A Bayesian adaptive basis algorithm for single particle reconstruction

Alp Kucukelbir [a,*], Fred J. Sigworth [a,b], Hemant D. Tagare [a,c]

[a] Department of Biomedical Engineering, Yale University, New Haven, CT 06520, United States
[b] Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520, United States
[c] Department of Diagnostic Radiology, Yale University, New Haven, CT 06520, United States

ARTICLE INFO

ABSTRACT

Traditional single particle reconstruction methods use either the Fourier or the delta function basis to represent the particle density map. This paper proposes a more flexible algorithm that adaptively chooses the basis based on the data. Because the basis adapts to the data, the reconstruction resolution and signal-to-noise ratio (SNR) is improved compared to a reconstruction with a fixed basis. Moreover, the algorithm automatically masks the particle, thereby separating it from the background. This eliminates the need for ad hoc filtering or masking in the refinement loop. The algorithm is formulated in a Bayesian maximum-a-posteriori framework and uses an efficient optimization algorithm for the maximization. Evaluations using simulated and actual cryogenic electron microscopy data show resolution and SNR improvements as well as the effective masking of particle from background.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Fixed versus adaptive basis reconstruction

All single particle reconstruction (SPR) algorithms, explicitly or implicitly, use a *basis* to represent the three-dimensional (3D) particle density map. 'Basis' is a mathematical term which refers to a set of functions, such as sine and cosine harmonics, whose weighted sum represents a density map. The weights in the sum are called the *coefficients* of the density map.

Classical SPR algorithms use either the Fourier basis or the delta function basis, both of which are *fixed* bases. In contrast, this paper proposes an SPR algorithm that works by *adaptively* choosing a basis. The basis is selected from a collection of bases; the collection is called a *frame*. The algorithm is adaptive in that the basis it chooses depends on the data. Basis adaptivity improves the reconstruction resolution and signal-to-noise ratio (SNR) in comparison to reconstruction with a fixed basis.

The key idea behind this—that adapting the basis improves the SNR—is an important discovery of modern signal processing (Mallat, 1999; Elad, 2010). Yet it is largely unexplored in single particle reconstruction. An explanation of why adaptivity improves SNR is given in Section 2. This section is meant to be explanatory and is primarily included for readers who may be unfamiliar with the concept.

Adaptive basis selection works in a Bayesian framework. The Bayesian framework requires a prior and a parameter to balance the effect of the prior with that of the data. We use a *sparsity prior* (Elad, 2010) and a data-adaptive method (Wainwright, 2009) to set this parameter, both explained in Section 2.

The algorithm in this paper can be used with any frame, but a specific frame is necessary for implementation. After preliminary analysis with a number of frames (including frames containing the Fourier basis), we discovered that a multi-resolution, stationary scaling function and wavelet frame is well suited for SPR. The bases in this frame have finite spatial support and using this frame corresponds to assuming that the particle has a finite, but unknown, spatial extent in the reconstructed volume. Reconstruction with this frame suppresses background and automatically masks the particle. Masking is critical in combating particle over-refinement and is discussed below.

1.2. Relation to previous reconstruction algorithms

Popular electron cryo-microscopy (cryo-EM) packages such as EMAN, SPIDER, and FREALIGN (Ludtke et al., 1999; Shaikh et al., 2008; Grigorieff, 2007) use the Fourier basis to exploit the Fourier slice theorem for fast reconstruction. In a variation on the Fourier basis, spherical harmonics are used in (Yin et al., 2001) for fast 3D rotation.

* Corresponding author. Address: 333 Cedar Street, P.O. Box 208042, New Haven, CT 06520-8042, United States. Fax: +1 203 737 4273.
E-mail address: alp.kucukelbir@yale.edu (A. Kucukelbir).

Algebraic reconstruction techniques (ARTs) use the delta function basis, with the delta functions (Herman, 2009) or spherically symmetric blobs (Marabini et al., 1998; Bilbao-Castro et al., 2009) located at the centers of the voxels. A recent approach uses a fixed wavelet basis (Vonesch et al., 2011). To our knowledge, the idea of adaptively using bases to improve SNR has not been explored in single particle reconstruction.

Turning to the priors used in Bayesian formulations of SPR, we note that (Jaitly et al., 2010) uses a regularizing 'smoothing' prior. The effect of such smoothing priors is well understood: using them is mathematically equivalent to filtering the reconstruction with a low-pass filter (Wahba, 1980; Poggio et al., 1985). The parameter that balances the prior is set in an ad hoc manner.

The Bayesian formulation in (Scheres, 2011) uses a more sophisticated smoothing prior. This prior parametrically models the Fourier spectrum of the density map. The parameters of the prior and the density map are simultaneously updated in an iterative algorithm. A data-adaptive method is used to set the prior balancing parameter. As noted in (Scheres, 2011), this approach shares some similarities with a Wiener filtering approach to regularizing the density map.

The sparsity prior used in our algorithm has a very different effect than smoothing or Wiener filtering. The sparsity prior suppresses the background in the reconstruction without excessively smoothing the reconstructed particle. Suppressing the background prevents particle over-refinement which occurs when noise outside the particle gets iteratively aligned and appears as a part of the reconstructed particle.

Over-refinement is often prevented by manually masking the particle to suppress the background (Joyeux and Penczek, 2002; Sindelar and Grigorieff, 2011). Low pass filtering the particle and the images during alignment (Frank, 2006) is another frequently used method to combat over-refinement. A recently reported method uses mutual information (Shatsky et al., 2009). All of these methods require some algorithm parameters (the masking threshold, filter bandwidth, etc.) to be set manually. In contrast, the sparsity prior masks the particle automatically, adjusting the mask to the particle and the noise level.

Finally, we mention the maximum-likelihood (ML) approach to single particle reconstruction (Sigworth, 1998; Scheres et al., 2005, 2009; Tagare et al., 2008, 2010). This approach is closely related to the Bayesian approach, but does not incorporate priors.

2. Bases, frames, and sparsity: a brief tutorial

Changing the basis influences the noise and signal spectrum in different ways and this difference can be exploited to improve the SNR of the reconstruction. The argument is as follows:

1. White noise has a flat spectrum of the same magnitude in *any* orthonormal basis. This spectrum is sometimes referred to as the noise floor. Any signal component that is below the noise floor is difficult to reconstruct.
2. In contrast to noise, signal spectra are *not* flat. The shape of their spectra depends strongly on the basis. Fig. 1 shows the (hypothetical) spectrum of the same signal in two different bases. The signal spectrum in Fig. 1(a) is relatively flat, with a large part of the spectrum lying below the noise floor. The signal spectrum in Fig. 1(b) is highly peaked with most of the spectrum rising above the noise floor. The basis in Fig. 1(b) is preferred for reconstruction because it has a smaller portion of the signal below the noise floor.

The above argument can be made more precise by recalling a property of orthonormal bases. Whatever its shape, the signal
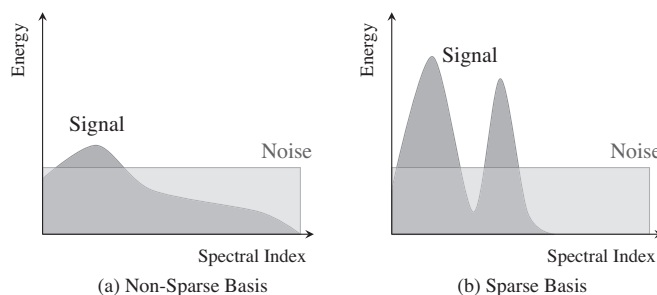

Fig. 1. Spectra of a hypothetical signal and noise shown in two different bases. The spectral index (e.g. 'frequency' for the Fourier basis) depends on the basis. Note the amount of the signal lying under the noise floor in the two cases. The higher the signal level with respect to the noise floor, the more reliably it can be recovered.

spectrum in any orthonormal basis conserves energy (the energy in the signal spectrum equals the signal energy). When the signal spectrum is peaky, most of this energy falls into a few large coefficients while the rest of the coefficients have very small values. That is, "peaky" signal representation is *sparse*. Given a choice of basis, we prefer the basis in which the signal has the sparsest representation, because in this basis most of the signal energy is compressed into a few strong peaks and can be reliably recovered.

How sparse are particle density maps in the classical Fourier and delta function bases? Consider Fourier first. Particle density maps are spatially compact. Fourier spectra of spatially compact signals are spread over much of the Fourier domain due to the uncertainty principle. Thus, particle density maps do not have a sparse representation in the Fourier basis.

Delta function bases are spatially compact, but they are problematic because they exist at too fine a spatial scale. Particle density maps at less than atomic resolution have smooth features that span many voxels and spread over a cluster of delta functions. This is not a sparse representation either.

The above suggests that a good basis for particle density maps should be capable of representing spatially compact maps with a few coefficients at many different resolutions. Scaling functions and wavelets provide just such a basis—they are spatially compact, they exist at a variety of different spatial scales, and they have been theoretically proven to represent piecewise-continuous signals sparsely (Donoho and Johnstone, 1994).

Since sparser bases are better, it is natural to consider many bases simultaneously and to adaptively choose the sparsest basis depending on the signal. Such a collection of bases is called a *frame*. Frames have led to further performance gains in many recent signal processing applications when compared to fixed bases (Starck et al., 2010).

The mechanism for adaptively switching between bases fits well within a Bayesian framework, specifically the maximum-a-posteriori (MAP) estimation framework. The MAP framework has two components: a conditional term that expresses the relation between the basis coefficients and the data, and a prior term that prefers coefficients satisfying certain properties. To switch between bases we use the *sparsity prior* which prefers a few large coefficients, with the rest being exactly or close to zero. This prior is expressed as a multi-variate Laplacian prior. A large mathematical literature justifies this form (e.g. Tibshirani, 1996; Wainwright, 2009; Elad, 2010).

3. Mathematical formulation

3.1. Cryo-EM image formation

The particle density map is an $L \times L \times L$ voxel array S, which we simply refer to as the *particle*. The particle S is projected along P

different directions $p = 1, \ldots, P$. The projection operator along the p-th direction is Ω_p, so that the projected image is $\Omega_p S$. The projection is filtered by the contrast transfer function (CTF) of the microscope, which can vary from image to image. Let C_k denote the CTF filtering operator, where $k = 1, \ldots, K$. The CTF-filtered projected image is thus $C_k \Omega_p S$.

Suppose that $\mathbf{x}_n$ with $n = 1, 2, \ldots, N$ are N particle images picked from micrographs. Each image is a noisy, translated, and rotated copy of some $C_k \Omega_p S$ and is associated with the projection direction p and CTF k. Let g be a function that associates the n-th image with projection direction p, so that $p = g(n)$. Similarly, let h be a function that associates the n-th image with CTF k, so that $k = h(n)$. Further, let $\tau_n = (\vec{t}_n, \gamma_n)$ be the translation $\vec{t}_n = (t_n^x, t_n^y)$ and the rotation γ_n parameters of an image transformation operator T_{τ_n}. Then, the image $\mathbf{x}_n$ is given by

$$T_{\tau_n} \mathbf{x}_n = C_{h(n)} \Omega_{g(n)} S + \eta, \tag{1}$$

where η is noise. While τ_n and $g(n)$ are unknown and have to be estimated along with S from the data, $h(n)$ is assumed to be known. (The CTF parameters for each image are often estimated from their corresponding micrographs prior to particle picking.) Formulae for the operators T_{τ_n}, $C_{h(n)}$ and $\Omega_{g(n)}$ are given in Appendix A.

Assuming that the noise η is i.i.d., white, and Gaussian with a standard deviation σ, and using $\theta = (\tau_1, \ldots, \tau_N, g)$ to denote the entire set of image formation parameters, the log-likelihood of all images $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ given θ and the particle S is

$$\log p(\mathbf{x}|S, \theta) \propto -\sum_{n=1}^{N} \left\{ \frac{\|T_{\tau_n} \mathbf{x}_n - C_{h(n)} \Omega_{g(n)} S\|^2}{2\sigma^2} \right\}, \tag{2}$$

where $\propto$ means 'proportional to' and $\|\cdot\|^2$ is the sum of pixel-values squared.

3.2. A Frame for the particle density map S

Let $\phi_1, \ldots, \phi_B$ be a collection of basis vectors for the density map S. These vectors form a frame, so that S can be written as $S = \sum_{b=1}^{B} \phi_b \alpha_b$, where α_b are the coefficients of expansion. Letting $\alpha = (\alpha_1, \ldots, \alpha_B)^\top$ denote the column vector of the coefficients, the density map S can be compactly expressed as $S = \Phi\alpha$ where Φ is the operator that takes α as input and produces the density map according to $S = \sum_{b=1}^{B} \phi_b \alpha_b$.

For the reconstructions in this paper, we use a two-level, stationary wavelet transform-based frame. We take a brief aside to explain this frame and its relation to the standard wavelet basis. The basis functions included in our frame are illustrated in Fig. 2. At the center of each voxel is a grid point. Conceptually, each grid point is surrounded by two cubes, one twice as large as the other

(in our case, 18 and 36 voxels). Each cube contains eight basis functions, formed as a product of three functions ϕ_x, ϕ_y, and ϕ_z along the sides of the cube. The functions ϕ_x, ϕ_y, and ϕ_z are either Coiflet scaling functions or wavelets (see Fig. 2(b)). The two levels of eight basis functions centered at every grid point give a set of $16 \times L^3$ functions which forms the frame.

Our frame is related to, but quite different from, the standard wavelet basis. The difference is twofold. First, our frame uses scaling functions in the larger and smaller cubes; the standard wavelet basis uses scaling functions only in the larger cube. Second, in our frame, the cubes are centered at every grid point. In the standard wavelet basis the smaller cube is centered only at every *second* grid point and the larger cube only at every *fourth* grid point. Centering both cubes at every grid point is what makes our frame *stationary*.

The design of a frame is a compromise between including many basis functions to improve sparsity and managing its size (and computational complexity). Some preliminary analysis is usually necessary to determine a good compromise (Starck et al., 2010). Our analysis showed that just using a standard wavelet basis was not very sparse; the frame described in Fig. 2 admitted a sparser representation. Further, adding the Fourier basis to the frame did not improve sparsity; the Fourier coefficients were not able to capture any part of the signal sparsely. Based on this analysis, we did not include the Fourier basis in our frame.

Returning from the aside back to our mathematical formulation, Eq. (2) can be expressed in terms of the coefficients α of the frame as

$$\log p(\mathbf{x}|\alpha, \theta) \propto -\sum_{n=1}^{N} \left\{ \frac{\|T_{\tau_n} \mathbf{x}_n - C_{h(n)} \Omega_{g(n)} \Phi\alpha\|^2}{2\sigma^2} \right\}. \tag{3}$$

Estimating the density map S is equivalent to estimating the coefficients α.

3.3. MAP estimation framework

MAP estimates of α and the parameters θ require a prior density $p(\alpha, \theta)$. Assuming that the prior densities of α and θ are independent gives $p(\alpha, \theta) = p(\alpha)p(\theta)$, and the MAP estimates are obtained as

$$\widehat{\alpha}, \widehat{\theta} = \arg\max_{\alpha, \theta} \{\log p(\mathbf{x}|\alpha, \theta) + \log p(\alpha) + \log p(\theta)\}. \tag{4}$$

The function in the curly braces in Eq. (4) is the *MAP objective function*. The MAP estimates $\widehat{\alpha}$ and $\widehat{\theta}$ are the values that maximize it. Since only the maximizing values of the argument are of interest, it is customary to drop all terms in the MAP objective function that are independent or are constant functions of the argument. To proceed, we specify the priors $p(\alpha)$ and $p(\theta)$.
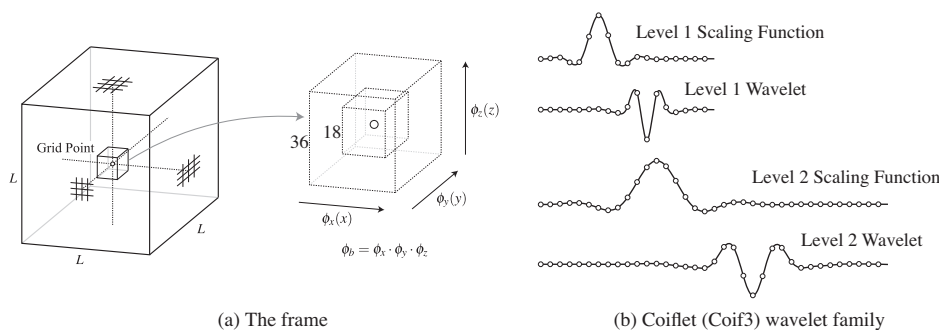

(a) The frame   (b) Coiflet (Coif3) wavelet family

Fig. 2. The two-resolution, stationary wavelet transform-based frame used to represent the particle S in the reconstruction. Every voxel in S is a grid point. Dotted lines in (a) orient the position of a grid point in 3D. The basis functions ϕ_b exist at two spatial scales (Level 1 and 2) centered around each grid point. The inner cube represents Level 1 (16 voxels) and the outer cube Level 2 (36 voxels). At every level, eight basis functions are formed by multiplying either a scaling function or a wavelet for each of the ϕ_x, ϕ_y, and ϕ_z. Hence, for every voxel in S there are sixteen coefficients in the frame. (b) The Coiflet scaling functions and wavelets with grid points denoted by white circles.

3.4. The sparsity promoting prior

A large body of signal processing literature shows that the Laplacian prior

$$p(\alpha) \propto \exp(-\lambda\|\alpha\|_1) \tag{5}$$

promotes sparsity (Mallat, 1999; Elad, 2010). In Eq. (5), λ is a non-negative constant and $\|\alpha\|_1 = \sum_{b=1}^{B}|\alpha_b|$ is the ℓ_1-norm of α. MAP estimates of α with the Laplacian prior have the property that many of the α_b's in the maximizing α are zero (Tibshirani, 1996). Which α_b's are set to zero depends on the data, and this gives the algorithm the means to adaptively choose a basis from the frame since only those functions ϕ_b whose coefficients α_b are not zero enter into the expression for the density map S.

3.5. The prior for θ

The vector θ contains translation and alignment parameters of the images as well as the function g. We assume that the priors for all of these are independent so that $p(\theta) = \prod_{n=1}^{N}\{p(tx_n^N, t_n^y) p(\gamma_n)\}p(g)$. Further, we assume that these priors are non-informative in that they do not prefer any one value over another. Specifically, $p(t_n^x, t_n^y)$ is uniform in a square, $p(\gamma_n)$ is uniform over all angles, and $p(g)$ is equal for all possible g's. This means that these priors are constant functions of their arguments and can be dropped from the MAP objective function.

3.6. The MAP objective function

With the above priors (and after dropping all constant terms), the MAP objective function becomes

$$\log p(\mathbf{x}|\alpha,\theta) + \log p(\alpha) \propto -\sum_{n=1}^{N}\left\{\frac{\|T_{\tau_n}\mathbf{x}_n - C_{h(n)}\Omega_{g(n)}\Phi\alpha\|^2}{2\sigma^2}\right\} - \lambda\|\alpha\|_1. \tag{6}$$

Maximizing this objective function with respect to α and θ gives the MAP estimates. After convergence, the value $\hat{\alpha}$ that maximizes the MAP objective function is available and the MAP estimate of the particle density map is calculated as $\hat{S} = \sum_{b=1}^{B}\phi_b\hat{\alpha}_b$.

3.7. Iterative maximization/refinement

Eq. (6) is maximized in an iterative fashion starting with an initial guess of α which is a coarse initial density map. Each iteration is composed of three steps. The first step maximizes the objective function with respect to θ for a fixed α. This is the *alignment* step where the images are aligned to the projected particle. The second step uses θ from the alignment step to obtain an appropriate value for λ. The third step maximizes the objective function with respect to α for a fixed θ, which is the *reconstruction* step. The three steps are sometimes referred to as the 'refinement loop' and the process of iterating until convergence as 'refinement.'

3.7.1. Alignment

Using the current (or initial) value of α, we maximize the objective function with respect to θ. Since α is held fixed in this step, the second term on the right hand side of Eq. (6) can be dropped from the objective function, giving the simplified objective function

$$-\sum_{n=1}^{N}\left\{\frac{\|T_{\tau_n}\mathbf{x}_n - C_{h(n)}\Omega_{g(n)}\Phi\alpha\|^2}{2\sigma^2}\right\}. \tag{7}$$

Maximizing this objective function with respect to $\theta = (\tau_1, \ldots, \tau_n, g)$ is the classical cryo-EM alignment step. In the results reported in Section 5, this maximization is carried out using an alignment

routine in the SPIDER software package. Any other implementation of the alignment step can also be used.

The maximization in the next two steps can be simplified by some algebraic manipulations. Let $\mathcal{N}_{p,k}$ be the set containing indexes of all images with projection direction p and CTF k. Let $|\mathcal{N}_{p,k}|$ be the number of images in the set. Then the class mean $\mu_{p,k}$ is the mean of all aligned images with indexes in $\mathcal{N}_{p,k}$ as

$$\mu_{p,k} = \frac{1}{|\mathcal{N}_{p,k}|}\sum_{n\in\mathcal{N}_{p,k}}T_{\tau_n}\mathbf{x}_n. \tag{8}$$

After some algebraic manipulations, Eq. (6) can be rewritten using the class mean as

$$\log p(\mathbf{x}|\alpha,\theta) + \log p(\alpha) \propto -\sum_{p,k}\left\{\sum_{n\in\mathcal{N}_{p,k}}\frac{\|T_{\tau_n}\mathbf{x}_n - \mu_{p,k}\|^2}{2\sigma^2}\right\} - \sum_{p,k}|\mathcal{N}_{p,k}|\frac{\|\mu_{p,k} - C_k\Omega_p\Phi\alpha\|^2}{2\sigma^2} - \lambda\|\alpha\|_1. \tag{9}$$

3.7.2. λ Estimation

Mathematical analysis shows that the appropriate value of λ depends on the extent of noise in the class means (Wainwright, 2009), and is determined by estimating the noise standard deviation in the class means and then using this estimate to calculate λ.

Assuming that small patches in the corners of the class mean images contain mostly noise, we concatenate these corner patches into a long vector κ. Then, the median absolute deviation (MAD) can be used to estimate $\hat{\sigma}$ directly from κ as

$$\hat{\sigma} = 1.4826 \times \text{median}(|\kappa - \text{median}(\kappa)|), \tag{10}$$

where the 1.4826 factor arises from the assumption that the noise is Gaussian (Hoaglin et al., 1983). MAD is preferable over the more common root-mean-square deviation (RMSD) because it is robust against outliers in the data, whereas RMSD is not. The parameter λ is calculated from $\hat{\sigma}$ by a standard formula (Wainwright, 2009):

$$\lambda = 3\left(2\sqrt{2}\hat{\sigma}\sqrt{2\log L^3}\right), \tag{11}$$

where L is the length of the particle along a single dimension. The scaling constant 3 in the above equation is necessary since our frame is not orthonormal and the projection operator Ω_p does not preserve norms (see Chapter 7 of Starck et al., 2010).

3.7.3. Reconstruction

Fixing the values of θ and λ from the previous steps, we maximize the objective function with respect to α. The first term on the right hand side of Eq. (9) is independent of α and can be dropped for this step, reducing the objective function to

$$-\sum_{p,k}|\mathcal{N}_{p,k}|\frac{\|\mu_{p,k} - C_k\Omega_p\Phi\alpha\|^2}{2\sigma^2} - \lambda\|\alpha\|_1. \tag{12}$$

Maximizing this simplified objective function with respect to α is carried out using Nesterov's algorithm (Nesterov, 2007), which is presented in detail in Appendix B.

Iterating the alignment, λ estimation, and reconstruction steps till convergence gives the adaptive basis reconstruction.

4. Methods

4.1. Simulated data

We used the atomic structure of the 50S ribosomal subunit from the Protein Data Bank (PDB ID: 1JJ2) along with a model for the water shell surrounding the particle (Shang and Sigworth, 2012) to create a $(3 \text{ Å})^3$ voxel density map in a $100 \times 100 \times 100$ voxel array. This is referred to below as the *true density map* S_{true}. The true density map was projected from 900 almost equally spaced directions in the northern hemisphere and projection images were created using the image formation model of Eq. (1). The CTF parameters were based on an electron energy of 120 keV, a defocus value drawn randomly from $-1.5 \pm 1.2 \, \mu\text{m}$, and an amplitude decay (B-factor) of 150 Å^2. These values were chosen to be similar to the CTF parameters in the National Resource for Automated Molecular Microscopy (NRAMM) 50S dataset, described in Section 4.2. We added sufficient white Gaussian noise to the images to simulate the noise present in class means calculated from approximately 5, 10, and 50 images of the NRAMM dataset per projection direction. This corresponds to SNR levels of approximately -3, 0, and 6 dB. Typical noisy class means from this simulation are shown in Fig. 3.

Our reconstruction algorithm (as well as others) requires knowledge of the CTF. In practice the CTF is not available exactly, and to simulate this, the reconstructions were carried out by setting the CTF defocus parameter of the reconstruction algorithm to the true defocus parameter perturbed by a deviation in the range $\pm 0.12 \, \mu\text{m}$.

4.2. Experimental data

The cryo-EM dataset for the 50S ribosomal subunit is from (Voss et al., 2010) and is available online. This dataset contains about 82000 images and their CTF parameters. We randomly picked a subset of approximately 8000 images from three different CTF classes with -1.3, -1.81, and $-2.35 \, \mu\text{m}$ defoci and phase-flipped the images. The class means were pre-whitened to conform to the white noise assumption in the adaptive basis algorithm.

4.3. Refinement loop

No alignment was necessary for the simulated data, since we directly simulated noisy class means from the true density map. For simulated data, we compared the adaptive basis algorithm to an algebraic reconstruction technique (ART). Details for ART are given in Appendix C.

We adopted two separate refinement strategies for the experimental data. First we used classical Fourier back-projection (SPIDER's BP 3F routine) within the refinement loop. In this case, low-pass filtering of intermediate reconstructions was necessary to ensure convergence. The low-pass filter cut-off frequency was set in an ad hoc, yet adaptive fashion as done in (Wang et al., 2007). We then used the final Euler angle assignments to compute class means and compared the adaptive basis algorithm to both ART and Fourier back-projection. In the second case of refinement, we used the adaptive basis method within the refinement loop with no intermediate filtering.

Both methods employed SPIDER's AP SH routine to perform the alignment with a fixed number of 1253 equally spaced projection directions over the northern hemisphere. Ten SPIDER iterations were executed with the initial volume being a 80 Å low-pass filtered version of the 'PDB experiment' model presented in (Voss et al., 2010). Both refinements appeared to converge within eight to nine iterations.

4.4. Performance measures

We measured the performance of the reconstruction algorithms using SNR (defined in Appendix D), Fourier Shell Correlation (FSC)
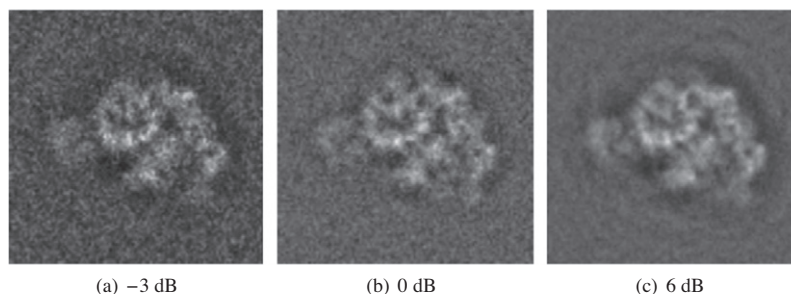

(a) -3 dB   (b) 0 dB   (c) 6 dB

Fig. 3. Class means of simulated data for 50S ribosomal subunit (PDB ID: 1JJ2) at SNRs of -3, 0, and 6 dB.

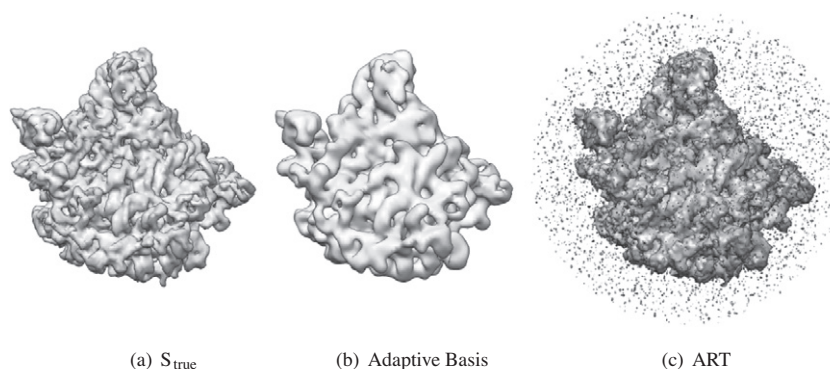(a) S_{true} (b) Adaptive Basis   (c) ART

Fig. 4. Volume renderings of the true density map S_{true}, the adaptive basis, and the ART reconstructions using -3 dB class means from the simulated data. All volumes are displayed at the same threshold. Note the higher SNR and lack of background artifacts in the adaptive basis reconstruction.
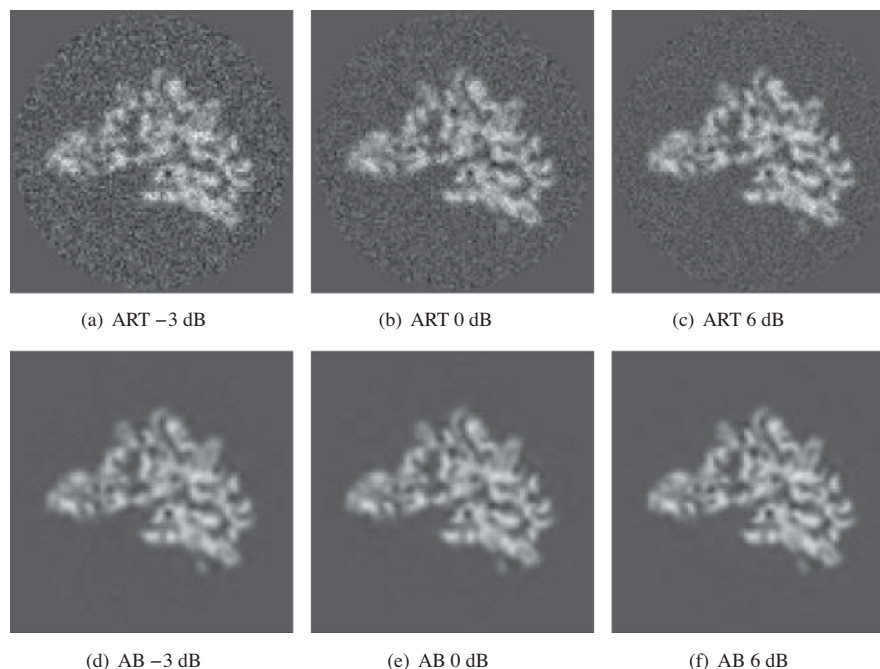
(a) ART −3 dB   (b) ART 0 dB   (c) ART 6 dB

(d) AB −3 dB   (e) AB 0 dB   (f) AB 6 dB

Fig. 5. Central slices from the simulated data reconstructions. All image are displayed using a common grayscale mapping. Noise levels represent reconstructions obtained using −3,0, and 6 dB class means. Adaptive basis reconstructions are denoted as AB. Note the maintained high-frequency details and the suppressed background in the AB slices.

(van Heel and Schatz, 2005), and a background suppression energy ratio metric described below.

As mentioned towards the end of Section 1, the adaptive basis algorithm is expected to suppress the background. To measure the effectiveness of background suppression, we manually create a mask that is unity outside the particle and zero inside. Using this mask, we isolate the voxels outside the particle and place them into a vector $\mathbf{c}$ of length M. The energy per voxel in the background is calculated as

$$E = \frac{1}{M}\sum_{m=1}^{M}\mathbf{c}_m^2,\qquad(13)$$

where higher values of E imply a stronger background in the reconstruction.

To compare the effectiveness of two reconstruction algorithms at suppressing the background, we calculate the ratio of their background energies (as E_1/E_2 where E_1 and E_2 are the energy per voxel of the background of two reconstructions as calculated by Eq. (13)) using the same mask. The mask is created to completely mask the particle in both reconstructions.

5. Results

The adaptive basis algorithm is expected to produce higher SNR and higher resolution reconstructions than traditional methods. The algorithm is also expected to suppress background in the reconstructions, eliminating the need for intermediate filtering in the refinement loop. To test both claims, we evaluated our algorithm using a well-known particle, the 50S ribosomal subunit. We began by testing our algorithm on simulated data which allow for SNR and resolution comparisons to the true signal. We continued by evaluating our method on actual cryo-EM
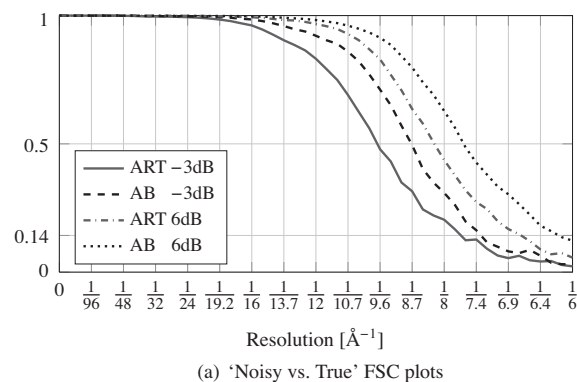
(a) 'Noisy vs. True' FSC plots
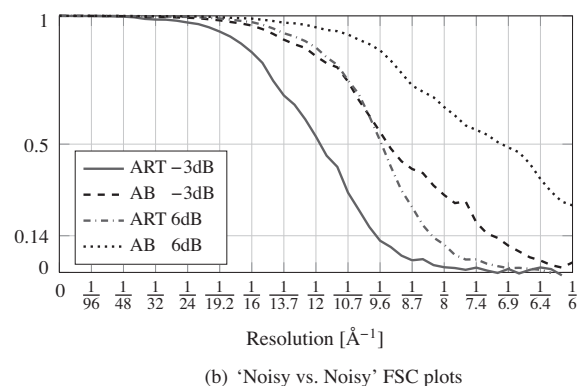
(b) 'Noisy vs. Noisy' FSC plots

Fig. 6. FSC plots of the simulated data reconstructions. Adaptive basis reconstructions, denoted as AB, outperforms ART in all cases. Both sub-figures show FSC plots for reconstructions obtained using −3 and 6 dB class means. Results for 0 dB are omitted for readability.

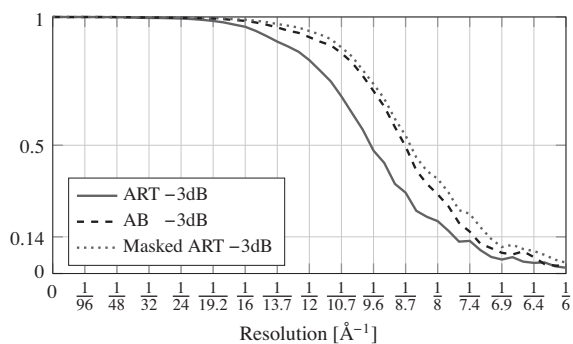images of the 50S subunit from (Voss et al., 2010). More on this in Section 5.2.

Fig. 7. 'Noisy vs. True' FSC plots of the simulated data reconstructions. Adaptive basis reconstruction, denoted as AB, exhibits a very close FSC curve to the manually masked ART reconstruction.

Table 1
SNR, background energy ratios, and runtime measurements for the simulated data reconstructions. Adaptive basis reconstruction, denoted as AB, shows markedly higher SNR measurements at all three noise levels, and exhibits large background energy suppression ratios (E_{ART}/E_{AB}). AB also exhibits comparable execution times to ART. Bold values indicate better performance.

Noise (dB)	SNR (dB)		$\frac{E_{ART}}{E_{AB}}$	Time (s)	
	ART	AB		ART	AB
−3	5.5	**17.6**	605.7	734	**695**
0	10.6	**18.3**	440.1	718	718
6	13.4	**19.6**	373.7	776	**725**

5.1. Reconstructions from simulated data

Reconstructions using the adaptive basis and ART methods for the simulated noisy class means from the 50S ribosomal subunit are displayed in Figs. 4 and 5. Fig. 4 shows volume renderings of the reconstructed particle density using the −3 dB class means. All displayed volumes in Fig. 4 were thresholded at the same value using Chimera (Pettersen et al., 2004).

The adaptive basis reconstruction maintains the overall structure of the true density map without introducing background artifacts. Reconstructions from 0 dB and 6 dB are omitted for brevity in Fig. 4 and are similar to the −3 dB reconstruction. Fig. 5 displays a central slice through the reconstructions and shows that the adaptive basis reconstructions exhibit less overall noise while preserving much of the high frequency information.

The FSC plots of the reconstructions are shown in Fig. 6. Fig. 6(a) shows the FSC of the reconstructions compared to the true density map ('Noisy vs. True'), and Fig. 6(b) shows the FSC of the reconstructions obtained by splitting the noisy data into two halves and reconstructing two density maps ('Noisy vs. Noisy'). In both figures, the adaptive basis reconstructions consistently show a higher resolution (at FSC = 0.5 and at FSC = 0.143) than the ART reconstructions. FSC plots for 0 dB lie between the −3 and 6 dB plots and are omitted.

The improvement in the FSC for the adaptive basis algorithm is largely a result of background suppression. This is demonstrated by masking out the background in the reconstructions and then calculating the FSC. These masked-FSCs are shown in Fig. 7 for the adaptive basis reconstruction with −3 dB data (curves for the 0 and 6 dB data show a similar effect and are omitted). The plots in Fig. 7 were obtained by masking the ART reconstructions with the soft mask shown in Fig. 8(b) to suppress the background and then calculating FSC curves. After masking the background, the FSC of the ART reconstruction is comparable to that of the adaptive basis algorithm (in fact, it is slightly better). The slight improvement in the masked ART reconstruction is likely due to the manual tailoring of the mask which makes the mask tighter and more effectively suppress the background.

Table 1 shows other performance measures for the algorithms: the SNR of the reconstructions (compared to the true density map), the ratios of the background energy per voxel in the two reconstructions, and the execution times for the reconstructions. The binary mask shown in Fig. 8(c) was used to calculate the background energy per voxel.

The adaptive basis SNRs are consistently higher than the SNRs for ART reconstructions. Table 1 also shows a large reduction in the background for adaptive basis as compared to ART reconstructions. This is consistent with the FSC curves of Fig. 8. The execution times of both methods are also comparable.

5.2. Reconstructions from experimental data

There is a subtlety in testing a reconstruction method with experimental data because of the refinement process. It is difficult to assess whether resolution gains are intrinsic to the reconstruction method or due to better alignment (or both). To clarify this, we adopted two different refinement strategies. The reconstructions from the first establish resolution gains intrinsic to our method, while those from the second show our method's performance within the refinement loop.

In the first strategy, we used Fourier back-projection within the refinement loop and saved the Euler angles from the final iteration. To prevent over-refinement, the intermediate reconstructions were low-pass filtered. Using the final Euler angles, we performed three reconstructions using Fourier back-projection, ART, and our adaptive basis method. Performance gains measured between these
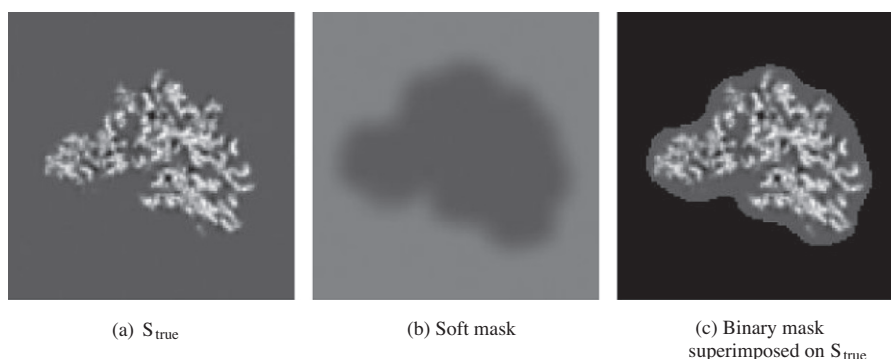

(a) S_{true} (b) Soft mask   (c) Binary mask superimposed on S_{true}

Fig. 8. Soft and hard masks used in evaluating the simulated reconstructions. The soft mask in (b) is used in the 'Masked FSC' plots in Fig. 7, while the binary mask in (c) is used in the background energy ratio calculations. The true signal S_{true} is shown for reference in (a).

(a) NRAMM

(b) Adaptive Basis
ab initio

(c) Adaptive Basis

(d) ART filtered at
FSC = 0.5 (21Å)

(e) ART filtered at
FSC = 0.143 (16Å)

(f) ART filtered at 9Å

(g) Fourier filtered at
FSC = 0.5 (23Å)

(h) Fourier filtered at
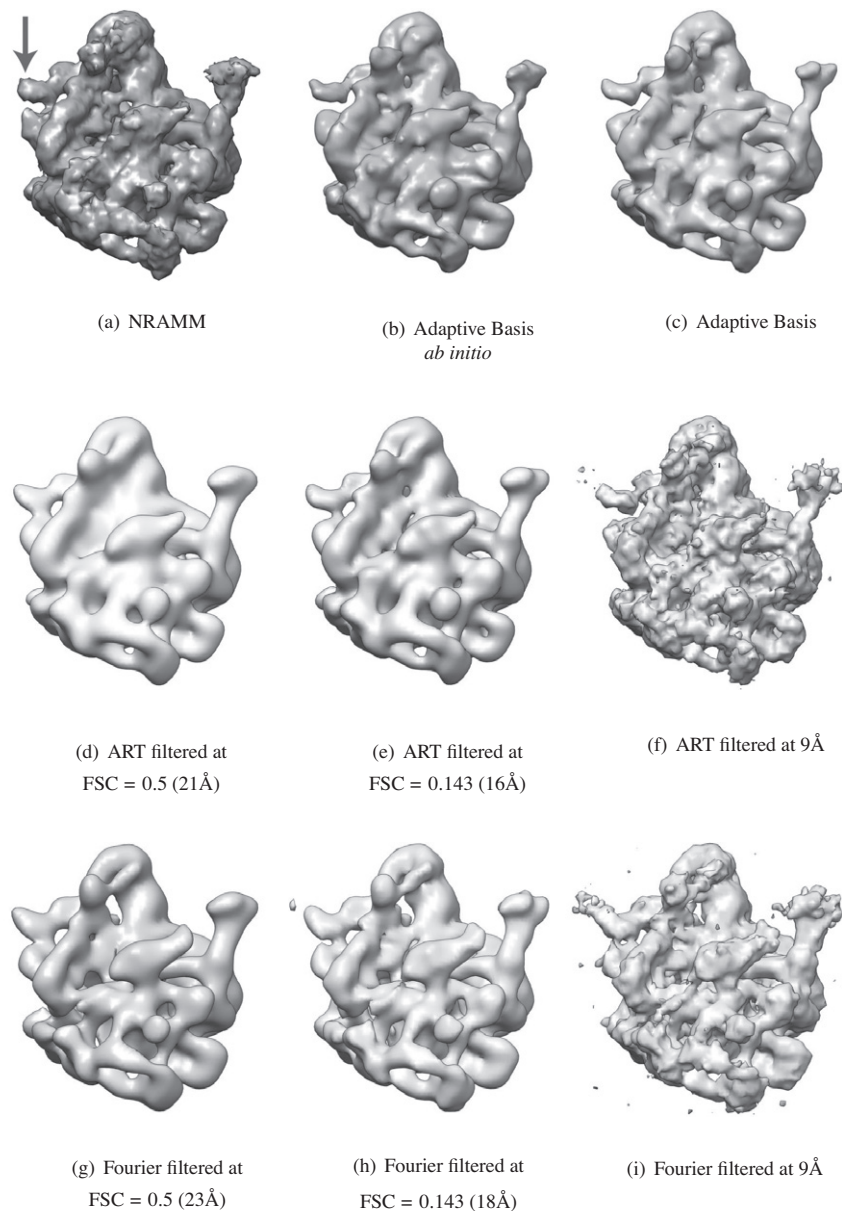FSC = 0.143 (18Å)

(i) Fourier filtered at 9Å

Fig. 9. Volume renderings of the experimental data reconstructions. The adaptive basis reconstructions display higher frequency information than the other reconstructions, without exhibiting background artifacts. This is particularly evident in the region indicated by the arrow in sub-figure (a). All volumes are thresholded at the same value.

three reconstructions are solely representative of their respective reconstruction methods, since everything else is the same.

In the second strategy, we used the adaptive basis algorithm *within* the refinement loop and obtained another adaptive basis reconstruction we call 'AB *ab initio*.' In this case, no filtering was done to the intermediate reconstructions. Therefore, this second strategy provides, in a sense, 'real-world proof' of our method's background suppression effectiveness.

The results of both refinements and corresponding reconstructions are displayed in Figs. 9 and 10. Fig. 9 presents volume renderings of the reconstructed particle density maps. Because ART and Fourier reconstructions were too noisy to render, we low-pass filtered them at the frequencies corresponding to FSC = 0.5 (21–23 Å) and FSC = 0.143 (16–18 Å), as well as at an aggressive 9 Å. The density map reconstructed from the entire NRAMM dataset (using all ≈82,000 images) is published in (Voss et al., 2010) and is available online (nramm.scripps.edu). We denote this map as 'NRAMM,' and

use it as the 'ground truth.' Note that in contrast to the NRAMM map, all other density maps in Figs. 9 and 10 are reconstructed using approximately one tenth of the images. All densities displayed in Fig. 9 are thresholded at the same value using Chimera.

The adaptive basis reconstructions exhibit a higher level of detail than the other reconstructions without showing significant noise artifacts. This is most visible in the top left corner of the particle indicated by the arrow in Fig. 9(a). Detail in that region of the particle is either lost or obscured by noise in the ART and Fourier reconstructions.

Fig. 10 displays a central slice through the reconstructions. The adaptive basis reconstructions show significant suppression of background while maintaining high frequency information. Such detail is lost in the ART and Fourier reconstructions filtered at FSC = 0.5 and FSC = 0.143. Also note the general adherence of the adaptive basis and ART density values to the NRAMM reconstruction, in comparison to the dark undershoots around and within
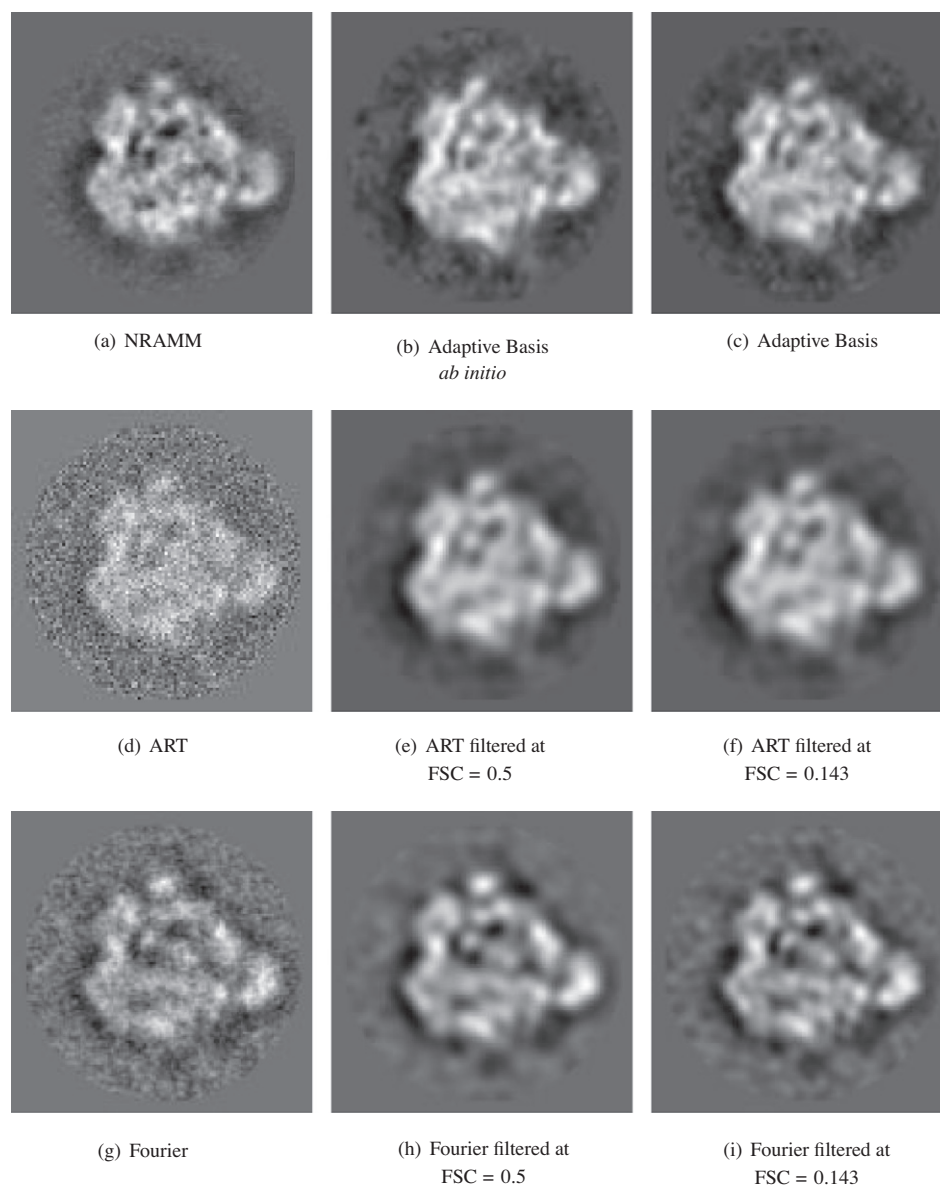
Fig. 10. Central slices from the experimental data reconstructions. The adaptive basis reconstructions show increased background suppression while maintaining high frequency information. The Fourier reconstruction exhibits undershoots within and around the particle.

the particle in the Fourier reconstruction. These undershoots probably arise from the lack of CTF correction in this reconstruction.

The FSC curves of the reconstructions are shown in Fig. 11. Fig. 11(a) presents the FSC of the reconstructions obtained by splitting the data into halves and reconstructing two density maps. Fig. 11(b) displays masked FSC results where a soft mask (based on the downloaded NRAMM reconstruction) is used to suppress the background. In both figures, the adaptive basis reconstructions consistently show a higher resolution (at FSC = 0.5 and at FSC = 0.143) than both ART and Fourier reconstructions. A close look at the FSC curves reveals a resolution improvement of about 2–3 Å at FSC = 0.5 and at least 1 Å at FSC = 0.143. Note that both adaptive basis reconstructions exhibit similar FSC curves.

Table 2 presents the ratios of the background energy per voxel, where a binary mask (also based on the downloaded NRAMM reconstruction) was used. The measurements reported in Table 2 show increased background suppression for both adaptive basis reconstructions as compared to ART and Fourier reconstruction, with and without low-pass filtering at FSC = 0.5 and FSC = 0.143.

This is consistent with the slices presented in Fig. 10 and the FSC curves of Fig. 11.

6. Discussion

The adaptive basis reconstruction algorithm shows improved performance for simulated as well as actual cryo-EM data when compared to ART and Fourier back-projection. By using a frame of spatially compact basis functions, the adaptive basis algorithm simultaneously suppresses background while retaining high frequency details of the particle.

In the experimental data reconstructions, adaptive basis outperforms both Fourier back-projection and ART in terms of resolution and particle masking. Slices, volume renderings, FSC, and background energy ratios all display improvement. The first refinement strategy shows that even with identical Euler angle assignments, the adaptive basis algorithm reconstructs a higher resolution density map. The second refinement strategy demonstrates the adaptive basis algorithm successfully operating within the refinement
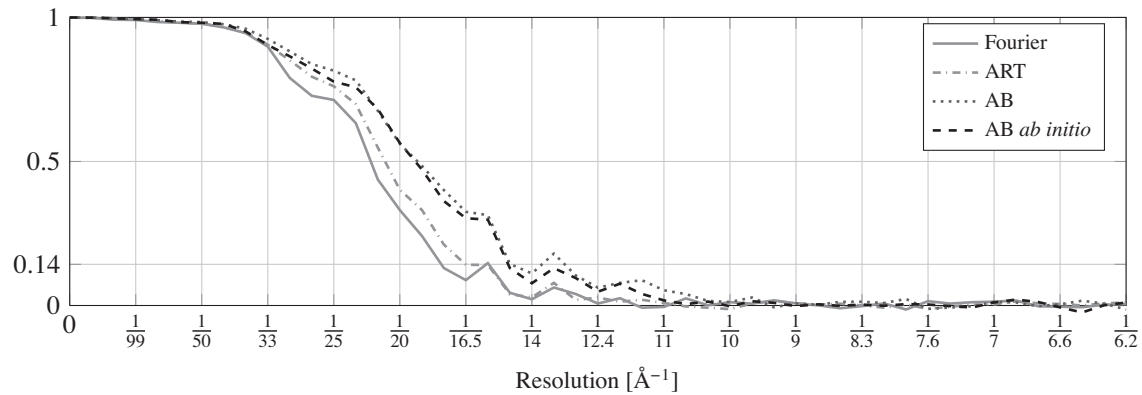
(a) 'Noisy vs. Noisy' FSC plots
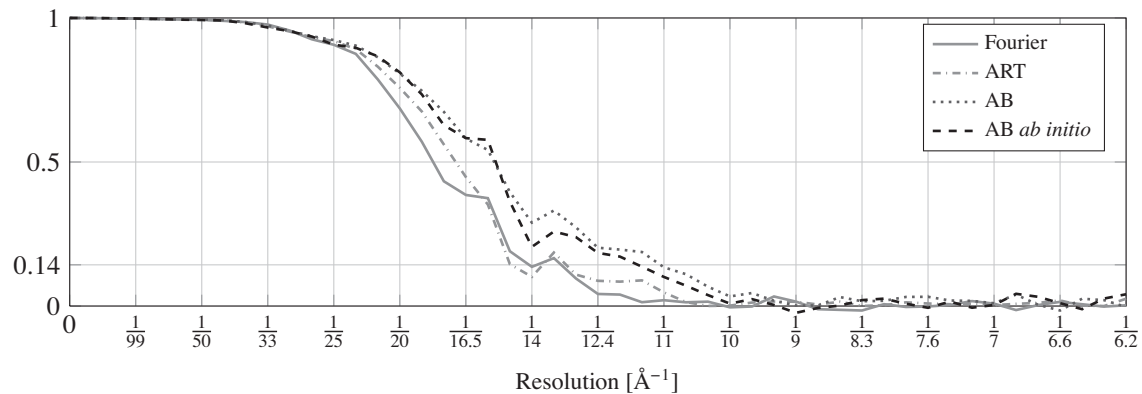
(b) Masked 'Noisy vs. Noisy' FSC plots

Fig. 11. FSC plots of the experimental data reconstructions. Adaptive basis reconstructions, denoted as AB, outperforms ART and Fourier with or without masking the particle. Moreover, both adaptive basis reconstructions exhibit highly similar FSC curves.

Table 2
Background energy ratios for the experimental data reconstructions. Values are reported as the energy ratios of the reconstruction in the row over the reconstruction in the column. (For instance $E_{ART}/E_{AB} = 16.67$.) Adaptive basis reconstructions, denoted as AB, show stronger background suppression than all other reconstructions, even after filtering at FSC = 0.143 and FSC = 0.5.

	Adaptive basis	
	AB	AB *ab initio*
ART		
Unfiltered	16.67	14.72
FSC = 0.143	1.37	1.21
FSC = 0.5	1.22	1.07
Fourier		
Unfiltered	6.52	5.76
FSC = 0.143	1.77	1.53
FSC = 0.5	1.35	1.19

loop, thereby eliminating the need for any intermediary masking or filtering. In both cases, the adaptive basis reconstructions appears to match the NRAMM density map more accurately than the ART and Fourier reconstructions.

There are several ways in which our adaptive basis algorithm may be further improved. A better characterization of noise in cryo-EM images would be useful. The white noise assumption, and the accompanying Fourier pre-whitening as a pre-processing step may be satisfactory when a Fourier basis is used, but our algorithm could benefit from a methodology that characterizes cryo-EM noise in way that can be used with non-Fourier bases. Even

further gains may be obtained by designing and using other frames, especially frames that are tailored to structures commonly found in particles imaged with cryo-EM.

MATLAB code for the adaptive basis algorithm is available online at: mathworks.com/matlabcentral/fileexchange/36040.

Acknowledgments

Appendix A. Image formation model details

The standard two-dimensional (2D) and 3D Cartesian spaces are $\mathbb{R}^2$ and $\mathbb{R}^3$. Points in these spaces are described as $\vec{u} = (u^x, u^y)$ in 2D and $\vec{u} = (u^x, u^y, u^z)$ in 3D. The dimension of the vector is stated explicitly unless it can be inferred from context. Unit coordinate vectors are defined in the standard form of $(\hat{\imath}, \hat{\jmath})$ in $\mathbb{R}^2$ and $(\hat{\imath}, \hat{\jmath}, \hat{k})$ in $\mathbb{R}^3$.

The density map we seek to reconstruct is defined as $S : \mathbb{R}^3 \to \mathbb{R}$. Its value $S(\vec{u})$ gives the electron scattering density of the protein at point $\vec{u} \in \mathbb{R}^3$. S is restricted to a $L \times L \times L$ regular grid with the coordinates at the centers of the voxels.

An image is defined as $\mathbf{x} : \mathbb{R}^2 \to \mathbb{R}$. The value of an image $\mathbf{x}(\vec{u})$ is its intensity at point $\vec{u} \in \mathbb{R}^2$. An image is restricted to a $L \times L$ regular grid with the coordinates at the centers of the pixels.

A set of unit orthonormal vectors $p = \{\hat{\imath}', \hat{\jmath}', \hat{k}'\}$ in 3D forms the projection frame. The ray projection operator Ω_p attached to the projection frame p operates on a particle S to produce an image whose x,y coordinates are aligned along $\hat{\imath}', \hat{\jmath}'$ and whose value $\mathbf{x}(\vec{u})$ at the point $\vec{u} \in \mathbb{R}^2$ is given by

$$\mathbf{x}(\vec{u}) = (\Omega_p S)(\vec{u}) = \int S(u^x \hat{\imath}' + u^y \hat{\jmath}' + z\hat{k}')dz. \tag{A.1}$$

This is often referred to as the *forward projection* operator. Its inverse operates on an image $\mathbf{x}$ to produce a volume by smearing the image along the $\hat{k}$ axis of the projection frame p. Therefore, the value of the volume at $\vec{v} \in \mathbb{R}^3$ is simply

$$\left(\Omega_p^\top \mathbf{x} \right)(\vec{v}) = \mathbf{x}(v^x \hat{\imath}' + v^y \hat{\jmath}'). \tag{A.2}$$

This is often referred to as the *back projection* operator.

The operator C_k simply applies CTF k to the image. In real space this is a convolution with a circularly symmetric kernel, assuming no astigmatism.

The 2D rigid transformation operator T_{τ_n} jointly applies a rotation by γ_n and translation by $\vec{t}_n$ to an image. Specifically, it rotates the image by γ_n counter-clockwise around its origin (center) and translates each point in the image by $\vec{t}_n$.

Appendix B. Nesterov algorithm details

The simplified objective function of Eq. (12) is concave and has a single maximum. However, the objective function is not differentiable (the $\|\alpha\|_1$ term is not differentiable at $\alpha = 0$). Therefore standard gradient-based techniques cannot be used to maximize this objective function. But alternative optimization methods do exist; we use one such method, called Nesterov's method (Nesterov, 2007). It has the advantage that it can optimize the objective function of Eq. (12) while providing a second-order convergence rate to the global maximum.

Nesterov's method is presented in Algorithm 1. Nesterov optimization is designed to *minimize* an objective function. Therefore, to *maximize* the MAP objective function in Eq. (6), we simply take its negative. We can do so because the MAP objective function is concave and has a single maximum, therefore its negative is convex and has a single minimum.

Algorithm 1: Nesterov optimization

Data $\{\mu_{p,k}\},\{C_k\}$, step size δ, stopping rule Eq. (B.5)
Result The reconstruction: $\widehat{S} = \boldsymbol{\Phi}\widehat{\alpha}$
Initialize $\alpha^{(0)} = \mathbf{0}$
Set internal parameters $i = 0, \kappa_0 = 0, \xi^{(0)} = \mathbf{0}$
Estimate λ via Eq. (11)
while *change in S is above stopping threshold* **do**
   First Proximal Computation
      $v^{(i)} \leftarrow \mathrm{prox}_{\kappa_i \lambda \|\alpha\|_1}(\alpha^{(0)} - \xi^{(i)})$
   Auxiliary
      $a_t \leftarrow \frac{\delta + \sqrt{\delta^2 + 4\delta\kappa_i}}{2}, \omega^{(i)} \leftarrow \frac{\kappa_i \alpha^{(i)} + a_i v^{(i)}}{\omega_i + a_i}$
   Second Proximal Computation
      $\alpha^{(i+1)} \leftarrow \mathrm{prox}_{\frac{\delta}{2}\lambda\|\alpha\|_1}(\omega^{(i)} + \frac{\delta}{2}\nabla f_d(\omega^{(i)}))$
   Accumulate Gradient Directions
      $\xi^{(i+1)} \leftarrow \xi^{(i)} - a_i(\nabla f_d(\alpha^{(i+1)}))$
   Auxiliary
      $\kappa_{i+1} \leftarrow \kappa_i + a_i, i \leftarrow i + 1$
end
$\widehat{\alpha} \leftarrow \alpha^{(i)}$

The proximity function for the ℓ_1-norm function is the following operator

$$\mathrm{prox}_{\lambda\|\alpha\|_1}(\alpha) = \left(1 - \frac{\lambda}{|\alpha_b|} \right)_+ \alpha_b, \forall b \tag{B.1}$$

$$(\cdot)_+ = \max(\cdot, 0) \tag{B.2}$$

where Eq. (B.1) is applied to each element b of α independently. It is also known as the *soft thresholding* operator.

The first part of the negative of the objective function in Eq. (6) is

$$f_d(\alpha) = \sum_{p,k} |\mathcal{N}_{p,k}| \frac{\|\mu_{p,k} - C_k \Omega_p \boldsymbol{\Phi}\alpha\|^2}{2\sigma^2} \tag{B.3}$$

and is also known as the *data fidelity* term of the objective function. The gradient of the data fidelity term is

$$\nabla f_d(\alpha) = \boldsymbol{\Phi}^\top \sum_{p,k} |\mathcal{N}_{p,k}| \frac{\Omega_p^\top C_k^\top (\mu_{p,k} - C_k \Omega_p \boldsymbol{\Phi}\alpha)}{2\sigma^2}, \tag{B.4}$$

where $\boldsymbol{\Phi}^\top$ is the *forward* transform of the frame. In our case, it is not explicitly stored, yet a wavelet transform-based algorithm is available to compute both *forward* and *backward* transforms.

We define our stopping threshold as

$$\left(\frac{\|\widehat{S}^{(i)} - \widehat{S}^{(i-1)}\|^2}{\|\widehat{S}^{(i)}\|^2} \right) \leqslant 10^{-4}, \tag{B.5}$$

where $\widehat{S}^{(i)}$ denotes the reconstruction at iteration i of the Nesterov algorithm.

Appendix C. Algebraic reconstruction details

Algebraic reconstruction techniques (ARTs) model the projection in the spatial domain and reconstruct by minimizing an objective function. The ART objective function is

$$\sum_{n=1}^{N} \left\{ \frac{\|T_{\tau_n}\mathbf{x}_n - C_{h(n)}\Omega_{g(n)}S\|^2}{2\sigma^2} \right\} + \lambda\|S\|^2. \tag{C.1}$$

When $\lambda > 0$, the $\lambda\|S\|^2$ term ensures that the ART objective function is strictly convex. We take $\lambda = 50$ which guarantees convexity without affecting the reconstruction significantly. The ART objective function of Eq. (C.1) is easily minimized using the conjugate gradient method, and this is the method we used to obtain ART reconstructions in this paper.

Appendix D. Signal-to-noise ratio details

The SNR of the reconstructed particle $\widehat{S}$ is evaluated by comparing it with the true density map S_{true} as

$$\mathrm{SNR}_{\text{dB}} = 10 \log \left(\frac{\|\widehat{S}\|^2}{\|\widehat{S} - S_{\text{true}}\|^2} \right). \tag{D.1}$$

SNR can only be evaluated for simulated data reconstructions, because the true density is not known for experimental data.

References

Bilbao-Castro, J., Marabini, R., Sorzano, C., García, I., Carazo, J., Fernández, J., 2009. Exploiting desktop supercomputing for three-dimensional electron microscopy reconstructions using art with blobs. Journal of Structural Biology 165, 19–26.

Donoho, D., Johnstone, J., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika 81, 425.

Elad, M., 2010. Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer Verlag.

Frank, J., 2006. Three-dimensional Electron Microscopy of Macromolecular Assemblies. Oxford.

Grigorieff, N., 2007. Frealign: high-resolution refinement of single particle structures. Journal of Structural Biology 157, 117–125.

van Heel, M., Schatz, M., 2005. Fourier shell correlation threshold criteria. Journal of Structural Biology 151, 250–262.

Herman, G., 2009. Fundamentals of Computerized Tomography: Image Reconstruction From Projections. Springer Verlag.

Hoaglin, D., Mosteller, F., Tukey, J., 1983. Understanding Robust and Exploratory Data Analysis, vol. 3. Wiley, New York.

Jaitly, N., Brubaker, M., Rubinstein, J., Lilien, R., 2010. A bayesian method for 3D macromolecular structure inference using class average images from single particle electron microscopy. Bioinformatics 26, 2406.

Joyeux, L., Penczek, P., 2002. Efficiency of 2D alignment methods. Ultramicroscopy 92, 33–46.

Ludtke, S., Baldwin, P., Chiu, W., 1999. Eman: semiautomated software for high-resolution single-particle reconstructions. Journal of Structural Biology 128, 82–97.

Mallat, S., 1999. A Wavelet Tour of Signal Processing. Academic Pr.

Marabini, R., Herman, G., Carazo, J., 1998. 3D Reconstruction in electron microscopy using art with smooth spherically symmetric volume elements (blobs). Ultramicroscopy 72, 53–65.

Nesterov, Y., 2007. Gradient methods for minimizing composite objective function. ReCALL, 76.

Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E., Ferrin, T., 2004. Ucsf chimera—a visualization system for exploratory research and analysis. Journal of Computational Chemistry 25, 1605–1612.

Poggio, T., Torre, V., Koch, C., 1985. Computational vision and regularization theory. Nature 317, 314–319.

Scheres, S., 2011. A bayesian view on cryo-em structure determination. Journal of Molecular Biology.

Scheres, S., Valle, M., Carazo, J., 2005. Fast maximum-likelihood refinement of electron microscopy images. Bioinformatics 21, ii243.

Scheres, S., Valle, M., Grob, P., Nogales, E., Carazo, J., 2009. Maximum likelihood refinement of electron microscopy data with normalization errors. Journal of Structural Biology 166, 234–240.

Shaikh, T., Gao, H., Baxter, W., Asturias, F., Boisset, N., Leith, A., Frank, J., 2008. Spider image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nature Protocols 3, 1941–1974.

Shang, Z., Sigworth, F., 2012. A hydration-layer model for cryo-EM image simulation. Journal of Structural Biology, http://dx.doi.org/10.1016/j.jsb.2012.04.021.

Shatsky, M., Hall, R., Brenner, S., Glaeser, R., 2009. A method for the alignment of heterogeneous macromolecules from electron microscopy. Journal of Structural Biology 166, 67–78.

Sigworth, F., 1998. A maximum-likelihood approach to single-particle image refinement. Journal of Structural Biology 122, 328–339.

Sindelar, C.V., Grigorieff, N., 2011. An adaptation of the wiener filter suitable for analyzing images of isolated single particles. Journal of Structural Biology 176, 60–74.

Starck, J., Murtagh, F., Fadili, J., 2010. Sparse image and signal processing: wavelets, curvelets, morphological diversity. Cambridge Univ Pr.

Tagare, H., Barthel, A., Sigworth, F., 2010. An adaptive expectation-maximization algorithm with GPU implementation for electron cryomicroscopy. Journal of Structural Biology 171, 256–265.

Tagare, H., Sigworth, F., Barthel, A., 2008. Fast, adaptive expectation-maximization alignment for cryo-em. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008, 855–862.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267–288.

Vonesch, C., Wang, L., Shkolnisky, Y., Singer, A., 2011. Fast wavelet-based single-particle reconstruction in cryo-em. In: International Symposium on Biomedical Imaging, pp. 1950–1953.

Voss, N., Lyumkis, D., Cheng, A., Lau, P., Mulder, A., Lander, G., Brignole, E., Fellmann, D., Irving, C., Jacovetty, E., et al., 2010. A toolbox for ab initio 3-D reconstructions in single-particle electron microscopy. Journal of Structural Biology 169, 389–398.

Wahba, G., 1980. Spline bases, regularization, and generalized cross-validation for solving approximation problems with large quantities of noisy data. Approximation Theory III 2.

Wainwright, M., 2009. Sharp thresholds for high-dimensional and noisy sparsity recovery using l1-constrained quadratic programming (lasso). IEEE Transactions on Information Theory 55, 2183–2202.

Wang, H., Wang, J., Ding, F., Callahan, K., Bratkowski, M., Butler, J., Nogales, E., Ke, A., 2007. Architecture of the yeast rrp44–exosome complex suggests routes of rna recruitment for 3′ end processing. Proceedings of the National Academy of Sciences 104, 16844.

Yin, Z., Zheng, Y., Doerschuk, P., 2001. An ab initio algorithm for low-resolution 3-D reconstructions from cryoelectron microscopy images. Journal of Structural Biology 133, 132–142.