reconstruction problems in range sensing [20]. In recent years, many fast numerical solvers have been developed for the minimization problems (1.5) and (1.4). For example, the proximal forward-backward splitting (PFBS) scheme [14, 21, 22, 23] and the accelerated proximal gradient (APG) method [24, 25, 26] are used to solve (1.5) in various image restoration tasks. The split Bregman iteration [27, 28, 16] has been widely used to solve the model (1.4) adopted in many imaging applications. The linearized Bregman method [29, 30] has been used in [31] for solving a modified version of the synthesis model (1.3) in image restoration. The connection between the balanced model (1.5) and the modified synthesis model solved by the linearized Bregman iteration is also studied in [31].

The wavelet tight frame based regularizations are based on the assumption that the signal can be sparsely approximated under some wavelet tight frame system. For a wide range of bio-molecular particles, their 3D density maps can be roughly modeled by piecewise smooth functions in $\mathbb{R}^3$. It is known that piecewise smooth signals can be sparsely approximated by spline framelets [32, 13]. Thus, the sparsity prior of the 3D density map in some 3D framelet system is likely to hold true for most bio-molecular particles. Motivated by the success of the framelet methods in image recovery and CT reconstruction (see e.g. [14, 15, 16, 17, 18, 19]), we are interested in developing a powerful wavelet tight frame based 3D reconstruction method that can outperform the existing state-of-the-art EM techniques.

The scale of the reconstruction problem in EM is very large. For example, considering the cryo-EM Adenovirus dataset [33] with total 9621 images of size 408×408, the matrix A in (1.1) is of size $\mathrm{O}(10^9) \times \mathrm{O}(10^7)$ and there is no known transform to diagonalize it. For such a large-scale problem, the computational efficiency of the 3D reconstruction algorithm not only depends on its CPU time usage, but also depends on its memory usage. Any of the three sparsity-based regularizations under some 3D wavelet tight frame can be used to solve the linear system (1.1) related to the 3D reconstruction problem in EM. However, the existing popular numerical solvers for these models have serious issue on the usage of physical memory. In the next, we give an examination on the memory usage of three popular numerical solvers under a typical EM configuration.

Suppose that we use the analysis-based model (1.4) for recovery and solve it via the split Bregman iteration [28]:

$$\left\{\begin{array}{rcl} u_{k+1} &=& (A^\top A + \mu I)^{-1}(A^\top g + \mu W^\top (d_k - b_k)); \\ d_{k+1} &=& T_{\lambda/\mu}(W u_{k+1} + b_k); \\ b_{k+1} &=& b_k + \delta(W u_{k+1} - d_{k+1}) \end{array}\right. \tag{1.6}$$

for some parameters $\mu > 0$ and $\lambda \in \mathbb{R}^L_+$, where T_λ is defined by

$$T_\lambda([\alpha_1, \alpha_2, \cdots, \alpha_L]^\top) = [t_{\lambda_1}(\alpha_1), t_{\lambda_2}(\alpha_2), \cdots, t_{\lambda_L}(\alpha_L)]^\top, \tag{1.7}$$

with the soft-thresholding operator $t_{\lambda_i}(\alpha_i) = \mathrm{sgn}(\alpha_i)\max(|\alpha_i| - \lambda_i, 0)$. The first step in the above iteration requires solving a linear system whose associated matrix $A^\top A + \mu I$ is of extremely large size and cannot be diagonalized by some known fast transform. One may either directly solve it which is expensive in terms of CPU time usage, or find an approximate solution using some gradient descent method which is more efficient. It is empirically observed that the later uses less CPU time than the former to obtain a satisfactory result. It can be seen that, during each iteration, the scheme (1.6) needs to store one intermediate frame coefficient vector b_k for the usage in the next iteration. Considering a ET HIV-1 dataset [34] with about 70-80 images of size 2048×2048. For a density map of size $2048 \times 2048 \times 200$, it needs 84 gigabytes (GB) memory to store one frame coefficient vector under the single-level un-decimal 3D tensor linear framelet system. In other words, in a typical EM setting, the iterative scheme (1.6) has to aggressively call virtual memory which significantly slows down the computation.

Suppose that we use the linearized Bregman iteration [29, 30] for 3D density map reconstruction:

$$\left\{\begin{array}{l} \alpha_{k+1} = \frac{1}{\kappa} T_\lambda(\beta_k), \\ \beta_{k+1} = \beta_k + W A^\top (g - A W^\top \alpha_{k+1}) \end{array}\right. \tag{1.8}$$

with some $\kappa > 0$. The recovered density map is defined as $W^\top \alpha_{K+1}$ after K iterations. It is shown in [31] that the iteration scheme (1.8) above indeed minimizes a modified synthesis model (1.3) with an additional least squares term $\frac{\kappa}{2}\|\alpha\|_2^2$. It can be seen that the iteration (1.8) updates its estimation in wavelet frame domain. Thus, same as the split Bregman iteration, the scheme (1.8) also needs to store at least one frame coefficient vector of the intermediate 3D structure estimation during each iteration.

If we consider (1.5) for the recovery of density map and solve it via the APG

method [24, 25]:

$$\begin{cases} \beta_k & = & \alpha_k + \frac{t_{k-1}-1}{t_k}(\alpha_k - \alpha_{k-1}); \\ \alpha_{k+1} & = & T_{\lambda/l}\left(\beta_k - \nabla F(\beta_k)/l\right); \\ t_{k+1} & = & (1+\sqrt{1+4t_k^2})/2 \end{cases} \tag{1.9}$$

where $t_{-1} = 0$, $t_0 = 1$, $F(\alpha) = \frac{1}{2}\|AW^\top\alpha - g\|_2^2 + \frac{\kappa}{2}\|(I - WW^\top)\alpha\|_2^2$ with l-Lipschitz continuous gradient. The recovered density map is defined as $W^\top\alpha_{K+1}$ after K iterations. Same as (1.8), the iteration (1.9) also updates the estimation in frame coefficient domain, and it requires to store two wavelet coefficient vectors β_k and α_{k+1} during each iteration. As a result, the iteration (1.9) doubles the already excessive memory usage of the iteration (1.6) or (1.8).

In summary, in terms of computational efficiency, the existing sparsity-based models and their numerical solvers are not very suitable when being applied to solve 3D reconstruction problems in EM. All of them will lead to excessive usage of physical memory when processing real EM datasets. Motivated by the need of memory efficient methods for 3D reconstruction in EM, we aim at developing a wavelet frame based method that not only has good performance in EM reconstruction but also is efficient on both CPU time usage and physical memory usage.

1.2. Basic ideas. Algorithm (1.9) is efficient in terms of CPU time usage, but it updates the estimation in the wavelet tight frame space which requires excessive memory usage. The solution to reduce memory usage is to directly update the estimation in 3D image domain, rather than in 3D wavelet tight frame domain. One such iterative scheme is proposed in [14] for image inpainting

$$u_{k+1} = (I - P_\Lambda)W^\top T_\lambda(Wu_k) + P_\Lambda g, \tag{1.10}$$

where P_Λ denotes the projection operator. It is shown in [14] that the iteration (1.10) is solving the minimization problem of the balanced approach (1.5) with $A = P_\Lambda$. The iteration (1.11) is specialized by two special properties of the projection operator P_Λ: $P_\Lambda^\top = P_\Lambda$ and $P_\Lambda^\top P_\Lambda = P_\Lambda$. A more general expression of (1.10) indeed should be

$$u_{k+1} = (I - P_\Lambda^\top P_\Lambda)W^\top T_\lambda(Wu_k) + P_\Lambda^\top g.$$

Motived by the iterative scheme (1.10) for image resolution enhancement, we proposed the following iteration for general matrix A:

$$u_{k+1} = (I - \mu A^\top A)W^\top T_\lambda(Wu_k) + \mu A^\top g, \tag{1.11}$$

where μ is some pre-defined positive parameter. In contrast to (1.9), the iteration (1.11) only requires the storage of the intermediate estimation u_k in the image space, whose size is only 1/54 of that of two wavelet frame coeffcient vectors used in the APG method. It is noted that the analysis operator W and the synthesis operator $W^\top$ are computed via a bunch of convolution operations. Each convolution operation can be executed independently of the other convolutions. There is no need to store the whole vector Wu_k. More detailed analysis on the memory usage of the proposed iteration (1.11) will be presented in Section 2. Moreover, the iteration (1.11) is very efficient on memory usage, but is not as efficient as (1.9) in terms of CPU time usage. Thus, we proposed an accelerated version of (1.11) using the APG strategy. The proposed numerical method is not only computationally efficient on CPU time usage, but also is efficient on physical memory usage.

The iteration (1.11) can be understood by the PFBS scheme discussed in [22] and the similar idea is also used in the iterative thresholding method [35] for non-redundant system. Indeed, iteration scheme (1.11) can also be derived from the proximal point algorithm for solving the following balanced model:

$$\frac{\mu}{2}\|AW^\top\alpha - g\|_2^2 + \frac{1}{2}\|(I - WW^\top)\alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1, \tag{1.12}$$

whose iteration is given as follows,

$$\begin{cases} \alpha_{k+1} = T_\lambda(Wu_k); \\ u_{k+1} = (I - \mu A^\top A)W^\top\alpha_{k+1} + \mu A^\top g. \end{cases}$$

One may merge two steps into one and then it becomes the iteration scheme (1.11). However, we would like to point out that the solution $u^* = \lim_{k\to\infty} u_k$ obtained from the iteration (1.11) is different from the solution from the balanced model (1.12), which usually is defined by

$$v^* = W^\top\alpha^*, \text{ where } \alpha^* = \lim_{k\to\infty} \alpha_k.$$

The relationship between u^* and v^* can be described as

$$v^* = W^\top\alpha^* = W^\top T_\lambda(Wu^*).$$

We note that as long as $\lambda \neq 0$, we have $v^* \neq u^*$. Empirically, there are also many noticeable differences between the 3D density maps corresponding to these two solutions, but it is hard to tell in general which result is better as there are always some parts on which one is better than the other one and vice versa.

In addition to the gain on memory efficiency, the iteration scheme (1.11) also leads to a better understanding on its underlying variational model. Although the wavelet coefficient α^* is well-understood as one minimizer of (1.12), it is not clear what exactly the cost functional v^* is minimizing from the relationship $v^* = W^\top\alpha^*$ when $W^\top$ is the synthesis operator of a redundant system such as wavelet tight frame. Given an image v, there is an infinite number of frame coefficient vectors α satisfying $v = W^\top\alpha$, as $W^\top$ is a many-to-one linear mapping for a redundant system. Thus, it is difficult to understand the objective functional that the solution v^* from (1.12) is corresponds to in terms of v^* itself or its canonical coefficient vector Wv^* (recall that $\alpha^* \neq Wv^*$). For a better understanding of the result obtained from the proposed method, it is good to know exactly what the cost functional the solution is minimizing in terms of the solution itself or its canonical coefficient vector. In Section 2.2 and Appendix A, we present a detailed discussion on the variational model that the iteration (1.11) corresponds to and it is different from (1.12).

1.3. Related work. The earlier works on 3D reconstruction in EM can be classified into three categories: Fourier methods, analytic methods and iterative methods. The Fourier methods [1, 2, 36, 37] are based on the central slice theorem [38] which says the Fourier transform of a 2D projection is equal to one central slice of the Fourier transform of 3D object. Thus, the Fourier methods reconstruct the 3D map in Fourier domain by interpolating all Fourier coefficients from its values on many slices with different orientations. The analytic methods are derived from the numerical implementation of the inversion formula for ray transform. One popular analytic method for 3D reconstruction is the filtered back-projection and its variations [39, 40].

Another analytic method widely used in EM is the weighted back-projection method (WBP) [41, 42] which performs the deconvolution on the back-projected map using the weighting function derived from the point spread function of the back-projection operator. The representative iterative methods include algebraic reconstruction technique (ART) [43, 44, 45], simultaneous iterative reconstruction technique (SIRT) [46] and their variations [47, 45]. The ART method, also known as Kaczmarz method, iteratively solves a linear system by sequentially solving the equations in the linear system, while the SIRT method iteratively solves a linear system by using the least squares rules on all equations.

In recent years, the ℓ_1 minimization based regularization approach has been widely used for solving linear inverse problems in imaging science including the 3D reconstruction problem in EM or CT. The representative regularization is the total variation (TV) based regularization first proposed by [48]. The TV-based method restores the data by solving the following minimization model:

$$\min_u \frac{1}{2}\|Au - g\|_2^2 + \mu\|\nabla u\|_1, \tag{1.13}$$

where u is the data for recovery, μ is some pre-defined regularization parameters and ∇ is the first-order difference operator or some higher-order difference operator. The model (1.13) has been widely used in many generic image restoration tasks (see e.g. [49, 50, 51]) and in the reconstruction problem in cone-beam CT (see e.g. [52, 53]). For 3D reconstruction problem in EM, the TV-L2GF regularization method is also developed in [54, 55] with very promising results. It is noted that there exists a deep connection between the wavelet tight frame based regularization and TV regularization or its variations. Indeed, it is shown in [56] that, by choosing parameters properly, a special case of the wavelet tight frame approach can be seen as a sophisticated discretization of minimizations involving the TV regularization or their generalizations.

Using an orthonormal wavelet basis [57], a Landweber iteration with soft thresholding is proposed in [58] for SPM reconstruction:

$$u_{k+1} = W^\top T_\lambda(W(u_k + A^\top(g - Au_k))). \tag{1.14}$$

The convergence of the iteration above is given in [35]. Since the underlying system is an orthonormal basis, all three models (1.3), (1.4), (1.5) are equivalent and the iteration (1.14) converges to a minimizer of any of them.

This paper is organized as follows. Section 2 presents an iterative algorithm for 3D reconstruction in EM and its accelerated version, followed by the discussions on the associated minimization model, the convergence rate and the feasibility in EM reconstruction. In Section 3, the algorithm is applied to solve the 3D reconstruction problem in both the SPM and the ET, together with some extensive experimental evaluations on several datasets. The functional that the method minimizes and the outlined proofs of the theoretical results presented in Section 2 are provided in Appendix A.

2. Main results. The 3D reconstruction method developed in this paper is built on the sparse approximation of the 3D density map under the wavelet tight frame system. Next, we first give a brief introduction to wavelet tight frame.

2.1. Preliminaries on wavelet tight frame. We first present here some basics of a tight frame in a Hilbert space $\mathcal{H}$. Interested readers are referred to [32, 13] for more

details. Let $\|\cdot\|$ denote the norm of the Hilbert space $\mathcal{H}$. A sequence $\{\phi_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ is a tight frame for $\mathcal{H}$ if

$$\|f\|_2^2 = \sum_{n\in\mathbb{N}} |\langle f, \phi_n\rangle|^2, \quad \forall\, f \in \mathcal{H}.$$

There are two operators associated with a given sequence $\{\phi_n\}_{n\in\mathbb{N}}$: the analysis operator

$$\mathcal{W} : f \in \mathcal{H} \longrightarrow \{\langle f, \phi_n\rangle\} \in \ell^2(\mathbb{N})$$

and its adjoint operator $\mathcal{W}^*$, also called the synthesis operator, which is defined by

$$\mathcal{W}^* : \{a_n\} \in \ell^2(\mathbb{N}) \longrightarrow \sum_{n\in\mathbb{N}} a_n\phi_n \in \mathcal{H}.$$

Thus, a sequence $\{\phi_n\} \subset \mathcal{H}$ forms a tight frame if and only if $\mathcal{W}^*\mathcal{W} = I$, where $I : \mathcal{H} \longrightarrow \mathcal{H}$ is the identity operator. Tight frame is a generalization of orthonormal basis in the sense that both have the same perfect reconstruction property:

$$f = \sum_{n\in\mathbb{N}} \langle f, \phi_n\rangle \phi_n, \quad \forall\, f \in \mathcal{H}. \tag{2.1}$$

In the discrete setting, for a given tight frame $\{\phi_n\}_{n=1}^L \subset \mathbb{R}^N$, the matrix representation of the associated analysis operator is $W = (\phi_1, \phi_2, \ldots, \phi_L)^\top$ and the synthesis operator is its transpose $W^\top$. Clearly, $\{\phi_n\}_{n=1}^L$ forms a tight frame for $\mathbb{R}^N$ if and only if $W^\top W = I_N$, where I_N is the N-by-N identity matrix.

One widely used tight frame in signal/image processing is the multi-resolution analysis (MRA) based wavelet tight frame. The wavelet tight frame for $\mathbb{R}^N$ can be constructed from the filters associated with some MRA-based framelets for the space of continuum (see [13] for more details). For simplicity, only single-level un-decimal wavelet tight frame system for $\mathbb{R}^N$ is introduced. Given a framelet system, suppose $\{a_0, a_1, \ldots, a_l\}$ are the associated filters with finite support. For a given filter a of finite length, let the N-by-N matrix, denoted by $\mathcal{S}_a$, be the Toeplitz matrix that represents the convolution operator by the mask a under Neumann boundary condition. Then, the columns of the following N-by-$(l+1)N$ matrix:

$$[\mathcal{S}_{a_0}^\top, \mathcal{S}_{a_1}^\top, \ldots, \mathcal{S}_{a_l}^\top] \tag{2.2}$$

forms a tight frame for $\mathbb{R}^N$. Let W denote the transpose of (2.2), then W is the analysis operator and its transpose $W^\top$ is the synthesis operator of the tight frame system (2.2). The perfect reconstruction property in matrix representation can be expressed as $W^\top W = I_N$.

The 3D framelet system system can be generated by the tensor product of three 1D framelet systems. Correspondingly, the associated analysis operator W for 3D signal is the Kronecker product of three analysis operators of 1D framelet systems. The 3D framelet system used in this paper is the tensor product of 1D linear spline framelet constructed via Unitary Extension Principle [32]. The 1D linear spline framelet filers are listed as follows:

$$a_0 = \frac{1}{4}(1, 2, 1); \quad a_1 = \frac{\sqrt{2}}{4}(1, 0, -1); \quad a_2 = \frac{1}{4}(-1, 2, -1). \tag{2.3}$$

It is noted that writing W as a matrix is for notational convenience. In the actual computation, the matrix multiplication by W or by $W^\top$ is equivalent to the convolution operation using the filters associated with the framelet system. Taking 3D tensor linear spline framelet as an example, there are totally 27 associated 3D filters. For a signal $f \in \mathbb{R}^N$, then the corresponding framelet coefficients are the collection of the outputs of the signal convoluted by these 27 filters. We have then $Wf \in \mathbb{R}^{27N}$ for $f \in \mathbb{R}^N$. In other words, the size of a single-level framlet coefficient vector is 27 times of that of the signal.

2.2. Numerical algorithm and analysis. In this section, we will first establish the convergence of the iteration scheme (1.11), the functional it minimizes and its connection to the balanced model (1.5). Then we present the accelerated version of (1.11) using the same idea as the APG method.

In practice, not only the theoretical convergence of (1.11) is needed for guaranteeing the numerical stability, but also it is important to know the exact variational problem that the limit of the iteration corresponds to. In our case, the iteration scheme is developed first from the computational need, and the underlying functional that the iteration aims at minimizing is not clear. Hence, the available analysis techniques in optimization can not be applied before we find out the underlying objective functional. It is shown in [14] that the iteration scheme is minimizing a model closely related to the balanced model (1.5) for the case of $A = P_\Lambda$. Thus, one might hypothesize that for general operator A, the iteration (1.11) is also minimizing a model related to the balanced model (1.5) which actually does not hold true. The iteration (1.11) is solving a minimization model different from the balanced model (1.5); see Appendix A for the details of the process of finding the exact objective functional. For a positive vector $\lambda \in \mathbb{R}^L_+$, define the multivariate Huber function H by

$$H_\lambda([\alpha_1, \alpha_2, \dots, \alpha_L]^\top) = \sum_{i=1}^{L} h_{\lambda_i}(\alpha_i), \tag{2.4}$$

where

$$h_{\lambda_i}(\alpha_i) = \begin{cases} \frac{1}{2}\alpha_i^2, & |\alpha_i| < \lambda_i; \\ \lambda_i(|\alpha_i| - \frac{1}{2}\lambda_i), & |\alpha_i| \geq \lambda_i. \end{cases} \tag{2.5}$$

THEOREM 2.1. *Let $\{u_k\} \subset \mathbb{R}^N$ be the sequence generated by (1.11) for an arbitrary $u_0 \in \mathbb{R}^N$. Suppose that $\mu > 0$ is chosen such that $D := (I - \mu A^\top A)^{-1}$ is positive definite. Then, the sequence $\{u_k\}$ converges and its limit minimizes the following functional:*

$$F(u) := \frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2 + H_\lambda(Wu), \tag{2.6}$$

where $\|\cdot\|_D$ is the norm defined by $\|x\|_D = (x^\top Dx)^{\frac{1}{2}}$ for any $x \in \mathbb{R}^N$.

Proof. See Appendix A. □

There are two parts in the minimization (2.6) solved by the iteration (1.11). The first part is the fidelity term consisting of two components: one is the often seen least squares fidelity term; and the other is a weighted least squares fidelity term with the weighting matrix given by $A(I - \mu A^\top A)^{-1}A^\top$. The second part is the regularization term $H_\lambda(Wu)$ which is indeed the minimum value of the following functional:

$$\frac{1}{2}\|Wu - \alpha\|_2^2 + \|\operatorname{diag}(\lambda)\alpha\|_1$$

corresponding to the minimizer $T_\lambda(Wu)$. Although the minimization model (2.6) appears different from the balanced model (1.5), these two models are closely related as shown in the following theorem.

THEOREM 2.2. *Let $\{u_k\} \subset \mathbb{R}^N$ be the sequence generated by (1.11) for an arbitrary $u_0 \in \mathbb{R}^N$. Let u^* denote its limit, i.e., $\lim_{k\to\infty} u_k = u^*$. Then the vector $\alpha^* := T_\lambda(Wu^*)$ is a minimizer of the following balanced model functional:*

$$\frac{\mu}{2}\|AW^\top\alpha - g\|_2^2 + \frac{1}{2}\|(I - WW^\top)\alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1. \tag{2.7}$$

Proof. See Appendix A. □

Compared to (1.9), the iteration (1.11) has a much smaller memory footprint, but it is not as efficient as the iterative method using the APG idea in terms of CPU time usage. Let $\{u_k\}$ be the sequence generated by (1.11) from some initial guess u_0. By the same arguments in [59, Theorem 4.4], we have

$$F(u_k) - F(u^*) \le \frac{\|u^* - u_0\|_2^2}{2k} \tag{2.8}$$

for some minimizer u^* of (2.6). Motivated by the additional gain on computational efficiency using the APG idea, we propose the following iterative method for 3D reconstruction in EM.

Algorithm 1 Framelet based method for 3D reconstruction

Given $\lambda \in \mathbb{R}_+^L$ and $\mu > 0$,

(i) set initial guess v_0 and v_1, and set $t_{-1} := 0$, $t_0 := 1$.

(ii) For $k = 0, 1, \ldots$, iterate until convergence

$$\begin{cases} \overline{v}_k := v_k + \frac{t_{k-1}-1}{t_k}(v_k - v_{k-1}); \\ v_{k+1} := (I - \mu A^\top A)W^\top T_\lambda(W\overline{v}_k) + \mu A^\top g; \\ t_{k+1} := \frac{1+\sqrt{1+4t_k^2}}{2}. \end{cases}$$

Let $\{v_k\}$ be the sequence generated by Algorithm 1 from some initial guess v_0. Then by the same analysis of the convergence rate of the APG algorithm provided in [59, Theorem 4.5], we have

$$F(v_k) - F(v^*) \le \frac{2\|v^* - v_0\|_2^2}{(k+1)^2} \tag{2.9}$$

for some minimizer of v^* of (2.6). It is seen that, for any given error tolerance $\epsilon > 0$, within $K = O(\epsilon^{-\frac{1}{2}})$ iterations, the sequence $\{v_k\}$ generated by Algorithm 1 achieves ϵ-optimality: $\|F(v_k) - F(v^*)\| < \epsilon$. In contrast, the sequence generated by the iteration (1.11) needs K^2 iterations to achieve the same ϵ-optimality.

It is noted that, there is a small loss on the efficiency of memory usage when replacing the iteration (1.11) by Algorithm 1, as there is one additional intermediate estimate of the 3D density map need to be stored during the iteration. However, compared to (1.9), the additional memory required for storing such an additional map is minimal, since there are two intermediate wavelet tight frame coefficient vectors

needed to be stored for (1.9). Recall that the size of two wavelet coefficient vectors is 54 times of that of one 3D density map. In summary, for Algorithm 1, a small loss on the memory usage leads to a big gain on the efficiency of CPU time usage.

2.3. Implementation details for parallelization. The typical computing environment for 3D reconstruction in EM is high-end computer cluster with many connected computing nodes. The structure of Algorithm 1 allows a highly parallel implementation. The only non-trivial step in Algorithm 1 is Step (ii) which reads

$$u_{k+1} = v_k - \mu A^\top (Av_k - g), \tag{2.10}$$

where $v_k = W^\top T_\lambda(W\overline{u}_k)$. The matrix A, W and vector g all have the following block structures:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}, g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{bmatrix}, W = \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_l \end{bmatrix},$$

where g_i denotes the i-th image, A_i denotes its associated measuring matrix, and W_j denotes the channel associated with the j-th wavelet filter. There are totally m images and $l+1$ wavelet channels. Then we can rewrite (2.10) as the following:

$$v_k = \sum_{j=0}^{l} W_j^\top T_{\lambda_j}(W_j\overline{u}_k);$$
$$u_{k+1} = v_k - \mu \sum_{i=1}^{m} \left(A_i^\top (A_i v_k - g_i)\right).$$

It is seen that the computation of u_{k+1} can be distributed to $s(\leq m)$ processes when computing $A_i^\top(A_i v_k - g_i)$ for $i = 1, 2, \ldots, m$, and the computation of v_k can also be distributed to l processes when computing $W_j^\top T_{\lambda_j}(W_j\overline{u}_k)$ for $j = 0, 1, 2, \ldots, l$. In practice, considering the overhead for data communication among processes using message passing interface, we do not parallelize the computation of v_k and only parallelize the computation of u_{k+1} by distributing the computation of $\{A_i^\top(A_i v_k - g_i)\}_{i=1}^m$ to s processes. In the next, we give a detailed analysis on the memory usage of the parallel implementation of Algorithm 1 in practical scenarios for the SPM and the ET.

For SPM, consider two 3D density maps obtained from the atomic structure file 1J4Z.pdb[1] of the chaperonin GroEL of bacterium. One is of size $256 \times 256 \times 256$ and the other is of size $512 \times 512 \times 512$. For each density map, construct three SPM datasets with 10^3, 10^4 and 10^5 images with random orientations. The Algorithm 1 is parallelized using MPI (message passing interface). Then the 3D reconstruction of these two density maps is done by running Algorithm 1 on the 13 computing nodes of the computer cluster[2] with totally 100 processes. Only one thread is run on each process. The details of the memory usage and the CPU time consumption are summarized in Table 2.1.

[1] `http://www.rcsb.org/pdb/explore.do?structureId=1j4z`

[2] Computer cluster "LSSC3" in ICMSEC at CAS with totally 282 nodes annd 24 GB memory each

TABLE 2.1
Memory usage and CPU time of Algorithm 1 for SPM

number of images	image size	3D map size	thread	process	memory usage (GB) each node	wall time for one iteration
10^3	$2^8 \times 2^8$	$2^8 \times 2^8 \times 2^8$	100	100	2.00	54s
10^4	$2^8 \times 2^8$	$2^8 \times 2^8 \times 2^8$	100	100	2.16	223s
10^5	$2^8 \times 2^8$	$2^8 \times 2^8 \times 2^8$	100	100	3.92	1551s
10^3	$2^9 \times 2^9$	$2^9 \times 2^9 \times 2^9$	100	100	15.60	586s
10^4	$2^9 \times 2^9$	$2^9 \times 2^9 \times 2^9$	100	100	16.32	1610s
10^5	$2^9 \times 2^9$	$2^9 \times 2^9 \times 2^9$	100	100	23.36	12078s

TABLE 2.2
Memory usage and CPU time of Algorithm 1 for ET

number of images	image size	3D map size	thread	proc.	memory usage (GB) each node	wall time one one iteration
200	$2^9 \times 2^9$	$2^9 \times 2^9 \times 200$	32	4	1.22	159s
200	$2^{10} \times 2^{10}$	$2^{10} \times 2^{10} \times 200$	32	4	4.89	598s
200	$2^{11} \times 2^{11}$	$2^{11} \times 2^{11} \times 200$	32	4	19.56	2141s

For ET, we use the same atomic structures as the experiments of the SPM. Three ET density maps with different sizes are considered: $512 \times 512 \times 200$, $1024 \times 1024 \times 200$, and $2048 \times 2048 \times 200$. For each density map, 200 images with the tiling angle randomly chosen within the range $[-70°, +70°]$ are used for the reconstruction. In practice, the ET density map is usually significantly larger than the SPM map. Thus, Algorithm 1 is parallelized using MPI plus OpenMP (Open Multi-Processing) and is run on 4 computing nodes, each node runs one process with 8 OpenMP threads. The memory usage and the CPU time are summarized in Table 2.2.

From Table 2.1 and Table 2.2, it can be seen that the memory usage of Algorithm 1 is well constrained for the SPM datasets with hundreds of thousands of images of size 512×512 and for the ET datasets with hundreds of images of size 2048×2048. For most practical EM 3D reconstruction tasks, the memory usage of Algorithm 1 is well below the physical memory profile of typical computer clusters used in EM.

3. Numerical experiments. In this section, Algorithm 1 is evaluated for the 3D reconstruction of the SPM and the ET on one simulated dataset and two real datasets. For the SPM reconstruction, Algorithm 1 is compared with the TV-L2GF regularization method proposed in [54], the WBP method [42], the Fourier-4NN method implemented in the SPARX [36], ART [44] and SIRT [46]. The TV-L2GF method proposed in [54] reconstructs the 3D density map via solving the minimization problem (1.13). In [54], the non-differentiable ℓ_1-norm related term in (1.13) is approximated by its smoothed version and then solved via the ℓ_2-gradient flow method. The WBP method proposed in [42] is defined by

$$u = \mathcal{F}^{-1}(\mathcal{F}(v) \cdot \omega)$$

where v denotes the back-projected map, $\mathcal{F}(\cdot)$ denotes the discrete Fourier transform, and ω denotes the weighting function given by some empirical formula. The Fourier-4NN method performs the nearest neighbor interpolation in the Fourier domain with

4-fold zero padding. The ART method [44] accesses the projections with the Euler angle being as different as possible from the previous ones, making the algorithm more effective. The computation of all experiments are done on the computer cluster LSSC3 at CAS, China.

Following the conventions in EM, the evaluation of the results is based on three factors: the visual details of the slices, the visual details of iso-surfaces and more importantly the resolution of the reconstructed 3D map. The resolution of a reconstructed 3D map is based on the so-called *Fourier shell correlation* (FSC) defined by

$$\mathrm{FSC}(r) = \frac{\sum_{r_i \in r} \mathcal{F}(f_1)(r_i) \cdot \overline{\mathcal{F}(f_2)}(r_i)}{\sqrt{\sum_{r_i \in r} |\mathcal{F}(f_1)(r_i)|^2 \cdot \sum_{r_i \in r} |\mathcal{F}(f_2)(r_i)|^2}}, \tag{3.1}$$

where f_1 denotes the reconstructed 3D density map and f_2 denotes the ground truth. When the ground truth is not available for the data from real bio-imaging experiments. the common practice is splitting the EM dataset randomly into two sub-sets, and running the reconstruction algorithm on each sub-set independently. Then, the widely used resolution estimation of the result is defined as the $1/2$ cut-off of the FSC between the outputs of two sub-sets, that is, the reciprocal of the frequency at $\mathrm{FSC} = 1/2$. When the structure from X-ray crystallographic methods is available, the FSC is computed between the EM map and the X-ray map. As the X-ray structure is usually less erroneous than the EM map, the cut-off threshold is then chosen higher than $1/2$. For example, $\mathrm{FSC} = 0.82$ is used in [60].

In the experiments of SPM, Algorithm 1 stops after 20 iterations. For the ART method with projections ordered by the algorithm proposed in [44], 1 or 2 iterations is usually sufficient to obtain valuable results. 50 iterations are used in the SIRT method [46] and the TV-L2GF method. All the parameters in the methods for comparison are tuned up to reach optimal results. Interested readers are referred to [61, 62] for a detailed discussion on the relaxation parameters of ART and SIRT. For the ET reconstruction, Algorithm 1 is compared against the WBP method and the TV-L2GF method. For the ET dataset used in the experiment, Algorithm 1 stops after 20 iterations and the TV-L2GF method stops 50 iterations. All iterative methods explored in the SPM and ET experiments start from initial zero.

3.1. Simulated SPM dataset: E. coli ribosome. In this experiment, the SPM dataset is simulated as follows. A 3D density map of E. coli ribosome with pixel size 2.82 Å is first generated using the crystal structure 3I1M.pdb[3] and 3I1N.pdb[4]. Then 5000 2D images with random orientations are generated from this 3D density map. Each image is of size 131×131 pixels. Since real EM images have extremely low SNR, Gaussian white noise is added on each image such that the SNR of the resulting noisy image is only 0.1. See Fig. 3.1 for a visual illustration of some simulated noisy EM images.

Algorithm 1 and the other five methods for comparison are tested on such a simulated dataset. The resolutions of the results are defined by using the 0.82 cut-off of the FSC curves; see Fig. 3.2 for the FSC graphs of the results from the six methods.

[3] `http://www.rcsb.org/pdb/explore.do?structureId=3i1m`
[4] `http://www.rcsb.org/pdb/explore.do?structureId=3I1N`

FIG. 3.1. *Simulated 2D noisy EM images of E. coli ribosome with* SNR $= 0.1$.

TABLE 3.1
Comparison of the six methods: Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1 on the simulated dataset experiment.

	thread	process	iteration	resolution	total time (sec.)
Fourier-4NN [36]	1	1	non.	20.80 Å	294
ART[44]	32	32	2	20.84 Å	2172
WBP [42]	1	1	non.	20.76 Å	2114
SIRT [46]	32	32	50	20.75 Å	1083
TV-L2GF [54]	32	32	50	18.51 Å	24148
Algorithm 1	32	32	20	17.65 Å	733

The iteration numbers for ART, SIRT, TV-L2GF and Algorithm 1 are 2, 50, 50, and 20 respectively. The numerical results are summarized in Table 3.1. It is seen that Algorithm 1 obtained a 3D density map with higher resolution than the other five methods. The visual comparison of the iso-surfaces and slices of the results is shown in Fig. 3.3. The ground truth of the 3D density map is obtained from the 3D density map used for synthesis with a resolution low-pass filtered to 10 Å. Clearly, the 3D density map reconstructed from Algorithm 1 has less artifacts than those from the other five methods.

3.2. Experimental cryo-EM SPM dataset: 50S ribosomal subunit. In this experiment, we use the public cryo-EM SPM 50S ribosomal subunit dataset[5] published in [11] to evaluate the Algorithm 1. This dataset contains 82575 images with the image size 100×100 and the pixel size 3.26 Å. These images have been CTF corrected using a Weiner filter with ACE2 [63]. A projection matching refinement was performed using the SPIDER software [64] to estimate the orientations of these particles.

We compare these algorithms with different sizes of datasets, a small dataset with randomly selected 825 images, a median dataset with randomly selected 8257 images and the entire dataset with 82575 images. For each dataset, we reconstruct the density map using the six methods: Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1. For the small dataset, 48 processes are used to parallelize the iterative methods including ART, SIRT, TV-L2GF and Algorithm 1. For the median and entire dataset, 96 processes are used for these iterative methods. For ART, one iteration is sufficient. For SIRT and TV-L2GF, 50 iterations are performed to obtain good approximation and for the Algorithm 1, 20 iterations. The resolutions of these reconstructed 3D maps are computed by randomly splitting each dataset into two halves, performing two reconstructions and computing the FSC between the two maps using (3.1) with the 1/2 cut-off. The FSC curves are shown in Fig. 3.5.

The numerical results are given in Table 3.2, 3.3, and Fig. 3.4, 3.6. From Table 3.2,

[5] `http://ami.scripps.edu/redmine/projects/ami/wiki/Ab_initio_model_datasets`

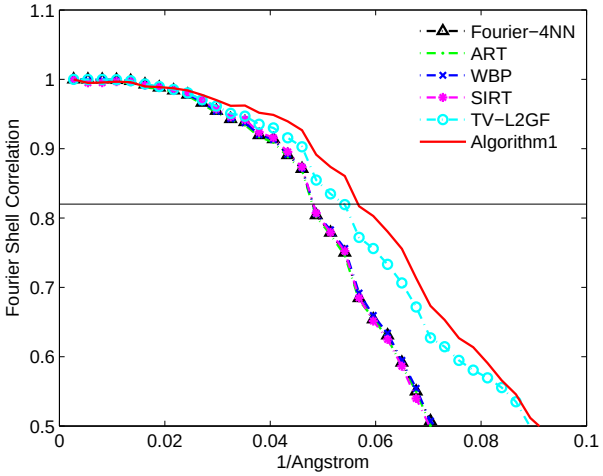


FIG. 3.2. *The comparison of the FSC graphs of the reconstructions from the following six methods: Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1.*

TABLE 3.2
Resolutions of the six methods for the three "50S ribosomal subunit" datasets.

dataset	Fourier-4NN	ART	WBP	SIRT	TV-L2GF	Algorithm 1
825	31.18Å	28.34Å	27.83Å	27.38Å	27.64Å	25.68Å
8257	20.65Å	17.80Å	17.48Å	17.92Å	16.00Å	15.11Å
82575	13.25Å	12.33Å	12.48Å	12.11Å	11.89Å	11.80Å

we can conclude that the Algorithm 1 obtains better resolutions than the other five methods. The main reason for Algorithm 1 improving the resolution is that the soft-thresholding operation in Algorithm 1 smooths out the noise during the iterations so as to increase the consistency between two maps when computing the FSC. With the increasing of the dataset size, the ability of improving the resolution for Algorithm 1 decreases. For example, for the dataset 82575, the resolution of Algorithm 1 is 11.80 Å, a little higher than 11.89 Å, the resolution of TV-L2GF.

From Table 3.3 and Fig. 3.4, it can be seen that the Fourier-4NN method runs fast but produces more noise in the reconstructed maps than the other five methods. The WBP method runs fast for the small dataset but very slow for the entire dataset. The TV-L2GF method runs lowest among the six methods. In conclusion, Algorithm 1 is competitive in running time. Finally, by comparing one slice of the entire dataset reconstructions shown in Fig. 3.6, we can see that the Algorithm 1 has less artifacts than the other five methods.

Since the ET reconstruction is severely ill-posed (a wedge area is missed in the 3D Fourier domain for single-axis tilting ET), in the following ET reconstruction, we use an analytic method WBP (which is fast for small dataset) and two regularized iteration methods Algorithm 1 and TV-L2GF to perform the experiment. Since the 4-fold padding makes the Fourier-4NN method consume much memory, it is not available for ET reconstruction.

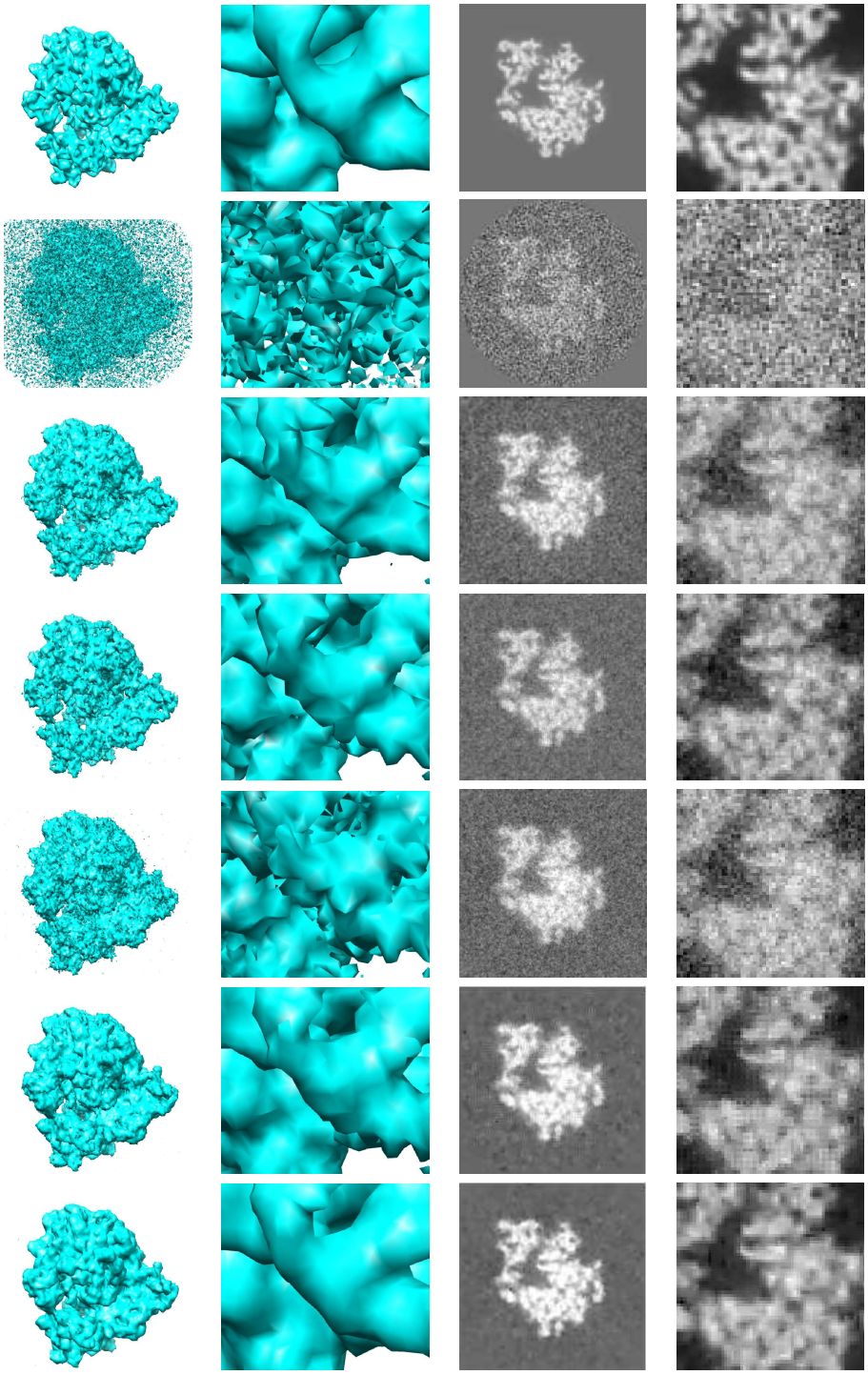


FIG. 3.3. *The comparison of ground truth and the reconstructed density maps from six methods. The maps from the first row to the last row are ground truth, the reconstructions from Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1, respectively. (a) The iso-surfaces of the seven maps; (b) one zoomed-in region of the first column; (c) one slice of the density map; and (d) one zoomed-in region of (c).*

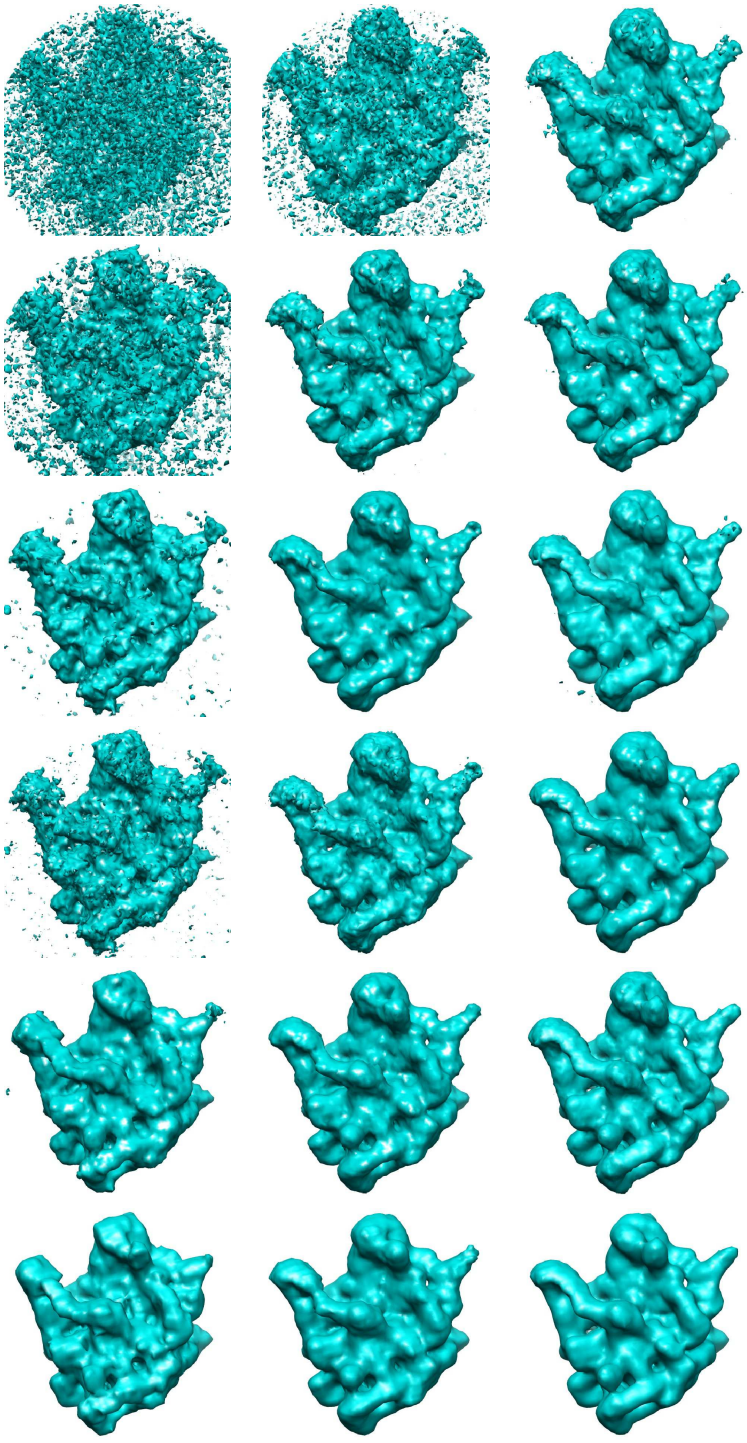
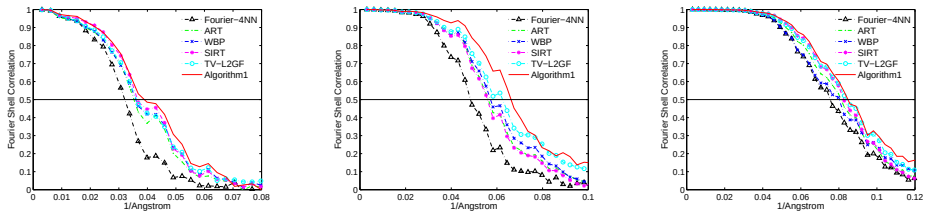


FIG. 3.4. *The reconstructed maps from all six methods for comparison. The maps from the first row to the last row are recovered from Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1 respectively. (a) The results from small dataset; (b) the results from the median dataset; and (c) the results from the entire dataset.*

TABLE 3.3
Total running time (in seconds) of six methods for the three "50S ribosomal subunit" datasets.

dataset	Fourier-4NN	ART	WBP	SIRT	TV-L2GF	Algorithm 1
825	72	107	73	85	1462	43
8257	184	1079	1069	264	5888	185
82575	1406	7002	23488	2552	55992	1854



(a) (b) (c)

FIG. 3.5. *The comparison of the FSC graphs of the reconstructions using the following six methods: Fourier-4NN, ART, WBP, SIRT, TV-L2GF and Algorithm 1. (a) The FSC graphs for the small dataset; (b) the FSC graphs for the median dataset; and (c) the FSC graphs for the entire dataset.*

3.3. Experimental ET dataset: simian immunodeficiency virus (SIV). ET technique using ice-embedded specimens is often called cryo-ET. In this experiment, we run the test on a real cryo-ET dataset "SIV" published in [65]. This dataset contains 55 images with size 512×512, measured with tilt angles between $[-70°, +70°]$. These tilt series have been aligned before running 3D reconstruction. Algorithm 1, the WBP method and the TV-L2GF method are tested on this dataset. In this experiment, the TV-L2GF method and Algorithm 1 stop after 50 and 20 iterations, respectively. The results are shown in Table 3.4. The visual comparison of the slices and iso-surfaces is shown in Fig. 3.7. By comparing the three columns in Fig. 3.7, we see that the result from the WBP method is more noisy than that from the other two methods. The spikes are important structures on the virus envelope. In Fig. 3.7, the region containing a single spike is marked out for each result in the second row; and is zoomed-in in the third row. It is seen that the result from Algorithm 1 has clearer spikes than that from both the WBP method and the TV-L2GF method.

3.4. Discussions. It is seen from the experiments that the proposed framelet-based Algorithm 1 outperformed the five methods in the SPM experiments and the two methods (WBP and TV-L2GF) in the ET reconstruction. The results from the proposed method have higher resolution than those from the other five in the experiments of the SPM. The improvement in terms of quantitative measure is also consistent with the visual improvement of the reconstructed 3D structure. The results from Algorithm 1 have less artifacts and clearer spikes than those from the two existing methods in the ET experiment. In the near future, we would like to investigate how to further improve the reconstruction quality by explicitly addressing other error sources in EM such as the estimation of the projection orientation and the CTF correction.

Acknowledgement. The authors would like to thank the associate editor and the reviewers for careful reading and helpful comments that improve the quality of this manuscript, particularly on the presentation of Section 1.2. The work of Ming Li

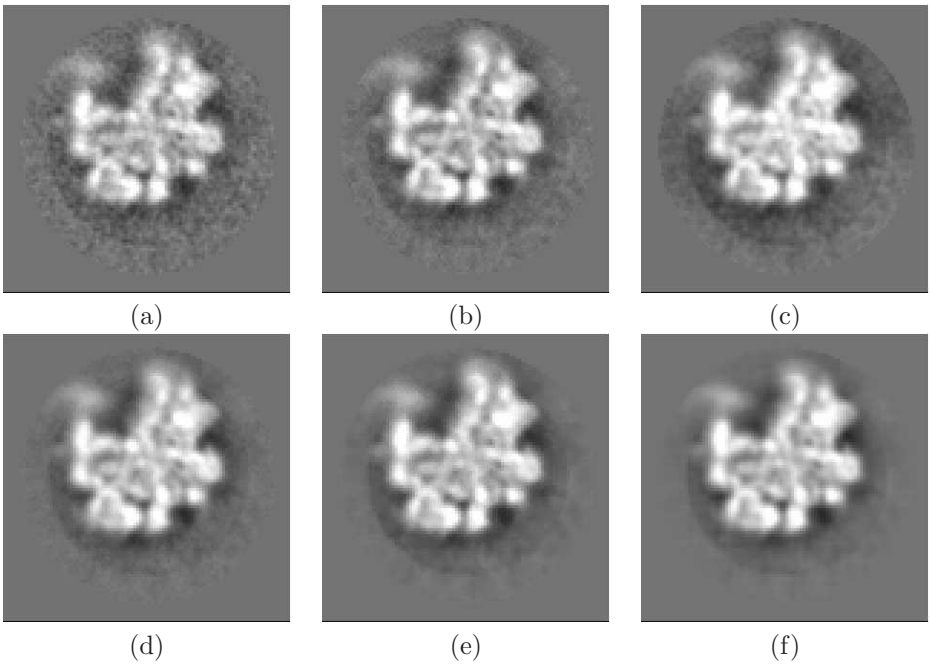


FIG. 3.6. *The comparison of one slice of the six reconstructed maps from entire dataset. (a) Fourier-4NN; (b) ART; (c) WBP; (d) SIRT; (e) TV-L2GF; and (f): Algorithm 1.*

TABLE 3.4
Comparison of the WBP method [42], the TV-L2GF method [54] and Algorithm 1 for experimental cryo-ET dataset SIV.

	thread	process	iteration	total time (sec.)
WBP	1	1	non.	331
TV-L2GF	11	11	50	99634
Algorithm 1	11	11	20	6739

was partially supported by the research grant #11101401 from NSFC, China; and the work of Hui Ji and Zuowei Shen was partially supported by Singapore MOE Research Grant R-146-000-165-112 and MOE2011-T2-1-116.

REFERENCES

[1] R. A. Crowther, D. J. DeRosier, and A. Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc. Roy. Soc., Ser. A*, 317:319–340, 1970.
[2] D. J. DeRosier and A. Klug. Reconstruction of three dimensional structures from electron micrographs. *Nature*, 217:130–134, 1968.
[3] J. Frank. *Electron Tomography Methods for Three-Dimensional Visualization of Structures in the Cell*. Springer, New York, second edition, 2006.
[4] J. Frank. *Three-dimensional electron microscopy of macromolecular assemblies*. Oxford University Press, New York, 2006.
[5] H. R. Saibil. Macromolecular structure determination by cryo-electron microscopy. *Acta Cryst. D*, 56:1215–1222, 2000.
[6] M. V. Heel, B. Gowen, R. Matadeen, E. V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan. Single-particle electron cryo-microscopy:

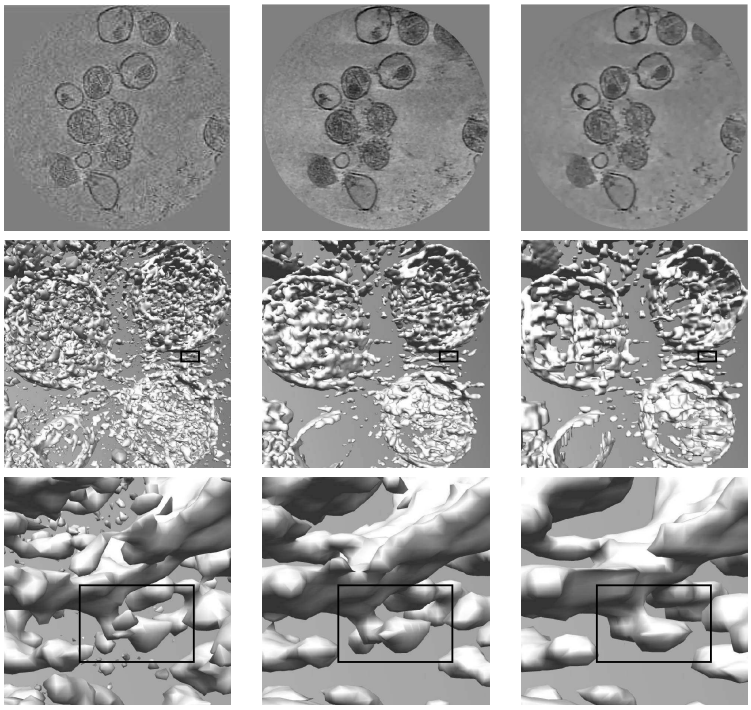


FIG. 3.7. *The reconstructed results from the three methods. (a) The result from the WBP method [42]; (b) the result from the TV-L2GF method [54]; and (c) the result from Algorithm 1. The three rows from top to bottom show the slices, iso-surfaces and zoomed-in regions of the reconstructed density maps respectively.*

towards atomic resolution. *Q. Rev. Biophys.*, 33:307–369, 2000.

[7] M. V. Heel. Angular reconstruction: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21:111–124, 1987.

[8] P. A. Penczek, R. A. Grasucci, and J. Frank. The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, 53:251–270, 1994.

[9] R. McIntosh, D. Nicastro, and D. Mastronarde. New views of cells in 3D: an introduction to electron tomography. *Trends Cell Biol.*, 15:43–51, 2005.

[10] H. Winkler. 3D reconstruction and processing of volumetric data in cryo-electron tomography. *J. Struct. Biol.*, 157:126–137, 2007.

[11] N. R. Voss, D. Lyumkis, A. Cheng, P. W. Lau, A. Mulder, G. C. Lander, E. J. Brignole, D. Fellmann, C. Irving, E. L. Jacovetty, A. Leung, J. Pulokas, J. D. Quispe, H. Winkler, C. Yoshioka, B. Carragher B, and C. S. Potter. A toolbox for ab initio 3-D reconstructions in single-particle electron microscopy. *J. Struct. Biol.*, 169(3):389–398, 2010.

[12] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. In *EUSIPCO*, 2006.

[13] Z. Shen. Wavelet frames and image restorations. In *Proceedings of the international congress of mathematicians*, volume 4, pages 2834–1863, 2010.

[14] J. F. Cai, R. Chan, and Z. Shen. A framelet-based image inpainting algorithm. *Appl. Comput. Harmon. Anal.*, 24:131–149, 2008.

[15] J. F. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion debulrring from a single image using sparse approximation. In *CVPR*, pages 104–111, 2009.

[16] J. F. Cai, S. Osher, and Z. Shen. Split bregman methods and frame based image restoration. *Multiscale Model. Simul.: A SIAM Interdisciplinary Journal*, 8:337–369, 2009.

[17] B. Dong, H. Ji, J. Li, Z. Shen, and Y. Xu. Wavelet frame based blind image inpainting. *Appl. Comput. Harmon. Anal.*, 32:268–279, 2012.

[18] B. Dong, J. Li, and Z. Shen. X-ray CT image reconstruction via wavelet frame based regularization and radon domain inpainting. *J. Sci. Comput.*, 54:333–349, 2013.

[19] B. Dong, A. Chien, Z. Shen, and S. Osher. A new multiscale representation for shapes and its application to blood vessel recovery. *SIAM Journal on Scientific Computing*, 32(4):1724–1739, 2010.

[20] H. Ji, Z. Shen, and Y.-H. Xu. Wavelet frame based scene reconstruction from range data. *J. Comput. Phys.*, 229(6):2093–2108, 2010.

[21] J. F. Cai, R. H. Chan, L. Shen, and Z. Shen. Convergence analysis of tight framelet approach for missing data recovery. *Adv. Comput. Math.*, 31:87–113, 2009.

[22] P.L. Combettes and V.R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.: A SIAM Interdisciplinary Journal*, 4:1168–1200, 2006.

[23] E. T. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for ℓ_1-minimization: methodology and convergence. *SIAM J. Optim.*, 19(3):1107–1130, 2008.

[24] P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. *SIAM J. Optim.*, 2008.

[25] Z. Shen, K. C. Toh, and S. Yun. An accelerated proximal gradient algorithm for frame based image restorations via the balanced approach. *SIAM J. Imaging Sci.*, 4:573–596, 2011.

[26] A. Chambolle and P. Thomas. A first-order primalprima algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.*, 40(1):120–145, 2011.

[27] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation based image restoration. *Multiscale Model. Simul.*, 4(2):460–489, 2005.

[28] T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–343, 2009.

[29] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. *SIAM J. Imaging Sci.*, 1(1):143–168, 2008.

[30] S. Osher, Y. Mao, B. Dong, and W. Yin. Fast linearized bregman iterations for compressive sensing and sparse denoising. *Commun. Math. Sci.*, 8(1):93–111, 2010.

[31] J. F. Cai, S. Osher, and Z. Shen. Linearized bregman iterations for frame-based image deblurring. *SIAM J. Imaging Sci.*, 2:226–252, 2009.

[32] A. Ron and Z. Shen. Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator. *J. Funct. Anal.*, 148:408–447, 1997.

[33] A. J. Pérez-Berná, R. Marabini, S. H. W. Scheres, R. Menéndez-Conejero, I. P. Dmitriev, D. T. Curiel, W. F. Mangel, S. J. Flint, and C. San Martn. Structure and uncoating of immature adenovirus. *J. Mol. Biol.*, 392:547–557, 2009.

[34] P. Zhu, H. Winkler, E. Chertova, K. A. Taylor, and K. H. Roux. Cryoelectron tomography of HIV-1 envelope spikes: Further evidence for tripod-like legs. *PLoS Pathog.*, 4:1–9, 2008.

[35] I. Daubechies, M. Defrise, and C. DeMol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1457, 2004.

[36] M. Hohn, G. Tang, G. Goodyear, P. R. Baldwin, Z. Huang, P. A. Penczek, C. Yang, R. M. Glaeser, P. D. Adams, and S. J. Ludtke. SPARX, a new environment for Cryo-EM image processing. *J. Struct. Biol.*, 157(1):47–55, 2007.

[37] P. A. Penczek, R. Renka, and H. Schomberg. Gridding-based three-dimensional image reconstruction from projections with arbitrary directions. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, pages 936–939, 2004.

[38] F. Natterer and F. Wübbeling. *Mathematical methods in image reconstruction.* SIAM, Philadelphia, 2001.

[39] R. N. Bracewell and A. C. Riddle. Inversion of fan-beam scans in radio astronomy. *Astrophy. J.*, 150:427–434, 1967.

[40] G. N. Ramachandran and A. V. Lakshminarayanan. Three-dimensional reconstruction from radiographs and electron micrographs: Application of convolutions instead of Fourier transform. *Proc. Nat. Acad. Sci.*, 68:2236–2240, 1971.

[41] M. Radermacher. The three-dimensional reconstruction of single particles from random and non-random tilt series. *J. Electron Microsc. Tech.*, 9:359–394, 1988.

[42] M. Radermacher. Weighted back-projection methods. In J. Frank, editor, *Electron Tomography*, pages 91–115. Plenum Press, New York, 1992.

[43] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *J. Theor. Biol.*, 29:471–481, 1970.

[44] G.T. Herman and L.B. Meyer. Algebraic reconstruction techniques can be made computationally efficient [positron emission tomography application]. *IEEE Trans. Med. Imaging*, 12(3):600–609, 1993.

[45] R. Marabini, G. T. Herman, and J. M. Carazo. 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy*, 72:53–65, 1998.
[46] P. Gilbert. Iterative methods for the three-dimensional reconstruction of an object from projections. *J. Theor. Biol.*, 36:105–117, 1972.
[47] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm. *Ultrason Imaging*, 6:81–94, 1984.
[48] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena 60.1 (1992): 259-268.*, 60(1):259–268, 1992.
[49] C. Vogel and M. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17(1):227–238, 1996.
[50] T. Chan, G. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20(6):1964–1977, 1999.
[51] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20(1):89–97, 2004.
[52] E. Y. Sidky, C. Kao, and X. Pan. Accurate image reconstruction from fewviews and limited-angle data in divergent-beam CT. *J. X-Ray Sci. Technol.*, 14:119–139, 2006.
[53] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.*, 53:4777–4807, 2008.
[54] M. Li, G. Xu, C. O.S. Sorzano, F. Sun, and C. L. Bajaj. Single-particle reconstruction using L^2-gradient flow. *J. Struct. Biol.*, 176:259–267, 2011.
[55] G. Xu, M. Li, A. Gopinath, and C. L. Bajaj. Inversion of electron tomography images using L^2-gradient flows - computational methods. *J. Comput. Math.*, 29:501–525, 2011.
[56] J. F. Cai, B. Dong, S. Osher, and Z. Shen. Image restoration: total variation, wavelet frames, and beyond. *J. Amer. Math. Soc.*, 25(4):1033–1089, 2012.
[57] I. Daubechies. *Ten lectures on wavelets.* SIAM, Philadelphia, PA, 1992.
[58] C. Vonesch, L. H. Wang, Y. Shkolnisky, and A. Singer. Fast wavelet-based single-particle recostruction in cryo-EM. In *Proc. IEEE Int. Symp. Biomedical Imaging*, pages 1950–1953, 2011.
[59] B. Dong and Z. Shen. *MRA-Based Wavelet Frames and Applications.* IAS Lecture Notes Series, 2010 Summer Program on "The Mathematics of Image Processing", 2010.
[60] P. A. Penczek. Resolution measures in molecular electron microscopy. *Methods. Enzymol.*, 482:73–100, 2010.
[61] C. O. S. Sorzano, R. Marabini, N. Boisset, E. Rietzel, R. Schröder, G. T. Herman, and J. M. Carazo. The effect of overabundant projection directions on 3D reconstruction algorithms. *J. Struct. Biol.*, 133:108–118, 2001.
[62] C. O. S. Sorzano, R. Marabini, G. T. Herman, and J. M. Carazo. Multiobjective algorithm parameter optimization using multivariate statistics in three-dimensional electron microscopy reconstruction. *Pattern Recognition*, 38:2587–2601, 2005.
[63] S. P. Mallick, B. Carragher, C. S. Potter, and D. J. Kriegman. ACE: automated CTF estimation. *Ultramicroscopy*, 104(1):8–29, 2005.
[64] T. R. Shaikh, H. Gao, W. T. Baxter, F. J. Asturias, N. Boisset, A. Leith, and J. Frank. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nature Protocols*, 3:1941–1974, 2008.
[65] A. Bennett, J. Liu, D. V. Ryk, D. Bliss, J. Arthos, R. M. Henderson, and S. Subramaniam. Cryoelectron tomographic analysis of an HIV-neutralizing protein and its complex with native viral gp120. *J. Biol. Chem.*, 282:27754–27759, 2007.

Appendix A. Outline of the proof. We first show how we find the objective functional that the iteration (1.11) minimizes and then give the proof of Theorem 2.1 and 2.2. For this, we note that the iteration (1.11) can be reformulated as the following alternating iteration scheme:

$$\begin{cases} \alpha_{k+1} &= T_\lambda(Wu_k), \\ u_{k+1} &= (I-\mu A^\top A)W^\top \alpha_{k+1} + \mu A^\top g. \end{cases} \tag{A.1}$$

Each step in the above iteration aims at solving a specific minimization problem. Let μ be some positive constant such that $I-\mu A^\top A$ is invertible and let $D := (I-\mu A^\top A)^{-1}$. Then we have

LEMMA A.1. *Let α_{k+1} and u_{k+1} be defined as (A.1), then*

$$\alpha_{k+1} = \arg\min_{\alpha} \{\frac{1}{2}\|Wu_k - \alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1\}, \tag{A.2}$$

$$u_{k+1} = \arg\min_{u} \{\frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2 + \frac{1}{2}\|Wu - \alpha_{k+1}\|_2^2\}, \tag{A.3}$$

where $\|\cdot\|_D$ is the norm defined by $\|x\|_D = (x^\top Dx)^{\frac{1}{2}}$ for any $x \in \mathbb{R}^N$.

Proof. It is known that $T_\lambda(Wu_k)$ is the explicit solution of the minimization (A.2) (see e.g. [14, 22]). Note that the matrix D is positive definite, the objective function in the minimization (A.3) is convex and differentiable.

The minimizer of (A.3), denoted as u', is a zero point of the gradient of the objective function, i.e.

$$0 = \mu A^\top(Au' - g) + \mu^2 A^\top AD(A^\top Au' - A^\top g) + u' - W^\top \alpha_{k+1}. \tag{A.4}$$

Since

$$I + \mu A^\top AD = D,\ I + \mu DA^\top A = D, \tag{A.5}$$

we have then

$$u' = (I - \mu A^\top A)W^\top T_\lambda(Wu_k) + \mu A^\top g\,,$$

which is exactly u_{k+1} in (A.1). □

Lemma A.1 implies that the iteration (1.11) indeed can be viewed as an alternating iteration for minimizing the following functional:

$$\min_{u,\alpha}\ \frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2 + \frac{1}{2}\|Wu - \alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1. \tag{A.6}$$

Or equivalently, the iteration (1.11) is a possible iterative scheme for minimizing the following functional:

$$\min_{u} \frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2 + \Big(\min_{\alpha} \frac{1}{2}\|Wu - \alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1\Big). \tag{A.7}$$

Since the minimizer of the functional $\frac{1}{2}\|Wu - \alpha\|_2^2 + \|\mathrm{diag}(\lambda)\alpha\|_1$ is $T_\lambda(Wu)$, its minimum value can be written as $H_\lambda(Wu)$, the Huber function defined in (2.4), by a direct calculation. Using $H_\lambda(Wu)$ in (A.7), we have the functional (2.6) proposed in Theorem 2.1 as our possible choice of the underlying objective functional:

$$\min_{u}\ \frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2 + H_\lambda(Wu).$$

Thus, the iteration (1.11) is an iterative method for finding the minimizer of the objective functional above. Next, we will show that the iteration (1.11) indeed converges and the limit is a minimizer of the objective functional above.

To show that the sequence generated by the iteration (1.11) converges to a minimizer of the proposed objective functional, we use the convergence analysis of general

proximal forward-backward splitting scheme provided in [22]. Recall, in convex analysis, for any proper, convex, lower semi-continuous function φ which takes values in $(-\infty, +\infty]$, its proximal operator is defined as

$$\text{prox}_{\varphi}(x) := \arg\min_{y} \{\frac{1}{2}\|x-y\|_2^2 + \varphi(y)\}.$$

THEOREM A.2 ([22]). *Consider the minimization problem*

$$\min_{x} \; F_1(x) + F_2(x) \tag{A.8}$$

where F_1 with range in $(-\infty, +\infty]$ is a proper, convex, lower semi-continuous function, and F_2 with range in $\mathbb{R}$ is a convex, differentiable function with $1/b$-Lipschitz continuous gradient. Assume a minimizer of (A.8) exists and $b > 1/2$. Then for any initial guess x_0, the iteration

$$x_{k+1} = \text{prox}_{F_1}(x_k - \nabla F_2(x_k)) \tag{A.9}$$

converges to a minimizer of $F_1 + F_2$.

The rest is to show that iteration (1.11) is (A.9) for the objective functional that we proposed here, i.e. (2.6). For this, rewrite the objective function F in (2.6) as $F_1 + F_2$, where

$$\begin{cases} F_1(u) = \frac{\mu}{2}\|Au - g\|_2^2 + \frac{\mu^2}{2}\|A^\top Au - A^\top g\|_D^2, \\ F_2(u) = H_\lambda(Wu). \end{cases} \tag{A.10}$$

Then F_1 is clearly a proper, convex and differentiable function ranging in $(-\infty, +\infty]$, and F_2 satisfies properties in the following lemma:

LEMMA A.3. *Let F_2 be defined in* (A.10)*. Then F_2 is convex, differentiable with* 1*-Lipschitz continuous gradient.*

Proof. Since F_2 is the sum of $h_{\lambda_i}((Wu)_i), i = 1, \ldots, L$, where the Huber function h_{λ_i} is defined as (2.5), F_2 is convex by the fact that each h_{λ_i} is a convex function, and W is a linear operator. By the definition of h_{λ_i} in (2.5),

$$\nabla h_{\lambda_i}(\alpha_i) = \alpha_i - t_{\lambda_i}(\alpha_i)$$

where t_{λ_i} is the soft-thresholding operator in (1.7). Hence F_2 is differentiable and

$$\begin{aligned} \nabla F_2(u) &= W^\top[\nabla h_{\lambda_1}((Wu)_1), \nabla h_{\lambda_2}((Wu)_2), \ldots, \nabla h_{\lambda_L}((Wu)_L)]^\top \\ &= W^\top(Wu - T_\lambda(Wu)), \end{aligned} \tag{A.11}$$

which implies that F_2 has 1-Lipschitz continuous gradient. ∎

This shows that F in (2.6) is exactly the functional form of (A.8) in Theorem A.2. Next, we show that the iteration scheme (1.11) is also exactly the iteration in Theorem A.2.

LEMMA A.4. *Let F_1 and F_2 be defined as* (A.10)*. Then the iteration* (1.11) *is the same as the iteration (A.9) in Theorem A.2.*

Proof. By using (A.2), (A.3) and definition of proximal function, (1.11) now becomes

$$\begin{aligned} u_{k+1} &= \text{prox}_{F_1}(W^\top \alpha_{k+1}) \\ &= \text{prox}_{F_1}(W^\top T_\lambda(Wu_k)) \\ &= \text{prox}_{F_1}(u_k - W^\top Wu_k + W^\top T_\lambda(Wu_k)) \\ &= \text{prox}_{F_1}(u_k - W^\top(Wu_k - T_\lambda(Wu_k))\,). \end{aligned}$$

With (A.11), then (1.11) becomes

$$u_{k+1} = \text{prox}_{F_1}(u_k - \nabla F_2(u_k)), \tag{A.12}$$

which is exactly the iteration (A.9). □

The remained for proving the convergence of the iteration (1.11) using Theorem A.2 is to show that there exists at least one minimizer for (2.6).

LEMMA A.5. *Let F_1 and F_2 be defined as* (A.10). *The minimization of (2.6) has at least one solution.*

Proof. By proposition 3.1 in [22], $\min_u\{F_1(u) + F_2(u)\}$ has at least one solution if $F_1 + F_2$ is coercive, i.e. $F_1(u) + F_2(u) \to \infty$ whenever $\|u\|_2 \to \infty$. This follows from the facts that $\lambda \in \mathbb{R}^L$ is a positive vector and W has a left inverse. The detailed proof can be modified line by line from that of [21, Lemma 4.3]. □

A similar discussion can lead to the proof of Theorem 2.2. Let $\{u_k\}$ be the sequence generated by (1.11). Theorem 2.2 follows after we show that the sequence $\{\alpha_{k+1} := T_\lambda(Wu_k)\}$ as defined in (A.1) converges to a minimizer of the balanced approach (2.7). The objective function (2.7) in Theorem 2.2 can be written as $F_1 + F_2$ where

$$\begin{cases} F_1(\alpha) = \|\operatorname{diag}(\lambda)\alpha\|_1, \\ F_2(\alpha) = \frac{1}{2}\|(I - WW^\top)\alpha\|_2^2 + \frac{\mu}{2}\|AW^\top\alpha - g\|_2^2. \end{cases} \tag{A.13}$$

Clearly, the function F_1 with range in $\mathbb{R}$ is proper, convex, lower semi-continuous and coercive; the quadratic function F_2 is convex and differentiable with 1-Lipschitz continuous gradient. Moreover, one can show as Section 4.2 of [21] with a proper modification that

$$\alpha_{k+1} = \text{prox}_{F_1}(\alpha_k - \nabla F_2(\alpha_k)).$$

By Theorem A.2, we conclude that the sequence $\{\alpha_{k+1}\}$ converges to a minimizer of the balanced model (2.7) which implies Theorem 2.2.