TECHNISCHE UNIVERSITÄT MÜNCHEN
Chair for Computer-Aided Medical Procedures & Augmented Reality

3D Image Processing for Structural Analysis of Macromolecules Using Cryo-Electron Tomography

Yuxiang Chen

Vollständiger Abdruck der von der Fakultät für Informatik
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. Johann Schlichter
Prüfer der Dissertation:	1. Univ.-Prof. Dr. Nassir Navab
	2. Hon.-Prof. Dr. Wolfgang Baumeister

Die Dissertation wurde am 15.10.2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 24.02.2015 angenommen.

Abstract

Cryo-Electron Tomography (CET) is a three-dimensional (3D) imaging technique to study the structures of macromolecular complexes in their physiological environment. 3D image processing plays a pivotal role for the structural analysis of macromolecules depicted in cryo-tomograms. A typical processing workflow of CET comprises: (i) Reconstruct a tomogram from 2D projections of a sample. (ii) Identify macromolecules of interest in the reconstructed tomogram. (iii) Align subtomograms of the individual copies of macromolecules and average them to enhance the signal and increase the resolution of resulting average. (iv) Classify subtomograms depicting heterogeneous molecules into smaller, homogeneous classes.

In this dissertation novel 3D image processing methods for each of these processing steps are presented and assessed: (i) An iterative reconstruction algorithm based on nonuniform Fourier transform yields the most accurate reconstruction results. (ii) A supervised machine learning approach, combined with rotation-invariant features for 3D objects, enables accurate macromolecule identification. (iii) For subtomogram alignment, an algorithm is introduced, which allows efficient and accurate computation of a constrained correlation function via a generalized convolution theorem. (iv) A novel subtomogram classification algorithm is presented, which is able to automatically focus the similarity measurement to regions of highest structural variability. This autofocus ability does not require any prior knowledge or human intervention, which avoids hypothesis-driven bias of classification results.

In conclusion, a complete image processing workflow for molecular structural analysis in CET is covered here. All the introduced methods are thoroughly evaluated on various simulated and experimental datasets. Moreover, most of them allow rapid computation via parallelization and released as open source software to the community. The 3D image processing framework presented here provides a solid basis for users to process massive datasets rapidly and accurately.

Zusammenfassung

Kryoelektronentomographie (KET) ist ein Bildgebungsverfahren zur Erforschung der dreidimensionalen (3D) Strukturen makromolekularer Komplexe in ihrer physiologischen Umgebung. Bei der Strukturanalyse der abgebildeten Makromoleküle spielt 3D Bildverarbeitung eine Schlüsselrolle. Ein typischer Analysevorgang beinhaltet folgende Schritte: (i) die Rekonstruktion des 3D Tomogramms anhand von 2D Projektionen. (ii) Identifizierung spezifischer Makromoleküle im rekonstruierten Tomogramm. (iii) Alignierung von Subtomogrammen, die Kopien eines Makromoleküls abbilden, sowie deren Mittelung zur Verstärkung des Signals und zur Erhöhung der Auflösung. (iv) Klassifizierung von Subtomogrammen, die strukturell heterogene Moleküle abbilden, in kleinere, homogenenere Klassen.

In dieser Dissertation werden für jeden dieser vier Schritte neue 3D Bildverarbeitungsmethoden präsentiert und ausgewertet: (i) Ein auf nicht-uniformer Fouriertransformation basierender iterativer Rekonstruktionsalgorithmus erzielt höchst genaue Rekonstruktionen. (ii) Ein maschinelles Lernverfahren, das rotationsinvariante Kenngrößen nutzt, ermöglicht genaue Molekülidentifikation. (iii) Zur Subtomogrammalignierung wird ein Algorithmus vorgestellt, der die effiziente und genaue Berechnung der eingeschränkten Korrelationsfunktion unter Ausnutzung eines verallgemeinertem Konvolutionstheorems ermöglicht. (iv) Ein neuer Algorithmus zur Klassifikation von Subtomogrammen fokussiert sich automatisch auf die Regionen der höchsten Variabilität. Hierdurch sind biologische Vorinformation und subjektive Intervention nicht notwendig, wodurch hypothesegetriebene Beeinflussung der Klassifikation verhindert wird.

Zusammenfassend wird in dieser Arbeit ein umfassender rechnerischer Arbeitsablauf für molekulare Strukturanalyse mittels KET behandelt. Alle entwickelten Methoden sind gründlich evaluiert worden anhand verschiedener simulierter und experimenteller Datensätze. Die meisten Algorithmen erlauben schnelle Berechnungen durch Parallelisierung und die Quellcodes sind frei zugänglich. Die im Rahmen der Arbeit entwickelte 3D Prozessierungsplattform bietet Nutzern eine solide Basis, große Datensätze schnell und akkurat zu prozessieren.

Acknowledgements

I would like to express my sincere gratitude to all the people who assisted me during my PhD studies. In particular, I want to show appreciation to a few people in the following.

First of all, I want to thank my supervisor Dr. Friedrich Förster, who showed me the fascinating macromolecular world and guided me through the academic field. As a visionary and motivating supervisor, he set me an example of how to become a successful researcher. He trusted me and let me continue my own research ideas. During the course of my PhD studies, he offered me his guidance and insightful advices whenever I encountered problems. He also encouraged me to attend various international conferences to build up scientific connections with others and to broaden my vision.

I would also like to thank Prof. Dr. Nassir Navab, who gave me the chance to pursue my PhD at Technische Universität München. His group is top-tier in the field of medical image computing. During the meetings with him and other colleagues, I gained valuable feedback and suggestions, which was extremely helpful for my research work. He invited me every year to the chair's summer workshop, which was an exciting experience for me to interact with people from different fields.

I am very grateful to Prof. Dr. Wolfgang Baumeister for allowing me to conduct my work in the department of molecular structural biology of Max Planck Institute of Biochemistry at Martinsried. He created a world-renowned group with excellent working environment. It was always a pleasure for me to work and communicate with people from various backgrounds. I learned a lot from them.

Many thanks to my colleague Stefan Pfeffer, who acquired most of the experimental data used in this thesis and taught me knowledge about cryo-electron tomography; my former colleague Thomas Hrabe, who introduced me to this project and created a wonderful software platform for my work.

I thank Florian Beck and Luis Kuhn-Cuellar for their fruitful discussions; Diana Mateus and Olivier Pauly for their insights about machine learning and image processing.

Last but not least, I would like to dedicate this thesis to my parents, the most important people in my life. They are always supportive since my childhood. Especially when I decided to go abroad and continue my studies and life in a foreign country, they cleared the path for me and provided mental and financial support at their best effort. Without their concern and assistance, this thesis would be impossible. I love you all!

Contents

1

Introduction

As an extension of human vision into the smaller scale, various imaging techniques are the driving force behind numerous biological discoveries. They enable researchers to study questions at the organic, cellular, or even molecular level. As most of the basic, biological questions can be traced back to the cell, imaging the cell is pivotal in the life sciences. While the modern light microscope is able to examine the cell *in vivo*, its resolution is limited by the wavelength of the visible light (a few hundreds of nanometers). Although by using fluorescence labels it is able to accurately localize and dynamically track certain proteins in the living cell (fluorescence microscope and the confocal light microscope), it is still impossible for detailed studies of the structures of macromolecular complexes and their interactions in the cell.

To obtain 3D information of specific macromolecules at subnanometer resolution, high-resolution techniques can be used, such as X-ray crystallography, Nuclear Magnetic Resonance (NMR) and cryo-electron microscopy Single Particle Analysis (SPA). Take SPA for example, by imaging numerous copies of a structurally invariable macromolecule isolated from the cell, atomic resolution (a few angstroms) can be obtained [van Heel et al., 2000, Armache et al., 2010]. The resulting reconstruction can then be used to fit atomic models of the respective peptide or nucleotide chains for further interpretation. However, the complexes may dissociate during extensive purification steps or undergo conformational changes. Moreover, the spatial-temporal information of the complexes in the living cell is lost by the isolation. The study of the interactions between macromolecular complexes inside the cell can only be conducted in their physiological environment.

On the other hand, Cryo-Electron Tomography (CET) has the ability to visualize the cellular architecture and macromolecular assemblies three-dimensionally in their physiological settings [Lučić et al., 2005]. Due to the advances in the past few decades (sample preparation, automated image acquisition, microscope hardware and software), CET has become an important imaging tool for structural studies of macromolecules *in situ* at nanometer scale. Briefly, in CET the biological sample is first rapidly cooled in the liquid ethane (ca. $-180\,^{\circ}\mathrm{C}$). The plunge-freezing prevents

the formation of ice crystals and preserves the sample in a near-to-native condition. In contrast, other preparation methods that involve staining or chemical fixation lead to artefacts and limit the maximal attainable resolution [Lučić et al., 2005, Frank, 2006a]. Afterwards, the frozen-hydrated sample is placed in a Transmission Electron Microscope (TEM) and its 2D projections are acquired from different angles by tilting the sample holder. The automated data acquisition procedure is crucial here to accurately control the imaging process and to acquire the images under low-dose conditions [Koster et al., 1992, Dierksen et al., 1992, Dierksen et al., 1993]. Finally, the 2D projections are used to reconstruct the 3D map (or the so-called tomogram), which is subjected to further visualization, interpretation or processing. Due to the dose limit, the final attainable resolution of a raw tomogram is typically in the range of 5-10 nm [Grünewald et al., 2003].

Although the resolution of CET is inferior to the ones of X-ray crystallography, NMR and SPA, the uniqueness of CET is the capability to image not only the macromolecule of interest, but also its cellular context. It provides medium resolutions of macromolecular complexes without the need for extensive purification. It hence bridges the gap between low-resolution imaging techniques (e.g., X-ray computed tomography and light microscopy) and high-resolution imaging techniques (e.g., X-ray crystallography, NMR and SPA), linking cellular and molecular structural biology. By registering multiple copies of the macromolecule of interest and averaging them, the resolution can be further improved, making it a powerful tool to study, e.g., macromolecular complexes associated with the membranes (Figure 1.1) [Bartesaghi and Subramaniam, 2009, Pfeffer et al., 2012, Pfeffer et al., 2014], which are difficult for other imaging methods. In selected cases, even averages with subnanometer resolution have already been obtained by subtomogram averaging [Schur et al., 2013].

An ambitious goal of CET is towards the so-called "visual proteomics" [Nickell et al., 2006, Brandt et al., 2009, Förster et al., 2010], which aims at studying macromolecular complexes in the native cellular environment by detecting them, quantifying their abundances and analysing their spatial distributions and interactions. As a consequence, a macromolecular atlas can be build (Figure 1.2), which provides insights into the functional associations between the macromolecular complexes. Visual proteomics requires high-resolution tomograms of CET and accurate identification of various macromolecular complexes in the 3D maps. Currently, only large complexes, such as ribosomes, can be detected with reasonable fidelity. It remains a challenging task to identify complexes of smaller weights (< 1 MDa).

1.1 3D Image Processing in CET

Image processing techniques are of vital importance in CET to mine the structure information extensively. Starting with the projections (micrographs) obtained by

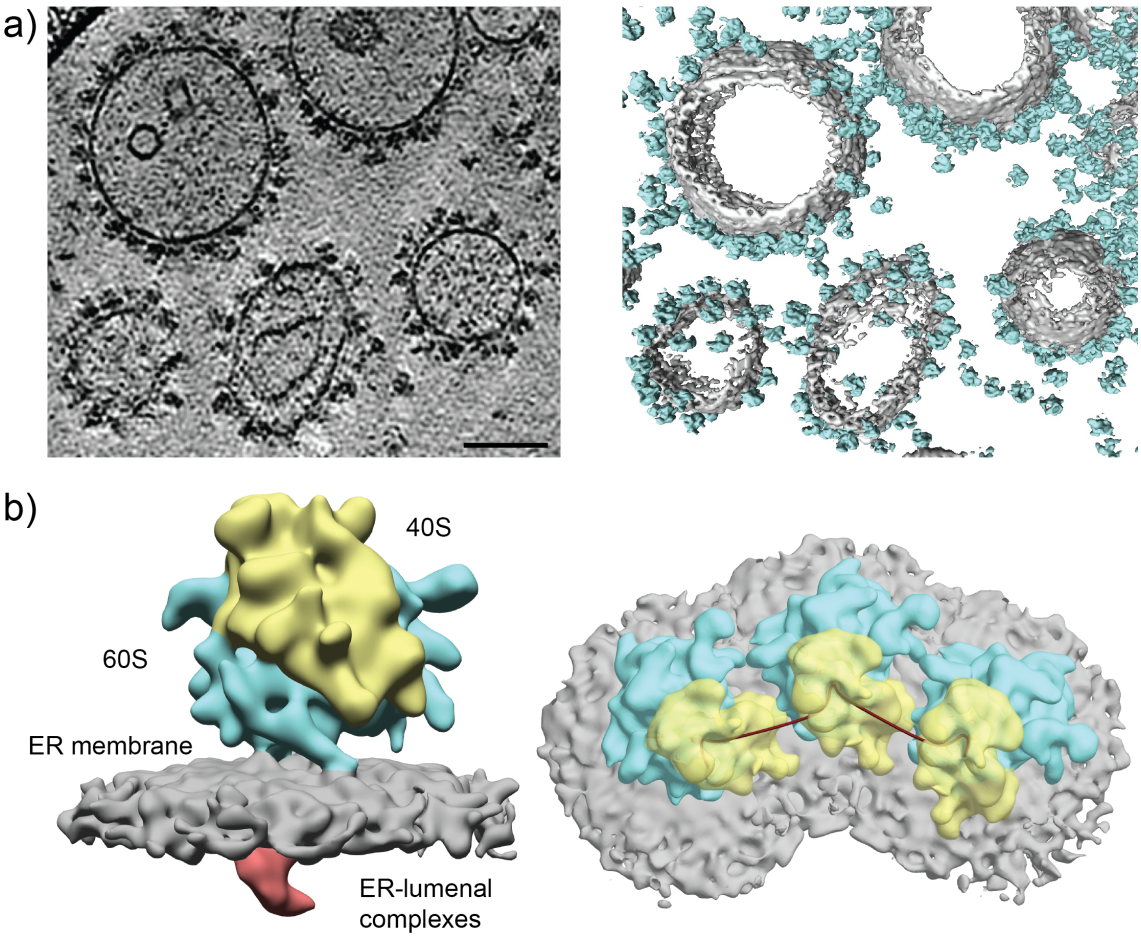

Figure 1.1: CET applied to study Endoplasmic Reticulum (ER) membrane-associated ribosomes. (a) Left: A 2D slice view of a tomogram of rough ER-derived microsomes. The image depicts densely populated, membrane-bound ribosomes. Scale bar: 100 nm. Right: The corresponding segmented map of the isosurface representation. The membrane is colored in gray and ribosome in cyan. (b) Left: The subtomogram average of the ER-associated ribosome at a resolution of approximately 3 nm. The 60S ribosomal subunit is colored in cyan, 40S subunit in yellow, membrane in gray and ER-lumenal part in red. Right: A preferred 3D arrangement of membrane-bound ribosomes *in situ*. A thread is drawn from the mRNA entry to the exit site of adjacent ribosomes, visualizing a possible pathway for interconnecting mRNA. Adapted from [Pfeffer et al., 2012].

a TEM, a typical processing workflow consists of several steps (Figure 1.3). First, the tomogram is reconstructed using the 2D projections from different projecting angles. CET shares the same principle with other tomography methods in medical imaging, which was first formulated by [Radon, 1917]. Radon proved that an object

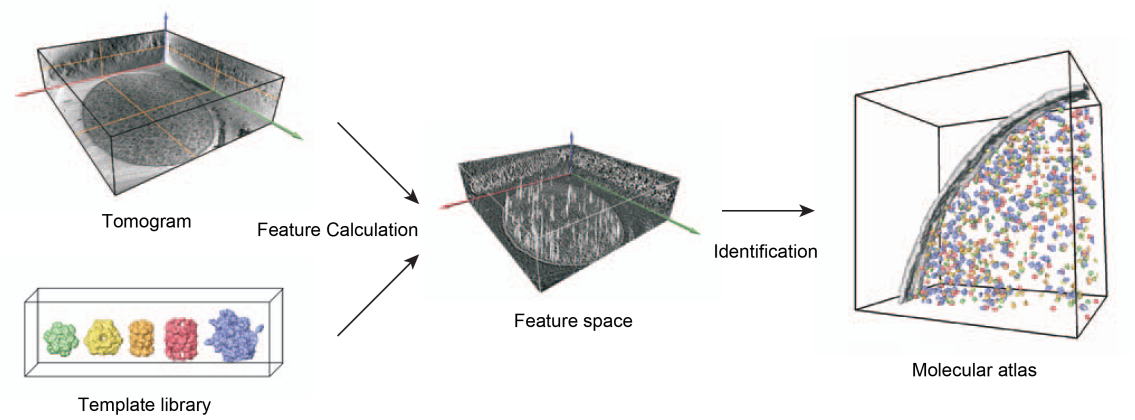

Figure 1.2: Idea of visual proteomics. The objective of visual proteomics is to detect macromolecular complexes comprehensively in a cell, allowing the studies of their abundances, spatial distributions and interactions. This can be achieved by first constructing a template library of macromolecules of interest, then calculating their features and finally identifying them inside the tomogram. As a result, a molecular atlas of the imaged cell can be built. Adapted from [Nickell et al., 2006].

can be reconstructed from its projections. This principle was first applied to CET in [DeRosier and Klug, 1968, Hart, 1968]. The resulting reconstruction is a 3D density map, indicating the density of the imaged specimen as a function of spatial coordinates.

Secondly, a specific type of macromolecular complex is often of interest, which has to be localized and identified in the 3D tomogram. Due to the low SNR of the tomogram and the large amount of data, manual labelling of the macromolecule by experts is rarely feasible. Instead, automated approaches utilize the structural signature of the macromolecule to identify and localize its occurrences within a tomogram. A common approach is to use the prior knowledge of the target macromolecule structure as the template to search the whole tomogram. With a descent similarity metric, the search can be efficiently computed, resulting in the spatial distribution information of the macromolecules in the tomogram.

Thirdly, subtomogram alignment and averaging can be conducted to get a 3D density of the target macromolecule with an improved resolution. The resolution of the raw tomogram is approximately 5-10 nm [Grünewald et al., 2003], which enables detecting large macromolecules. However, the resolution level is insufficient for depicting finer structural details. By aligning and averaging the subtomograms, extracted at the locations resulting from the identification step, the noise can be reduced, yielding a better-resolved structure. Normally, the alignment involves searching the best translational and rotational matches of 3D maps, which is a computationally intensive task. The obtainable resolution by subtomogram averaging

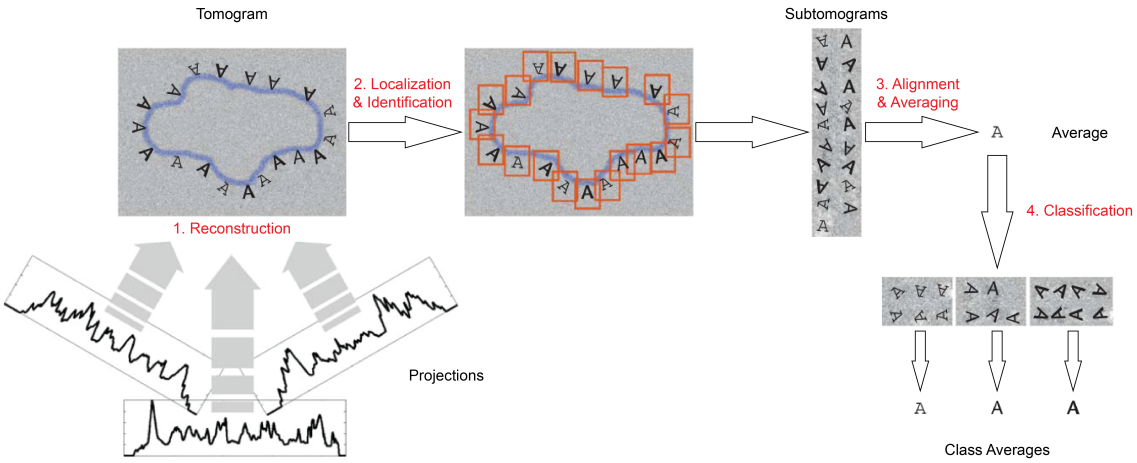

Figure 1.3: A typical image processing workflow of CET in 2D for simplification. It consists of four major steps: 1. The tomogram is reconstructed from the acquired projections of different angles. 2. The macromolecular complex of interest, denoted as 'A's with different fonts, is identified in the tomogram. 3. The subtomograms containing the targets are extracted from the tomogram. Aligning the subtomograms and averaging them can reduce the noise, resulting in a better resolved structure 'A'. 4. Further classification of the subtomograms helps to discriminate different conformations of the complex. Adapted from [Hrabe and Förster, 2011].

of thousands of subtomograms is typically in the range of 2-3 nm.

Finally, subtomogram alignment and averaging assume that the subtomograms contain copies of a specific macromolecule in the same conformation. If the imaged particles are structurally heterogeneous, the resolution of the averaged structure will be reduced. The heterogeneity could originate either from the false positive detections of the identification step, or from the conformational changes of the macromolecules when they fulfill their cellular functions, in which case the valuable information will be lost by simple averaging. This issue is addressed by subtomogram classification, where different, structurally more homogeneous classes are separated and their respective resolutions are improved ideally.

The aim of this thesis is to develop novel methods for these four steps: tomogram reconstruction, macromolecule localization and identification, subtomogram alignment and subtomogram classification to improve the image processing pipeline of CET. Although there are various software packages [Kremer et al., 1996, Frank et al., 1996, Sorzano et al., 2004, Nickell et al., 2005, Nicastro et al., 2006, Winkler, 2007, Heymann and Belnap, 2007, Castaño Díez et al., 2012] for each of these steps, the idea is to implement all the proposed methods in a unified open-source software, intended to provide a platform for accurate and efficient processing of CET data. All the methods should be applied to simulated and experimental datasets for

quantified evaluations of the performances.

1.2 Thesis Outline and Contributions

The remainder of the thesis is organized as follows: First, the background about the TEM and CET is presented in chapter 2. The knowledge about the components of a modern TEM helps understanding the principle of the image formation of a TEM, and therefore its power/weakness. A basic workflow of CET is presented to give an overview about the imaging technique.

In chapter 3, the first processing step, tomogram reconstruction, is discussed. Based on the projection-slice theorem, a Fourier-based 3D reconstruction method is adapted here. It formulates the reconstruction as an optimization problem and use the nonuniform fast Fourier transform as the forward (backward) projection operator in the optimization procedure, which can be solved iteratively using the conjugate gradient algorithm.

Chapter 4 demonstrates how to use machine learning to improve the identification of macromolecular complexes in the tomograms. A rotation-invariant descriptor for 3D data is proposed and integrated into a supervised learning framework. Compared to template matching, the approach yields a superior identification performance with a reduced false positive rate.

A fast and accurate alignment algorithm is presented in chapter 5 to align subtomograms to a common coordinate system. The major contribution is the generalization of the convolution theorem to the rotational space. Based on the spherical harmonic transform, the convolution can be efficiently computed. As evaluated on simulated and experimental datasets, the algorithm provides a speedup of up to three orders of magnitude and it is able to resolve the structures with 15-20 Å resolution, opening the possibility to process massive data in the future.

Chapter 6 introduces a subtomogram classification algorithm, which is able to automatically focus the classification on the regions of significant structural variability. This autofocus ability does not require any prior knowledge about the macromolecular structure. The algorithm can deconvolute different conformational states of macromolecular complexes *in situ* as demonstrated for ER-associated ribosomes.

All the above-mentioned algorithms and tools were implemented in PyTom [Hrabe et al., 2012] and released as an open-source software to the community.

Finally, the thesis is summarized in chapter 7. Detailed discussions are provided for each topic and the outlook for the future work is suggested.

2

Background

2.1 Transmission Electron Microscopy

The major components of a modern transmission electron microscope (TEM) are first described, followed by a brief discussion of the principles of the image formation mechanism in a TEM. They provide a basis for understanding the information content of electron micrographs, as well as the challenges faced when processing them.

2.1.1 Transmission Electron Microscope

As shown in Figure 2.1, the basic setup of a TEM resembles a light microscope, which has an illumination source, condenser lenses to focus the incident beam on the specimen, a specimen stage, objective lenses to obtain a real-space image from the scattered light, projector lenses to magnify the resulting image, and finally a camera to record the image. In detail, some parts are explained in the following.

Electron gun. The major difference between a TEM and a light microscope is the illumination source: electrons are used in the TEM instead of light. Electrons are emitted by the electron gun, accelerated and formed into an electron beam with high spatial and temporal coherence that transmits the specimen afterwards. A high-performance TEM is typically equipped with a field emission gun (FEG), which produces electron beams with significantly higher brightness than the cheaper LaB_6 sources. The resulting wavelength of electrons ranges from 0.037 Å to 0.020 Å (with the accelerating voltage between 100 kV and 300 kV), allowing structural studies at high resolution [Frank, 2006b]. To avoid interaction between the electron beam and air, the TEM column has to be kept at ultra-high vacuum.

Condenser lenses. Condenser lenses (C1 and C2) and their respective apertures condense the electron beam to a small spot size (typically on the order of 1 µm for bright-field imaging) [Reimer and Kohl, 2008]. For this purpose, electromagnetic coils are used as the lenses, which generate magnetic fields altering the direction of

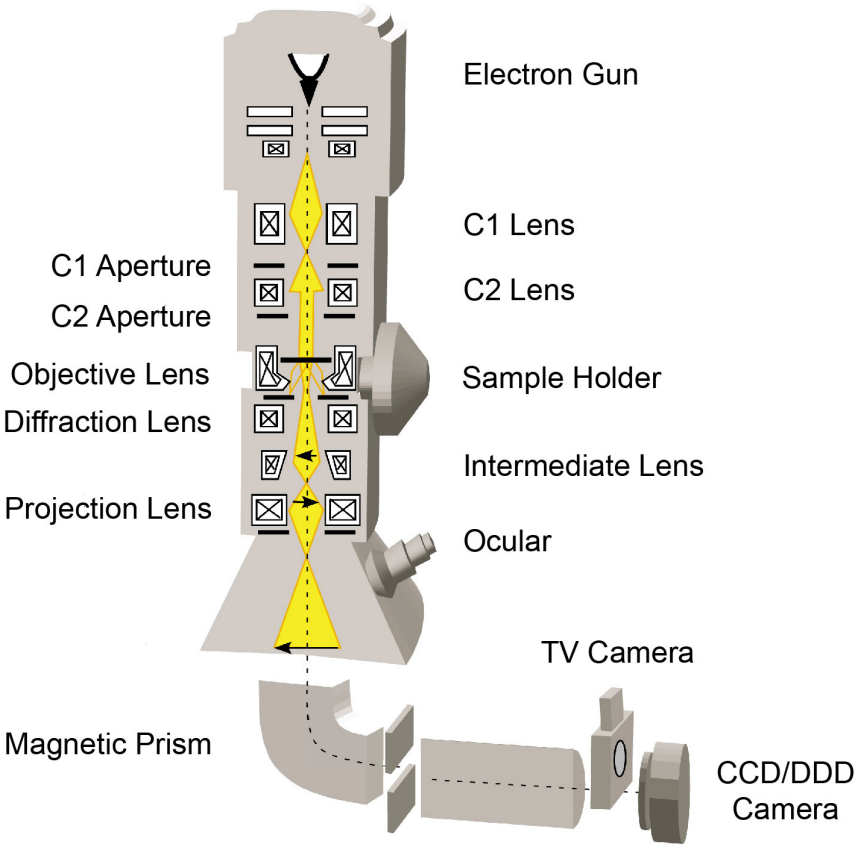

Figure 2.1: Schematic view of a modern transmission electron microscope. Adapted from [Schweikert, 2004].

electron beam (but not its energy). By applying different currents on the coils, the focal length of the lenses can be adjusted.

Specimen stage. The specimen is usually applied to a metal mesh grid, which is approximately 3 mm in diameter. The grid is then placed on a computer-controlled stage allowing precise translation and single-axis rotation of the sample.

Objective lenses, intermediate lenses and projector lens. The image of the specimen is first created by the objective lenses. At the back focal plane, the objective aperture removes the electrons with high scattering angles and improves the contrast. The image is further magnified by the intermediate lenses and finally projected to the detector by the projector lens.

Energy filter. When thick samples ($>$ 200 nm) are imaged, the chance of inelastic scattering of electrons increases (see section 2.1.2), which leads to blurring of the image due to the high chromatic aberration of the objective lenses. The blurring can be reduced by employing an energy filter, which can filter out the inelastically scattered electrons. One commonly used type is the post-column energy

filter [Krivanek et al., 1995], which consists of a magnetic prism dispersing the electrons according to their energy and an energy selective slit allowing only electrons within a certain energy range to pass.

Image Detector. The electrons are finally captured by the detector, forming an image visible to the human eye. A widely used detector is the scintillator-coupled Charge Coupled Device (CCD) camera, in which the electrons are first converted to photons by the scintillator. These photons are then transmitted to the CCD array, using either fiber optical or lens coupling, to generate the digital signal. During this process, the signal is first carried by electrons, then photons, and finally electrons again. The transformation of the carrier causes unnecessary information loss. This is addressed by the recently introduced Direct Detection Devices (DDDs), which are digital cameras capable of detecting the electrons directly. They have superior quantum efficiency and increased image contrast [McMullan et al., 2009, Faruqi and McMullan, 2011], which leads to a significant performance advance over CCD cameras.

2.1.2 Image Formation

Before describing the image formation mechanism of the TEM, we have to understand the interaction of the electrons with the specimen. There are four major types of interactions (Figure 2.2): 1. Most of the electrons are too far away from any atom and their paths will not be altered. They are called *unscattered* electrons. 2. If the electrons pass in the range of the electron clouds of the atoms, they will be scattered due to the Coulomb force. If a negligible amount of energy is transferred from the electrons to the specimen, we define this interaction as the *elastic scattering*, which is the main factor of forming the high resolution TEM image because this interaction contains information about the Coulomb potential distribution of the atoms in the specimen. 3. On the other hand, there is a significant energy loss of the electrons by *inelastic scattering*, which causes energy deposition in the specimen and thus radiation damage. Moreover, inelastic scattered electrons will introduce incoherence by generating new wavelengths, which results in the background noise. This becomes a severe issue, when the specimen is thick, increasing the chance of multiple scattering events (therefore also the chance of inelastic scattering). In such case, an energy filter is required to filter out the inelastic scattered electrons. 4. There are some electrons that travel close to the atomic nuclei. They will be attracted by the high Coulomb potential and be scattered at high angles. We call them *backscattered* electrons. They will be removed by the objective aperture, resulting in a decrease of transmitted electron intensity.

The image formation mechanism of a TEM comprises two parts: phase contrast and amplitude contrast. Phase contrast is due to the elastic scattering and is the main contribution of contrast in the image [Dubochet et al., 1988]. It arises from the interferences of electron waves at the image plane. Amplitude contrast, on the other

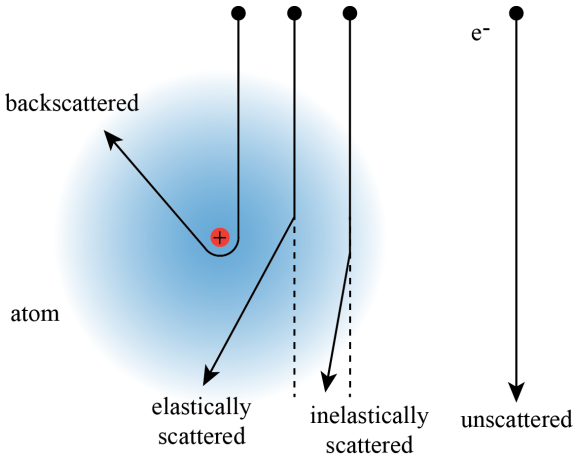

Figure 2.2: Interactions of the electron beam with an atom in the specimen.

hand, is a result of the backscattering and inelastic scattering, in which there is a loss of electron intensity. It contributes mainly to the low or medium resolutions. In the following, the phase contrast and amplitude contrast are discussed in detail.

Phase contrast. We can describe the incident electrons travelling through the specimen along the z-direction as a uniform planar wave $\Psi_0 = \exp(ikz)$, where $k = 1/\lambda$ is the wavenumber of the electrons. Due to the elastic scattering, the electron wave will undergo a phase shift $\Phi(x, y)$ [Frank, 2006b]:

$$\Psi(x, y) = \Psi_0 \exp(i\Phi(x, y)), \tag{2.1}$$

$$\Phi(x, y) = \int C(x, y, z)\mathrm{d}z. \tag{2.2}$$

Here, (x, y) is the coordinate of the specimen plane and $C(x, y, z)$ is the Coulomb potential distribution of the specimen. Under weak-phase approximation: $\Phi(x, y) \ll 1$, $\Psi(x, y)$ can be expanded to:

$$\Psi(x, y) = \Psi_0[1 + i\Phi(x, y) - \frac{1}{2}\Phi(x, y)^2 + \cdots]. \tag{2.3}$$

Omitting the Taylor series expansion of above the first order suggests that the electron wave behind the specimen can be approximated as a sum of an unscattered wave Ψ_0 and a weakly scattered wave of low amplitude $\Phi(x, y)$ with a $\frac{\pi}{2}$ phase shift.

However, this phase shift is hardly measurable because the intensity is dominated by the unscattered wave (Figure 2.3). This is essentially the same problem facing in light microscopy and it led to the invention of the phase contrast microscope. Here, on the other hand, the lens aberrations and defocusing are utilized to generate additional phase shift in TEM. Mathematically, the phase delay W depends on the frequency $\boldsymbol{k} = (k_x, k_y), k = |\boldsymbol{k}|$ and it can be expressed as [Reimer and Kohl, 2008]:

$$W(\boldsymbol{k}) = \frac{\pi}{2}(C_s\lambda^3k^4 - 2\Delta z\lambda k^2). \tag{2.4}$$

Herein, C_s is the spherical aberration coefficient of the lens; λ is the electron wavelength; and Δz is the defocus of the objective lens.

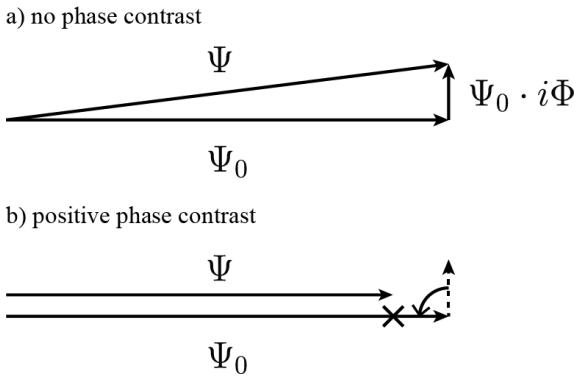

Figure 2.3: Phase contrast generation. (a) No phase contrast. The amplitude of the resulting wave Ψ is approximately the same as the one of the incident wave Ψ_0, making the phase shift hardly measurable. (b) Positive phase contrast. An additional $\pi/2$ phase shift is used to produce more contrast. Adapted from [Reimer and Kohl, 2008].

Therefore, the wave function at the back focal plane of the objective lens can be expressed in Fourier space as [Frank, 2006b]:

$$\Psi_{bf}(\boldsymbol{k}) = \mathcal{F}(\Psi(x,y)) \exp(iW(\boldsymbol{k})). \tag{2.5}$$

The wave function in the image plane can be obtained by an inverse Fourier transform:

$$\Psi_i(x,y) = \mathcal{F}^{-1}(\mathcal{F}(\Psi(x,y))A(\boldsymbol{k}) \exp(iW(\boldsymbol{k}))). \tag{2.6}$$

Here, $A(\boldsymbol{k})$ is the aperture transfer function:

$$A(\boldsymbol{k}) = \begin{cases} 1 & \text{for } |\boldsymbol{k}| = \theta/\lambda \leq \theta_1/\lambda \\ 0 & \text{elsewhere} \end{cases}, \tag{2.7}$$

where θ_1 is the angle corresponding to the radius of the objective lens aperture. Finally, the observed intensity in the image plane is equal to the magnitude of the incident wave function (ignoring the scaling factor):

$$I(x,y) = |\Psi_i(x,y)|^2 . \tag{2.8}$$

Assuming $\Phi(x,y)$ is real and ignore the terms above the first order, Equation 2.8 can be rewritten as [Reimer and Kohl, 2008]:

$$\mathcal{F}(I(x,y)) = O(\boldsymbol{k})A(\boldsymbol{k}) \sin(W(\boldsymbol{k})), \tag{2.9}$$

$$O(\boldsymbol{k}) = \mathcal{F}(\Phi(x,y)). \tag{2.10}$$

Here, the function $\sin(W(\boldsymbol{k}))$ is the so-called contrast transfer function (CTF). It is a function of the spatial frequency $\boldsymbol{k}$ and characterizes how the information of different frequencies is transferred (Figure 2.4). Due to the oscillatory nature of the sine function, some frequencies are transferred strongly while others are inverted or even eliminated. As seen from Equation 2.4, CTF is mainly controlled by the defocus value Δz. The effect of the CTF is illustrated in Figure 2.5.

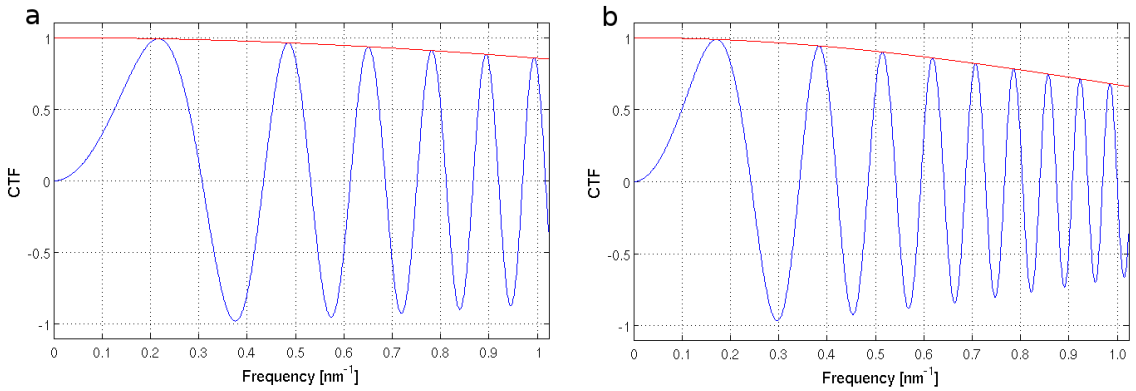

Figure 2.4: Theoretical contrast transfer functions (blue) and their envelopes (red). (a) The CTF is calculated with the defocus value of 5 µm, the accelerating voltage of 300 kV and the spherical aberration of 2 mm. The first zero-crossing of the CTF is at (3.1 nm)$^{-1}$. (b) If the defocus value is changed to 8 µm, the first zero-crossing of the CTF is at (4 nm)$^{-1}$.

Equation 2.9 assumes the coherent illumination with monochromatic electrons. However, the illumination in practice has both a finite divergence and a finite energy spread, which results in the damping of CTF towards higher frequencies. This can be formulated by introducing an additional "envelope function" $E(\boldsymbol{k})$:

$$\mathcal{F}(I(x,y)) = O(\boldsymbol{k})E(\boldsymbol{k})A(\boldsymbol{k})\sin(W(\boldsymbol{k})). \tag{2.11}$$

The descending property of $E(\boldsymbol{k})$ limits the obtainable resolution of TEM.

Amplitude contrast. Equation 2.1 does not take into account of the amplitude component. The Fourier transform of the amplitude component is transferred by $\cos(W(k))$ and Equation 2.9 can be rewritten as [Frank, 2006b]:

$$\mathcal{F}(I(x,y)) = O_r(\boldsymbol{k})A(\boldsymbol{k})\sin(W(\boldsymbol{k})) - O_i(\boldsymbol{k})A(\boldsymbol{k})\cos(W(\boldsymbol{k})), \tag{2.12}$$

where $O_r(\boldsymbol{k})$ and $O_i(\boldsymbol{k})$ are the real and imaginary parts of $O(\boldsymbol{k})$, respectively [Erickson and Klug, 1970]. The percentage of amplitude contrast varies according to the atom species. For a negatively stained specimen this percentage is higher, while for a thin cryo-specimen it is usually between 5%-7% [Orlova and Saibil, 2011].

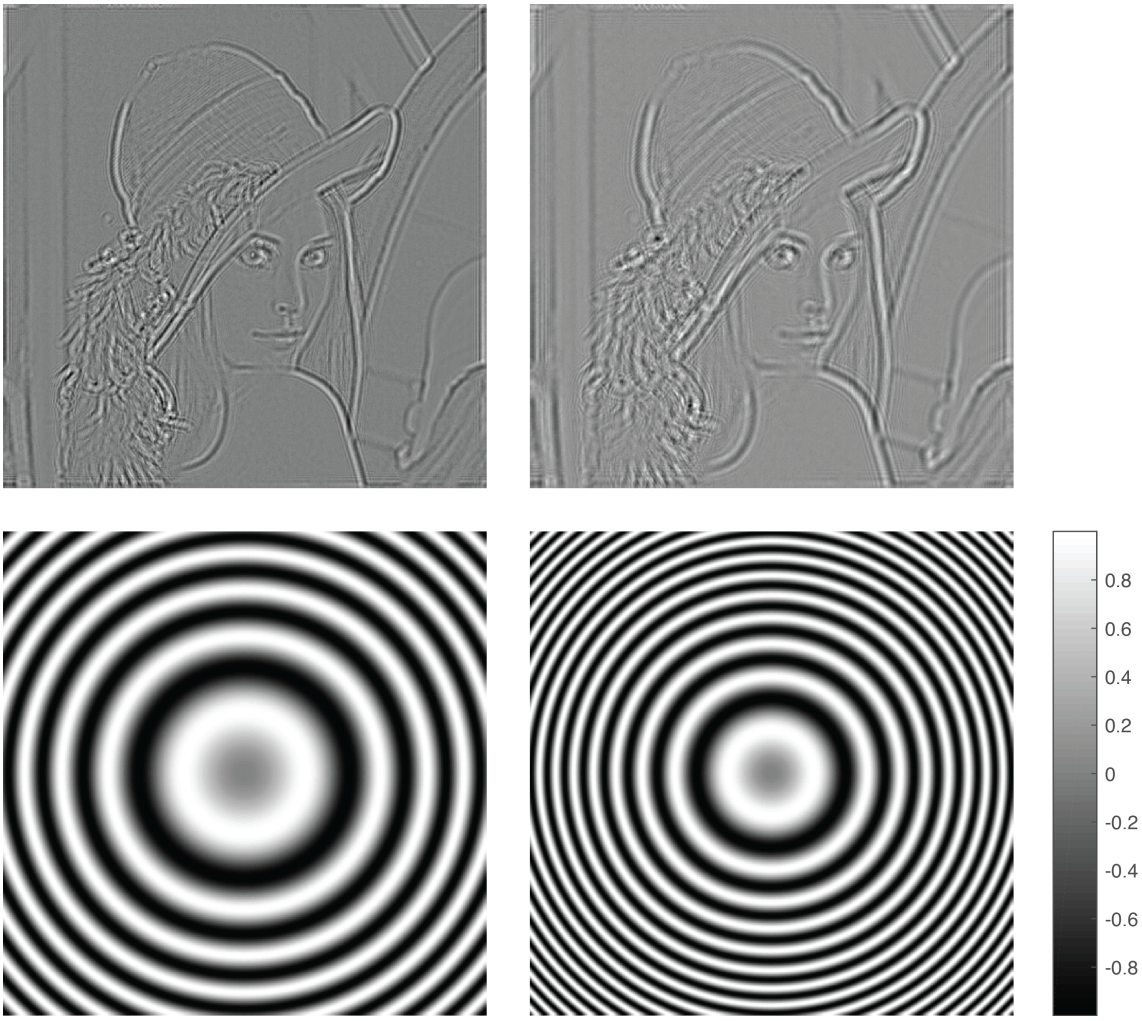

Figure 2.5: The effect of the CTF illustrated in 2D. The original image can be seen in Figure 2.6. Top row: CTF-affected images. Bottom row: Corresponding 2D CTFs.

2.2 Cryo-Electron Tomography

After the description of TEM, some important steps of CET are presented here, providing an overview about this imaging technique. The state-of-the-art methods for all the steps are introduced and some challenges are discussed.

2.2.1 Sample Preparation

Before the biological sample is examined under the microscope, it has to be fixated in a solid state. In molecular structural biology, the most common method is plunge freezing ([Dubochet et al., 1988]): the aqueous sample is first placed on a thin EM grid and then rapidly submerged in liquid ethane (ca. $-180\,^{\circ}\mathrm{C}$). During plunge freezing the temperature drops so quickly (10^4 K/s) below $-140\,^{\circ}\mathrm{C}$ that the formation of ice crystals is avoided, which would cause damage to the biological material. Instead, the ice is in an amorphous state (vitrification), which preserves the sample in a near-physiological condition.

2.2.2 Image Acquisition

After the sample is prepared, it is further examined using a TEM. A series of 2D projections is recorded from different angles by tilting the sample holder. Due to the dramatic increase of the sample thickness at high tilt angles, inelastic scattering of the electrons increases, which greatly degrades the image quality. The consequence is that the tilt angles are typically restricted to a certain range, e.g, -60° to $+60^{\circ}$ (limited angle tomography) with an angular increment of 2°-3°, which causes severe artefacts of the reconstruction (discussed in section 2.2.5). Automated data collection process is vital for the tilt-series acquisition [Koster et al., 1992, Dierksen et al., 1992, Dierksen et al., 1993], to automatically compensate for lateral movements of the specimen, to maintain invariant imaging conditions (e.g., keep the target centered under the beam and at a constant focus) and to take images with minimal electron dose (low dose) throughout the entire acquisition process. The development of such automated data acquisition tools (e.g., SerialEM [Mastronarde, 2005], TOM [Nickell et al., 2005], UCSF Tomography [Zheng et al., 2007], Leginon [Suloway et al., 2009]) makes it possible to obtain large amounts of data for structural studies with high resolution.

Another important issue during the data acquisition is the dose limitation. There is a limit of electron dose that can be applied on the biological sample because too much exposure leads to structural changes. The electron dose should be sufficiently low to avoid the structural damage of the specimen. Therefore, the obtainable resolution of CET is essentially limited by the applicable electron dose, typically below 100 $e^-/\text{Å}^2$ allowing resolutions in the range of 1-2 nm [Henderson, 2004]. Moreover, the electron dose has to be divided over the number of projections. As

a consequence, the SNR of each projection is extremely low. This is the major difference between SPA and CET. In SPA all the electron dose is applied on a single image. According to the dose-fractionation theorem [McEwen et al., 1995], the information contained in a projection and a tomogram is the same, if both are acquired with the same electron dose [Hegerl and Hoppe, 1976, McEwen et al., 1995]. Therefore, SPA resolves a 2D view in higher resolution, while CET yields a 3D map with lower resolution.

2.2.3 Tilt-Series Alignment

Although sophisticated automated acquisition tools exist for correcting large distortions, further processing of the recorded images/tilt-series is required for more accurate corrections of mechanical imperfections and shifts, rotations, magnification changes of the images [Frank, 2006a]. The goal of this step is to transform/align the images to a common coordinate system. This is typically achieved by tracking and aligning features with high contrast throughout the tilt-series.

A conventional method is to add colloidal gold particles into the biological samples as the fiducial markers, which is electron-dense and will create high contrast circle-like features. After the tilt-series is recorded, the fiducial markers are identified and localized either manually, semi-automatically or automatically across the whole tilt-series for further refinement. Because manual labelling of the gold markers is tedious, several algorithms were proposed to automatically track the markers [Brandt et al., 2001b, Amat et al., 2008]. There are also approaches which do not require the presence of gold particles; they are normally based on tracking of high-contrast features or patches [Brandt et al., 2001a, Brandt and Ziese, 2006, Castaño Díez et al., 2007, Sorzano et al., 2009, Castaño Díez et al., 2010]. They are particularly useful when adding the gold particles is impossible or inconvenient.

After the markers are localized, their coordinates are used to align the images by typically a least squares algorithm to minimize the alignment error as a function of shifts, rotations, magnification changes of the images, etc [Frank, 2006a, Lawrence et al., 2006, Amat et al., 2010a]. The resulting images are used for further processing.

2.2.4 CTF Correction

As discussed in section 2.1.2, the contrast of the images recorded by TEM is dominated by the phase contrast. In linear approximation, the micrograph is a projection of the specimen's electrostatic potential, convoluted with the inverse Fourier transform of the CTF, which describes the imaging properties of the TEM (lense aberration, defocus, etc). The CTF oscillates around zero, which not only modulates the amplitude of the signal in Fourier space, but also reverses its phase at some frequencies (Figure 2.4). As a consequence, the details of the acquired image at certain frequencies will have flipped signs (Figure 2.5). This is not a problem if the

expected resolution is below the first zero of the CTF. For example, if the defocus value is 4 µm and the accelerating voltage is 300 kV, the first zero is about 2.8 nm.[1] Lowpass filtering until the first zero is enough for structural studies at resolution 2.8 nm. However, if higher resolution is anticipated, e.g., by subtomogram averaging, CTF correction must be done.

Before correcting for the CTF it must be determined accurately. This is not trivial in CET because the SNR of the micrograph is low, which makes the power spectrum too weak to determine the CTF. A common strategy is the periodogram averaging [Fernández et al., 1997]: based on the assumption that the defocus along the tilt axis is constant, the regions (tiles) around the tilt axis in the micrograph will have the same defocus value. By averaging the power spectra of the tiles the defocus is determined by comparing the periodograms with the theoretical CTF model. From this defocus value the CTF at any point of the micrograph can be calculated according to the geometry. After the CTF is determined the correction can be done by phase flipping [Zanetti et al., 2009] or Wiener filtering [Fernández et al., 2006]. It is worth mentioning that with the development of phase plate [Danev et al., 2010] it will allow tilt-series acquisition close to focus without loss of contrast at low resolution.

2.2.5 Tomogram Reconstruction

After the projections are aligned and possibly CTF corrected, they can be used to reconstruct the 3D density map (tomogram). The principle connecting the projections and the tomogram is the so-called projection-slice theorem, which states that, for a 3D object, the Fourier transform of its 2D projection corresponds to a central slice of the 3D Fourier transform of the object [Bracewell, 1986]. This implies that, if the projection angles are fully covered from $-90°$ to $+90°$, the object can be uniquely recovered. However, as discussed in section 2.2.2, CET typically has a limited angular tilt range (e.g., from $-60°$ to $+60°$). As a consequence, a wedge-shaped region in the Fourier space is unsampled. This is called "missing wedge" problem (Figure 2.6b), which makes the reconstruction ill-posed (no unique solution exists) [Davison, 1983, Natterer, 2001]. This problem typically leads to severe artefacts of the reconstruction, especially along the direction of the "missing wedge". A remedy of this problem is to employ the double-tilt axis acquisition geometry [Penczek et al., 1995, Mastronarde, 1997] (Figure. 2.6a), which can, in principle, reduce the missing information dramatically. However, the double-tilt axis acquisition scheme is not widely used in CET due to the additional (mechanical and algorithmic) complexities introduced by the second tilt axis.

[1] Theoretically speaking, imaging close to focus can push the first zero to higher resolution. However, images under such situations will have low contrast, which makes it difficult to depict structural features.

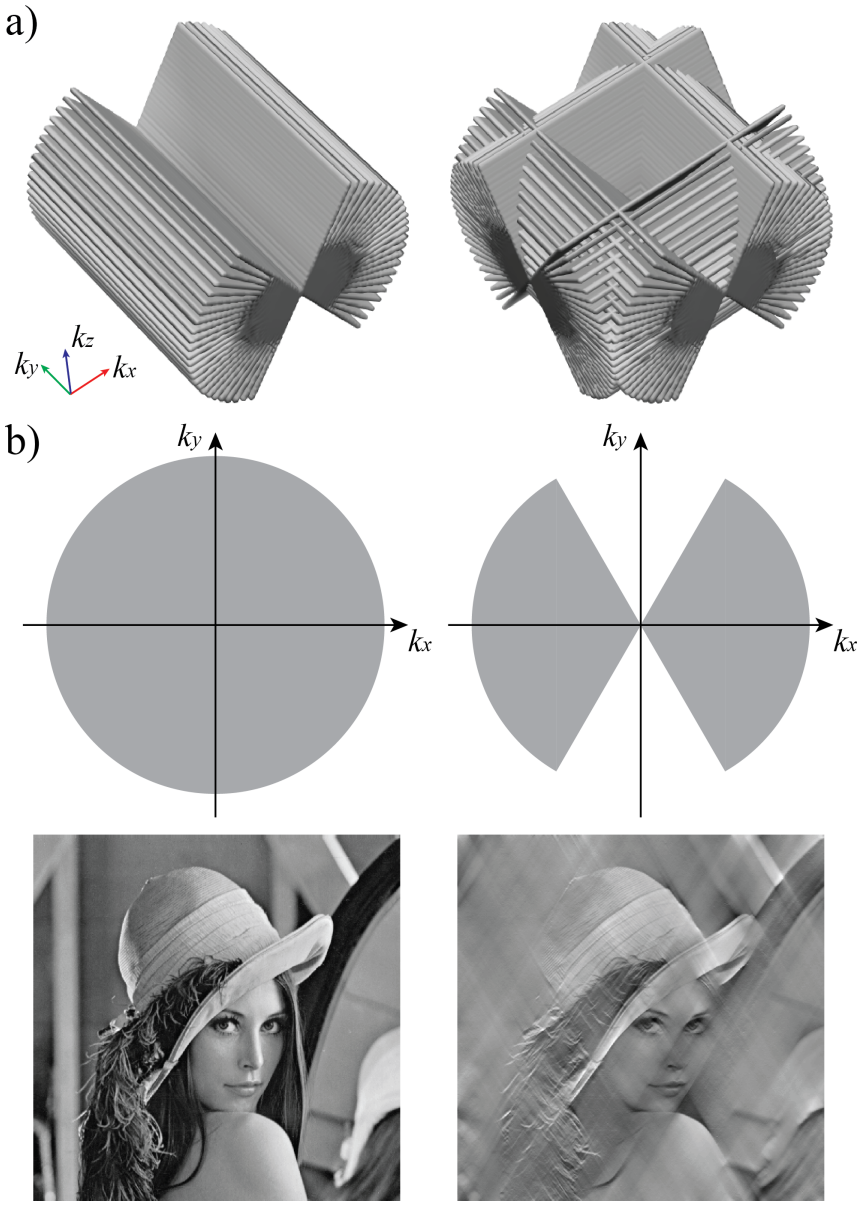

Figure 2.6: "Missing wedge" problem in CET. (a) Left: the sampling of the single tilt-axis tomography in Fourier space. The tilt axis is the y-axis and the tilt angle is limited from $-60°$ to $+60°$. A wedge-shaped area in Fourier space is left unsampled. Right: the "missing wedge" can be reduce to the "missing pyramid" using a double-tilt axis acquisition scheme. (b) The "missing wedge" effect illustrated in 2D. The upper and lower rows show the sampling area (gray) in Fourier space and the corresponding images in real space, respectively. If the "missing wedge" is absent, the Fourier space is fully sampled and the corresponding image is isotropically resolved. On the other hand, if the Fourier space is only partially sampled, the resulting image is severely deformed.

Arguably, the most commonly used method for tomogram reconstruction is the weighted backprojection (WBP) algorithm, which has been introduced several decades ago [Ramachandran, 1971, Harauz and van Heel, 1986, Radermacher et al., 1986, Radermacher, 1992]. In WBP the 2D projections are projected back to the 3D space according to the acquisition geometry. The projections have to be weighted properly prior to the backprojection to handle the unevenly distributed sampling points in Fourier space. Otherwise, the low frequency would be artificially enhanced. Due to its simplicity and efficiency, WBP is well understood and accepted in CET.

Other real space approaches also exist, such as Algebraic Reconstruction Technique (ART) [Gordon et al., 1970, Marabini et al., 1998], Simultaneous Iterative Reconstruction Technique (SIRT) [Gilbert, 1972, Penczek et al., 1992] and Simultaneous Algebraic Reconstruction Technique (SART) [Wan et al., 2011]. They formulate the reconstruction problem as a system of linear equations, which can be solved by minimizing the error between the observed projections and the expected projections calculated from the reconstruction. They are becoming popular because: (i) They normally produce results with more contrast than WBP. (ii) Regularization can be readily applied by incorporating prior knowledge, which is commonly done in medical imaging field. The regularization helps to stabilize the inversion procedure for ill-posed problem. Common strategies include total variation regularization [Wang et al., 2008], edge-preserving regularization [Yu and Fessler, 2002] and wavelet/curvelet regularization [Verhaeghe et al., 2008, Frikel, 2013], etc. However, such assumption or prior knowledge is hardly available for CET. Making a wrong assumption could lead to biased structure information. Advanced regularization is therefore rarely used for reconstruction in CET.

Another major category of reconstruction algorithms is based on Fourier transformation. The representatives are the fast Fourier summation algorithm [Sandberg et al., 2003], the gridding method [Penczek et al., 2004] and the nearest neighbor (NN) direct inversion method [Grigorieff, 1998, Zhang et al., 2008]. These Fourier-based methods have been shown to result in more accurate reconstructions than the algebraic methods (without imposing constraints) in terms of Fourier Shell Correlation (FSC, see section 2.2.9) [Penczek et al., 2004].

2.2.6 Tomogram Interpretation

The interpretation of the tomogram typically requires to decompose the tomogram into individual structural components. One category is based on segmentation. Although it is hampered by the low SNR of the tomograms and the "missing wedge" problem, it has been successfully applied to segment large structures with high contrast, such as membranes [Moussavi et al., 2010, Martinez-Sanchez et al., 2011, Martinez-Sanchez et al., 2013], microtubules [Weber et al., 2012] and filaments [Rigort et al., 2012]. Various automatic or semi-automatic segmentation methods were proposed [Volkmann, 2010], which include Watershed transform [Volkmann,

2002], density thresholding [Cyrklaff et al., 2005], normalized graph cut and eigenvector analysis [Frangakis and Hegerl, 2002], active contours [Bartesaghi et al., 2005], oriented filters [Sandberg and Brega, 2007], level-set [Whitaker and Elangovan, 2002], etc. Nevertheless, few of them are universally applicable and widely used due to the high noise level of cryo-tomograms and obscure parameters to tune. Manual segmentation is still commonly adopted.

Another category is based on identification, when multiple copies of specific macromolecules are present. The target is to localize and identify them in the tomograms. Most of the identification methods in CET originate from SPA in cryo-electron microscopy (CEM), in which 2D images of the biological samples containing presumably identical copies of macromolecules are acquired [Frank, 2002]. Compared to tomograms from CET, the micrographs from CEM are 2D and have relatively higher SNR, which makes the identification easier. Popular approaches include template matching methods [Ludtke et al., 1999, Roseman, 2003, Wong et al., 2004, Huang and Penczek, 2004], feature-based methods [Mallick et al., 2004, Zhu et al., 2003, Hall and Patwardhan, 2004, Volkmann, 2004] and some machine learning methods [Ogura and Sato, 2004, Mallick et al., 2004]. However, most of these methods are not widely adopted in CET due to several reasons: (i) The SNR of CET is lower than that of CEM. And the tomogram of CET is "missing wedge"-affected, which results in distortions in real space. Therefore, many feature-based approaches cannot be applied. (ii) The tomogram of CET is 3D, which makes the extension from some 2D methods difficult and the computation more intensive.

Among all the listed methods, the most widely used one is template matching [Frangakis et al., 2002, Roseman, 2003] (or matched filtering), when prior knowledge (template) about the structure of the interested macromolecule exists. This is achieved by exhaustively comparing the structural template against the noise-corrupted signal (tomogram) under scrutiny in different orientations. The resulting map contains similarity measures of the template and corresponding subregions of the tomogram. The maxima of the map indicate possible occurrences of the target macromolecule (candidates). More specifically, the local correlation function is typically used in CET as a measure of the similarity [Roseman, 2003]. Details of template matching technique are discussed in section 4.2. The template matching approach is widespread due to three major reasons: (i) It is is able to detect large macromolecules (> 1 MDa) with reasonable fidelity and is noise robust compared to many other approaches [Zhu et al., 2004]. (ii) The handling of the "missing wedge" problem can be integrated into the correlation score [Frangakis et al., 2002]. (iii) The computation is efficient using Fast Fourier Transforms [Roseman, 2003].

2.2.7 Subtomogram Alignment and Averaging

Higher resolution of a specific type of macromolecular complex can be obtained by aligning the subtomograms containing multiple copies of the same complex and

averaging them. This can reduce the noise level and fill in the missing information caused by the limited angle tomography.

Mathematically, given n subtomograms $V_1, \ldots, V_n$ depicting the same object S (no shift and rotation) and assuming each subtomogram V_i is corrupted by additive white Gaussian noise (i.i.d.): $V_i = S + N_i$, averaging all the subtomograms yields:

$$A = n \cdot S + \sum_n N_i, i \in [1, \ldots, n]. \tag{2.13}$$

Then the SNR of the average is:

$$SNR_A = \frac{\mathrm{Var}(n \cdot S)}{\mathrm{Var}(\sum_n N_i)} = \frac{n^2 \cdot \mathrm{Var}(S)}{n \cdot \mathrm{Var}(N)} = n \cdot SNR_{V_i}. \tag{2.14}$$

That means, theoretically speaking, the SNR of the average increases linearly with the number of subtomograms.

Prior to averaging all the subtomograms have to be aligned to a common coordinate system. This is an optimization problem where the objective is to maximize the similarity scores between the average A and all the subtomograms $V_1, \ldots, V_n$ with unknown shifts $\boldsymbol{\tau}$ and rotations $\boldsymbol{R}$:

$$A = \underset{A,\boldsymbol{\tau},\boldsymbol{R}}{\arg\max} \sum_{i=1}^{n} Score(T_{\boldsymbol{\tau}_i} \Lambda_{\boldsymbol{R}_i} A, V_i), \tag{2.15}$$

where T is the translation operator and Λ is the rotation operator. For the similarity metric, most software packages, including AV3 [Förster and Hegerl, 2007], PEET [Heumann et al., 2011], BSOFT [Heymann and Belnap, 2007], Dynamo [Castaño Díez et al., 2012] and Protomo [Winkler et al., 2009], normally use Constrained Cross-Correlation (CCC) because it accounts for the "missing wedge" problem and constrains the similarity calculation only to the commonly sampled region in Fourier space [Förster et al., 2008].

Equation 2.15 is a non-convex optimization problem with $6n$ degrees of freedom (DoFs) because each subtomogram has 3 translational parameters and 3 rotational ones to be determined. A common strategy for solving this problem is the expectation-maximization algorithm, where A is iteratively determined and used for calculating the new $\boldsymbol{\tau}$ and $\boldsymbol{R}$ for the next iteration [Hrabe et al., 2012]. For the translation determination there is an efficient algorithm based on Fourier transform [Roseman, 2003], while the rotation is normally sampled with a certain angular step to determine the maximum of the similarity score. As a consequence, the computational cost is enormous. The expectation-maximization algorithm requires an initial model, which can be obtained either by lowpass filtering a reference derived from other sources [Walz et al., 1997, Brandt et al., 2009], or by a *de novo* approach based on alignment by classification where the subtomograms are grouped according to the orientation [Bartesaghi et al., 2008, Winkler et al., 2009].

In the case of limited angle tomography, the averaging step in Equation 2.15 is typically not a direct sum of all the aligned subtomograms because the sampling in Fourier space is normally nonuniform. Mathematically, given a set of aligned subtomograms $V_i', \ldots, V_n'$ and their corresponding sampling regions in Fourier space $\omega_1, \ldots, \omega_n$, the average A can be calculated as:

$$A = FT^{-1}\left(\frac{FT(\sum_{i=1}^{n} V_i')}{\sum_{i=1}^{n} \omega_i}\right). \tag{2.16}$$

Here, FT and FT^{-1} are the forward and inverse Fourier transforms, respectively. In this way, the average is weighted in Fourier space according to the sampling density.

2.2.8 Subtomogram Classification

Subtomogram averaging assumes that all the subtomograms represent the same conformational state of a macromolecular complex. However, this is rarely the case because *in situ* macromolecules typically adopt various conformations to fulfil their tasks. Classification of subtomograms is beneficial to reveal these conformational changes and to improve the resolution.

Major classification methods include: (i) Principle Component Analysis (PCA)-based approaches [Walz et al., 1997, Bartesaghi et al., 2008, Förster et al., 2008], in which all the subtomograms are first aligned to a single reference and the similarity matrix is then calculated for each pair of the subtomograms. The CCC, in which two volumes are correlated only in their commonly sampled regions in Fourier space, is typically used as the similarity measure. Afterwards, the similarity matrix is subjected to PCA analysis to reduce the dimensionality and thus also the noise influence, followed by K-means or hierarchical clustering. Alternative PCA-based classification approaches are probabilistic principal component analysis with expectation maximization [Yu et al., 2010] and wedge-masked differences-corrected PCA [Heumann et al., 2011]. (ii) Maximum likelihood approaches [Scheres et al., 2009, Stölken et al., 2011], which formulate the classification problem statistically and calculate the probability of observing a subtomogram for a given reference. They try to estimate the hidden parameters from the observed data to maximize the probability. (iii) Multi-reference alignment and classification [Bartesaghi et al., 2008, Winkler et al., 2009, Xu et al., 2012, Frank et al., 2012]. Here, the basic idea resembles k-means clustering. The subtomograms are iteratively aligned to a set of references and their class labels are assigned to the most similar references. After each iteration the references are updated by averaging the subtomograms with the new class assignment. The whole procedure terminates when it converges or a predefined number of iteration is reached.

2.2.9 Resolution Estimation

After obtaining the subtomogram average, one important question is how to estimate its resolution, which indicates the maximum spatial frequency at which the information can be considered significant. Reliable resolution estimation in CET is nontrivial because: (1) the imaging parameters of the microscope are hard to determine; (2) the noise level is high; and (3) the data processing might introduce artefacts. As a result, there is no single universal criterion to estimate the resolution. All the current methods provide different insights into this topic from different perspectives.

The commonly used criterion in this field is based on Fourier Shell Correlation (FSC) [Saxton and Baumeister, 1982, Harauz and van Heel, 1986], which is a 1D function of spatial frequency containing correlation coefficients between two volumes in the Fourier space over shells of same resolution (Figure. 2.7). Mathematically, given two volumes V_1, V_2 and their corresponding Fourier transforms $F_1 = \mathcal{F}(V_1)$, $F_2 = \mathcal{F}(V_2)$, the FSC at band r is calculated as:

$$FSC(r; V_1, V_2) = \frac{\sum_{r_i \in r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt{\sum_{r_i \in r} \mid F_1(r_i) \mid^2 \cdot \sum_{r_i \in r} \mid F_2(r_i) \mid^2}}. \tag{2.17}$$

Computing FSC for each band r results in a 1D function, for which a cutoff value can be chosen as the resolution threshold. Typical choices of the threshold are 0.5, 0.33 and 0.143 [Rosenthal and Henderson, 2003, Penczek, 2010].

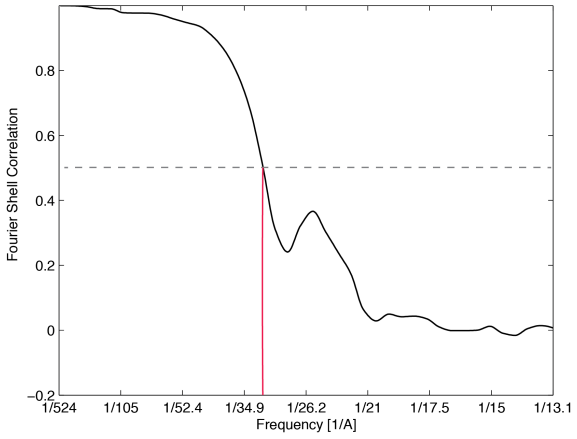

Figure 2.7: A typical Fourier shell correlation curve. The resolution determined here using 0.5 criterion is 29.9 Å.

There are three main types of FSCs: (1) Half-set (pairwise) FSC. It is calculated during the alignment procedure by splitting the subtomograms into two half sets and averaging them separately. The two half-set averages are then used for

calculating the FSC and determining the resolution of the overall average. Half-set FSC measures the self-consistency of the result. However, it has been reported that the noise might also be aligned by this approach, which leads to an overestimated resolution [Grigorieff, 2000, Stewart and Grigorieff, 2004]. (2) Gold-standard FSC. Here, the subtomograms are first split into two halves and aligned separately. The two resulting averages are compared using the FSC. The advantage is the reduction of the noise bias because the alignment procedures are carried out independently on each half set. (3) Cross-resolution FSC. When a high-resolution structure is available (e.g., from X-ray crystallography or SPA), it can be used to compute the FSC against the subtomogram average to estimate its resolution. In this thesis, these three resolution estimations are conducted as many as possible for each case in order to provide a comprehensive resolution analysis.

Overall, the FSC is easy to compute and is able to quantify the consistency of information contained in each frequency. Nevertheless, it has several shortcomings: (1) It is controversial for setting a proper FSC threshold for the resolution determination [van Heel and Schatz, 2005]. (2) The calculation of the FSC assumes that the signal and noise are uncorrelated. However, this might not be the case because the subtomograms have to be aligned prior to computing the FSC and the correlation might be introduced by the alignment. (3) Another assumption is that the signal contained in two averages is the same, which might also be violated if the dataset is heterogeneous.

3

Tomogram Reconstruction

3.1 Introduction

After the micrographs of the sample are acquired using a TEM, the tomogram can be reconstructed from those 2D images, which are approximately parallel projections of the object [Hawkes, 2006]. In practice, the performance of the reconstruction algorithm, which inverts the projection process, determines the accuracy of the 3D reconstruction of the sample. In CET, major challenges for the reconstruction process are the low SNR of the micrographs and the limitation of the projection angle (section 2.2.5).

Except for the commonly used methods mentioned in section 2.2.5, considerable advances have been made to solve the inverse problem of reconstructing an object from projections, especially in the medical imaging field. There is a trend towards iterative reconstruction algorithms. It is attractive to use Fourier-based interpolation methods in such iterative schemes due to their high accuracy and speed compared to real-space based approaches. For instance, [Fessler and Sutton, 2003] introduced the min-max interpolation for nonuniform fast Fourier transform and later combined it into an iterative procedure for 2D tomographic reconstruction [Matej et al., 2004]. Potts and co-workers [Knopp et al., 2007] introduced a method, which is referred to as Iterative Nonuniform fast Fourier transform (NUFFT) based Reconstruction method (INFR) in the following. In this method the reconstruction is formulated as an algebraic optimization problem, which is solved using the conjugate gradient method and NUFFT. INFR has been shown to result in excellent reconstructions when applied to magnetic resonance imaging data, but it has not been applied to CET data. In particular, it has not been characterized to what extent the excellent interpolation characteristics of INFR are beneficial to obtain meaningful information in parts of the missing wedge.

Here, INFR is adapted to reconstruct tomograms from cryo-electron micrographs and the reconstruction quality is compared to the state-of-the-art methods. Specially, the main contribution includes an efficient implementation based on the single-

axis tilt geometry and an analytic means to compute the density compensation matrix. Simulations show that the reconstructions obtained by INFR are more accurate than reconstructions using NN direct inversion method and WBP (section 2.2.5) for tilt series covering the complete angular range. More importantly, the behavior of INFR under the "missing wedge" situation (limited angle tomography) is studied in detail. For restricted angular sampling, INFR is capable of retrieving meaningful information in some regions of the missing wedge in Fourier space, in particular in the low frequency regime. When applied to experimental CET data, the improved reconstruction accuracy of INFR in the low frequencies has important consequences: sensitivity and accuracy of particle localization by template matching are increased considerably and subtomogram averaging yields higher resolution results due to more accurate subtomogram alignment.

This chapter is based on a previous publication [Chen and Förster, 2013].

3.2 Nonuniform Fast Fourier Transform

For reconstruction of cryo-electron tomograms INFR was implemented. In the following, the method and its specific implementation are explained.

First, the NUFFT [Keiner et al., 2009] is briefly discussed, which is the basis of the reconstruction algorithm described here. Given a function $f(x), x \in I_N$ and $I_N := \{x = (x_t)_{t=0,\ldots,d-1} \in \mathbb{Z}^d : -N/2 \leq x_t \leq N/2\}$ (the equispaced grid) as the input, NUFFT tries to evaluate the following trigonometric polynomial efficiently at the reciprocal points $k_j \in [-1/2, 1/2)^d, j = 0, \ldots, M-1$:

$$\hat{f}(k_j) := \sum_{x \in I_N} f(x) e^{-2\pi i x k_j}. \tag{3.1}$$

In contrast to the regular discrete Fourier transforms, k_j can be on an arbitrary nonuniform grid. In matrix vector notation, Equation 3.1 can be rewritten as

$$\hat{f} = Af \tag{3.2}$$

with the nonequispaced Fourier matrix $A := e^{-2\pi i x k_j}, x \in I_N, j = 0, \ldots, M-1$.

One approach for fast computation of Equation 3.2 is based on the factorization $A \approx BFD$ [Potts et al., 2001], where D is the inverse Fourier transform of a window function w, F is the oversampling Fourier matrix and B is a sparse matrix of the window function w with the cut-off parameter m, which contains at most $(2m+1)^d$ non-zero entries per row (Figure 3.1). The basic idea of this factorization, which resembles the reverse gridding method [Penczek et al., 2004], is the following: the accurate interpolation in Fourier space to a different grid is achieved by convolution with an appropriate window function w, which is compensated for via prior division by the inverse Fourier transform of w. The accuracy of this approach depends on the

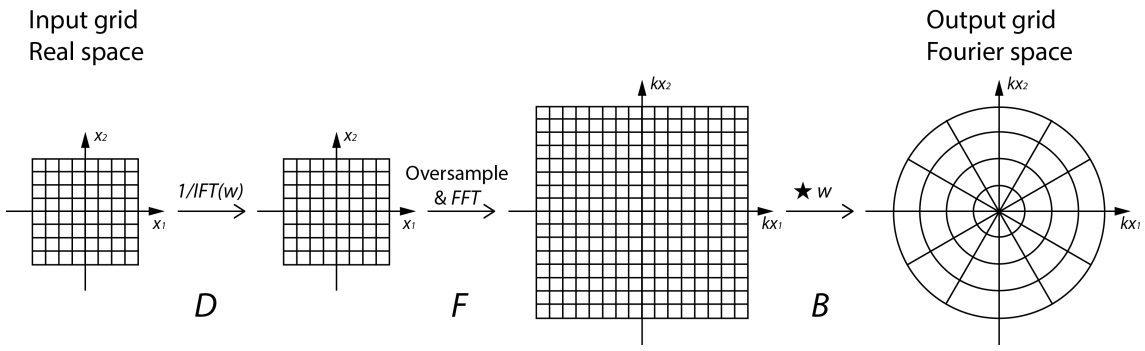

Figure 3.1: Fast computation of Equation 3.2, where $A \approx BFD$. It consists of three steps: 1. Multiply with a diagonal matrix D, which contains the inverse Fourier transform (IFT) of the window function w (simultaneously in space and frequency localized, e.g., Kaiser-Bessel window). 2. Do a Fourier transform and oversampling in Fourier space, i.e., multiply with a matrix F in matrix notation. 3. Convolute with a matrix B of the window function w. B is a sparse matrix that contains at most $(2m+1)^d$ non-zero entries per row and m is a predefined cut-off parameter specifying the maximum width of w for the sake of efficient computation.

oversampling factor and the choice of window function w and its cut-off parameter m. It has been suggested that the Kaiser-Bessel window function provides high accuracy and a typical choice of m would be 3 for an oversampling factor 2 [Fessler and Sutton, 2003, Jackson et al., 1991].

The adjoint (or conjugate transpose) NUFFT is defined as the sum

$$f'(x) := \sum_{j=0}^{M-1} \hat{f}(k_j) e^{2\pi i x k_j}, x \in I_N \tag{3.3}$$

or in matrix vector notation

$$f' = A^H \hat{f} \tag{3.4}$$

Its efficient computation can be analogously achieved by the factorization $A^H \approx D^T F^T B^T$. It has been shown that the gridding method can be seen as an efficient algorithm to compute A^H [Potts et al., 2001].

3.3 Iterative Reconstruction Scheme

Without loss of generality, the reconstruction of a 2D image from 1D projections is considered for the sake of simplicity. According to the projection-slice theorem the Fourier transform of a projection corresponds to a slice in the Fourier space of the object [Bracewell, 1986]. Given M observations b as the Fourier transforms

of projections, we define the matrix A as a nonuniform Fourier transform matrix depending on the sampling geometry:

$$A := e^{-2\pi i x k_j}, k_j \in [-1/2, 1/2)^2, j \in [0, \ldots, M). \tag{3.5}$$

The reconstruction problem is to recover $f(x), x \in \{-N_1/2, \ldots, N_1/2 - 1\} \times \{-N_2/2, \ldots, N_2/2 - 1\}$ defined on a regular grid, such that:

$$Af = b. \tag{3.6}$$

Solving Equation 3.6 can be formulated as an optimization problem:

$$\tilde{f} = \arg\min \parallel b - Af \parallel . \tag{3.7}$$

This is a least square problem and when the sampling density compensation [Pipe and Menon, 1999] is considered its solution requires solving the following equation:

$$A^H W A f = A^H W b. \tag{3.8}$$

Here $W := diag(w_m)$ is the density compensation matrix, which account for the nonuniform distribution of the sampling in Fourier space (Figure 3.2b). For example, in CET the sampling is very dense towards zero frequency and thus the projections contain to some degree redundant information in these frequencies [Crowther et al., 1970]. It is important to weigh the information in Fourier space according to the overall sampling pattern because low frequencies would be artificially enhanced otherwise.

Equation 3.8 can be solved using the conjugate gradient method on the normal equations [Saad, 2003], in which the matrix vector multiplications in this optimization algorithm are substituted by the (adjoint) nonuniform Fourier transform operators. The details are described in Algorithm 3.1. Throughout the iterative optimization the residual decreases and agreement of the reconstruction with the observations increases. Interestingly, it can be shown [Knopp et al., 2007] that the result of the first iteration of the conjugate gradient optimization starting from a void volume is similar to the solution of the gridding method. When optimization is continued, the iterative approach becomes more accurate than the gridding method [Bronstein et al., 2002].

Nevertheless, excessive optimization is harmful under the situation of the ill-posed inverse problem (limited angle tomography) and the high noise level. There is no straightforward or clear stopping criterion under such situations. The ill-conditioning will cause a semi-convergence behavior [Hansen, 1998, Qu et al., 2005], which manifests itself as an initial convergence towards the true solution and later divergence. This situation is aggravated by the noise. If the noise variance ε is known, in order to prevent the noise overfitting, a conservative criterion is to stop the iteration when:

$$\parallel A^H W b - A^H W A \tilde{f} \parallel < \varepsilon. \tag{3.9}$$

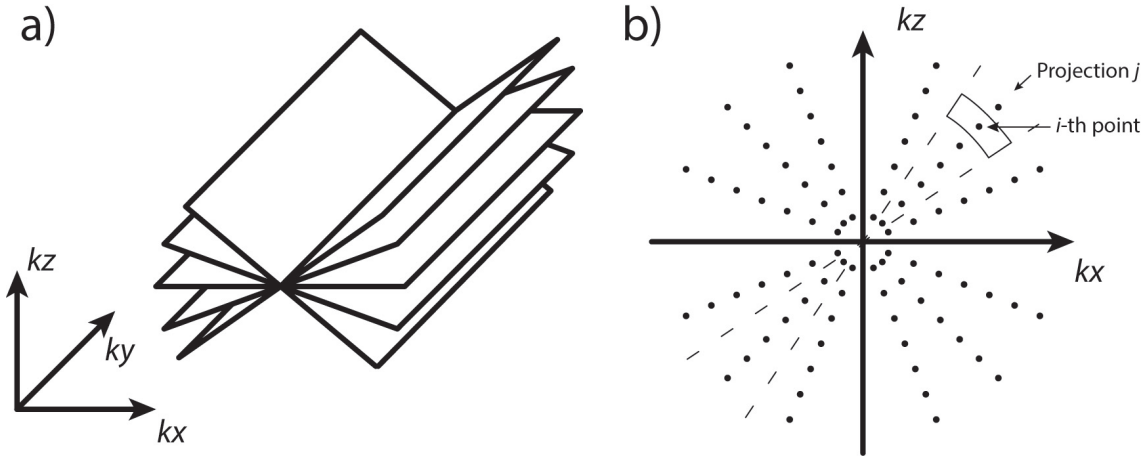

Figure 3.2: Sampling and weighting in Fourier space. (a) The sampling pattern of single-axis CET in k-space. (b) One slice view of the k_x-k_z plane, assuming y-axis is the tilt axis. The black dots are the sampling points and the outlined area around the sampling point is its corresponding weight.

Algorithm 3.1 Solving Equation 3.8 via the conjugate gradient method

Input: A, A^H, W, b, k: Number of iterations to run.
Output: $\tilde{f}_k$

$\tilde{f}_0 = \mathbf{0}$
$r_0 = b - A\tilde{f}_0$
$Z_0 = A^H W r_0$
$P_0 = Z_0$
for $i = 0, \ldots, k-1$ **do**
  $V_i = AP_i$
  $\alpha_i = Z_i^H Z_i / V_i^H W V_i$
  $\tilde{f}_{i+1} = \tilde{f}_i + \alpha_i P_i$
  $r_{i+1} = r_i - \alpha_i V_i$
  $Z_{i+1} = A^H W r_{i+1}$
  $\beta_i = Z_{i+1}^H Z_{i+1} / Z_i^H Z_i$
  $P_{i+1} = Z_{i+1} + \beta_i P_i$
end for
return $\tilde{f}_k$

Here, the number of iterations is a compromise between the accuracy of the reconstruction and the overfitting. Essentially all iterative reconstruction methods face this problem. The common approach is to simply terminate the reconstruction after certain number of iterations, which is specified by the user according to typically subjective criteria. Alternatively, it is suggested here to determine the stopping cri-

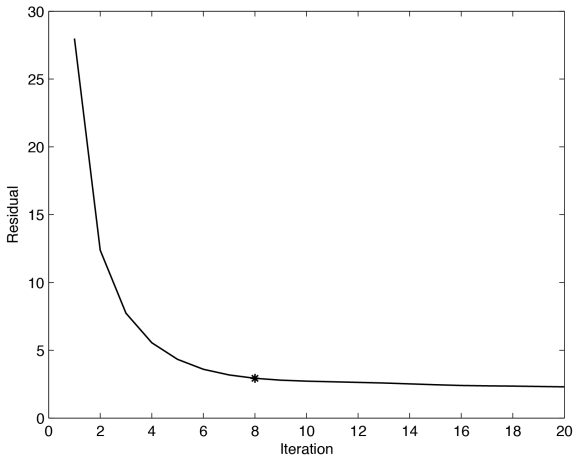

Figure 3.3: The residual of INFR reconstruction of an experimental tomogram (yeast lysate, Figure 3.5) against the number of iterations. It typically has an L-shape and the star labels the "elbow" of the L-curve, which is the lower bound of the necessary number of iterations. It can be used as an indicator of the stop criterion for INFR. According to the experience, a few more iterations beyond the elbow help INFR to achieve a better performance, as long as no visual artifact appears.

terion objectively based on the behavior of the residual as a function of the number of iterations. In most cases where the reconstructions are well-behaved, the residuals during the iterations typically adopts an "L" shape (Figure 3.3). As a criterion for the number of iterations to be used for the reconstruction, identifying the "elbow" of the L-curve is typically a reasonable option [Qu et al., 2005].

3.4 Reconstruction of Single-Axis Cryo-Electron Tomograms

The above-described reconstruction method is analogous in the 3D case. In this section, the density compensation for single axis tomography is explained. In particular, the efficient implementation of the algorithm is described here because the use of iterative reconstruction algorithms tends to be limited by extensive computation time, especially in the 3D case.

For general projection geometries, the density compensation matrix W can be determined by calculating the Voronoi diagram according to the sampling geometry. Although there exists an efficient algorithm for that [Barber et al., 1996] with the computational complexity of $O(n \log(n))$, the computation is nevertheless expensive, in particular in the 3D case. Moreover, this approach often faces difficulties at the sampling points on the periphery where the Voronoi cell is unbounded, or the areas

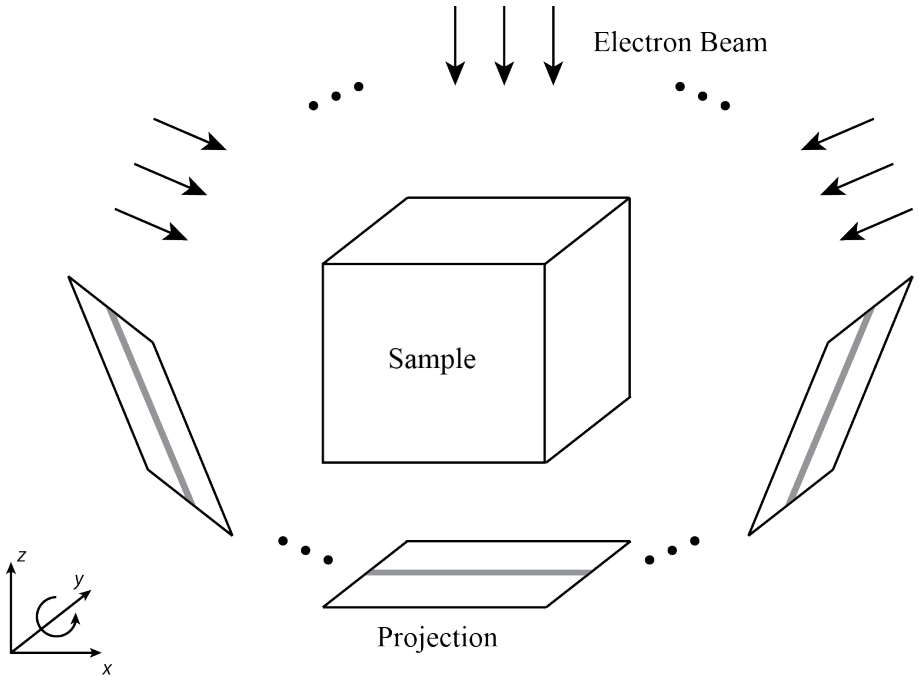

Figure 3.4: Single-axis tilt geometry in CET. Here, the tilt axis is y. All the grey strips with the same y index constitute the $b_y, y \in [1, ny]$. The corresponding sampling pattern in Fourier space is illustrated in Figure 3.2.

where there are major sampling gaps in Fourier space.

For the single-axis tilt geometry (Figure 3.4), an analytical way for the computation of W is proposed, which is simplified and computationally much faster than the Voronoi approach. Specifically, assuming 2D projections have been aligned (section 2.2.3) and the CTF corrections (section 2.2.4) have been conducted prior to the reconstruction, the original 3D reconstruction problem (Equation 3.6) can be reformulated as (the tilt axis is the y-axis):

$$\begin{cases} A' \cdot f_1 = b_1, \\ \dots \\ A' \cdot f_{ny} = b_{ny}. \end{cases} \tag{3.10}$$

Here, A' is the new 2D sampling matrix and is identical in the xz-plane for each y (Figure 3.2a). That means, solving Equation 3.6 can be decoupled into solving a set of independent linear equation systems, each of which is a 2D reconstruction with respect to the xz-plane. Therefore, the calculation of the weights in 3D can be reduced to 2D.

For each of the 2D pattern, the weight of i-th sampling point ($i = 0, \dots, N-1$)

along projection angle α_j ($j = 0, \ldots, M-1$) can be approximated as (Figure 3.2b):

$$w_{i,j} = \begin{cases} |\alpha_{j+1} - \alpha_{j-1}| \times |i - N/2|, & i \neq N/2, j \neq 0, M-1 \\ 2|\alpha_1 - \alpha_0| \times |i - N/2|, & i \neq N/2, j = 0 \\ 2|\alpha_{M-1} - \alpha_{M-2}| \times |i - N/2|, & i \neq N/2, j = M-1 \\ \pi/4M, & i = N/2. \end{cases} \tag{3.11}$$

In this way, the overweighting of the sampling points close to the missing wedge area in Fourier space is avoided. In general, these weights resemble the "ramp" function in the WBP algorithm, when the tilt angles are equally spaced [Harauz and van Heel, 1986].

After calculating the weights for the sampling points, the method described above can be used to reconstruct the tomograms. Due to the single axis tilt geometry the sampling pattern is of the type 2D×1D. Therefore, a significant speedup can be achieved by first performing 2D reconstructions in k_x-k_z planes and then an overall reconstruction by 1D FFTs along the k_y dimension. Currently, the implementation of the reconstruction algorithm is restricted to cubic volumes.

3.5 Tomogram Simulation

In this section, the means of tomogram simulation is presented, which serves as the foundation of performance evaluations on the simulated dataset. To simulate the tomograms as realistically as possible, the protocol described in [Beck et al., 2009, Förster et al., 2008] was followed. In short, the simulations were carried out as follows: the electron optical 3D density of an atomic model was approximated by summing up the atomic numbers on a Cartesian grid with the specified pixel size and was low-pass filtered. Subsequently, a 2D CTF function was simulated with a given defocus value. The density map was projected according to the tilt angles and angular increment. Gaussian-distributed noise (CTF noise) was added to each projection and the CTF was applied. Furthermore a non-CTF affected portion of noise (MTF noise) was added and finally the tomogram was reconstructed from the projections.

3.6 Experimental Tomograms from Yeast Lysate

For application of INFR to the experimental tomograms, the yeast cytosol was prepared by following a protocol that was slightly modified from [Brodsky, 2010]. Yeast cells were grown in YPD medium to mid log phase at 30 °C. Cells were harvested and washed in water, before they were resuspended in a minimal amount of Lysis buffer (20 mM Hepes at pH 7.7, 100 mM KOAc, 2 mM $Mg(OAc)_2$, 1 mM

DTT, 0.5 mM PMSF, protease inhibitor). The cells were added drop-wise into liquid nitrogen and broken by manual crushing with a pistil. The resulting powder was resuspended in a minimal amount of lysis buffer. The lysate was then centrifuged (15 min at 18000 rcf in an SS34 rotor) and 4.5mL aliquots of the supernatant were applied to a fine G25 column, equilibrated in TE-buffer (20 mM Hepes, 100 mM KOAc, 2 mM $Mg(OAc)_2$, 2 mM DTT, 0.5 mM PMSF, 10% glycerol). Fractions with high OD_{280} were pooled, aliquoted, flash-frozen in liquid nitrogen and stored at $-80\,°C$.

Ribosomes from 200 µl of thawed yeast lysate were pelleted in a Beckman Optima TLX ultracentrifuge (TLA 100, 30 min, 190,000 g, 4 °C) and were resuspended in 40 µl ribosome buffer (5 mM $MgCl_2$, 140 mM KCl, 10 mM HEPES pH 7.4, 1 mM DTT, protease inhibitor). 3 µl of resuspended ribosomes were applied to lacey carbon molybdenum grids (Ted Pella, USA). After an incubation time of 60 s, 3 µl of 10 nm colloidal gold in ribosome buffer were added to the grid and the sample was vitrified in liquid ethane using a Vitrobot Mark IV (FEI, Netherlands). The tilt series were acquired using a FEI Titan Krios transmission electron microscope (acceleration voltage of 300 kV) equipped with a 4k×4k FEI Falcon direct electron detector. Single-axis tilt series were acquired from $-60°$ to $+60°$ with an angular increment of 3° at different nominal defocuses (2, 3, 4, 5 µm) and an object pixel size of 0.288 nm using the FEI tomography acquisition software. The cumulative electron dose did not exceed 60 electrons/Å^2.

3.7 Results

3.7.1 Reconstruction of Simulated Data

INFR was first evaluated on simulated data. Following the protocol described in section 3.5, the 3D density of the *S. cerevisiae* translating 80S ribosome was simulated from atomic models (Protein Data Bank ID: 3IZB 3IZE 3IZF 3IZS) and projected it in different tilt angles using 3° as the angular increment. The projections were then used for the reconstruction. For comparison, the volumes were reconstructed using WBP and NN direct inversion methods implemented in the SPARX software [Hohn et al., 2007], which yielded the most accurate reconstructions among the software packages used in the tests here (TOM [Nickell et al., 2005], EMAN2 [Tang et al., 2007], SPARX [Hohn et al., 2007], PyTom [Hrabe et al., 2012]). The reconstruction accuracy was finally measured by FSC.

The results are plotted in Figure 3.5a. If the tilt angles range from $-90°$ to $90°$ (no missing wedge), the FSC of the reconstruction using INFR exceeds that of the other reconstructions essentially over the entire frequency range. In particular, the FSC remains essentially one until very close to the Crowther frequency [Crowther et al., 1970] for the simulation.

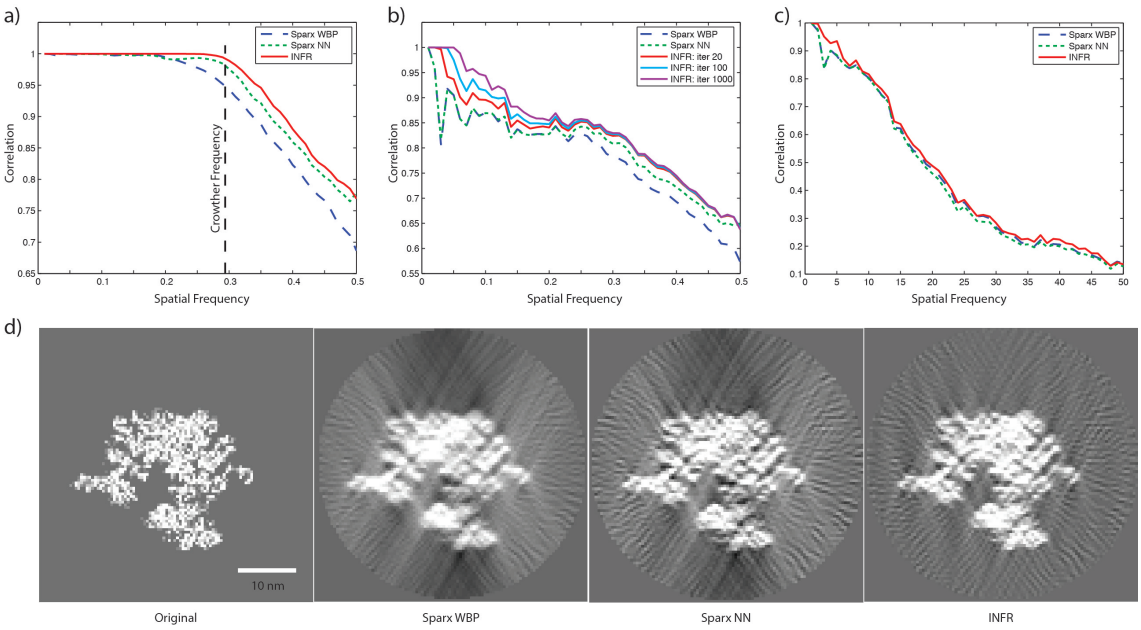

Figure 3.5: Comparison of reconstructions of simulated 80S ribosomes. (a) The FSC curves of the reconstructions using projections from $-90°$ to $90°$ with an angular increment of $3°$. (b) The FSC curves of the reconstructions using projections from $-60°$ to $60°$ ($3°$ angular increment). (c) The FSC curves of the reconstructions using projections with SNR=1 from $-60°$ to $60°$ ($3°$ angular increment). (d) Central slice views of XZ plane of reconstructions from $-60°$ to $60°$ using various methods.

If the tilt angles are limited from $-60°$ to $60°$, the improvement in reconstruction quality achieved by INFR is more apparent (Figure 3.5b). Strikingly, INFR could even fill some missing wedge areas in Fourier space with meaningful information, if the sampling is dense enough around that area. By design, the other two methods set the coefficients in the missing wedge area essentially to zero throughout and consequently no meaningful information is placed in those sectors. The filling of meaningful information in the part of the missing wedge area gets apparent when the FSC is plotted throughout the course of the conjugate gradient optimization: the reconstruction accuracy continuously improves, due to the reconstructed coefficients in the missing wedge area.

To compare the noise sensitivity of the three methods, Gaussian white noise was added to the projections of the simulated tilt series (SNR=1) and again reconstructed the volume from limited tilt angles ranging from $-60°$ to $60°$. The FSC was then measured against the original tomogram and the results are shown in Figure 3.5c. This analysis indicates that the reconstruction accuracy of the INFR is also superior in the presence of noise. In particular, the generation of meaningful information by INFR in the missing wedge area, especially in the low-frequency regime,

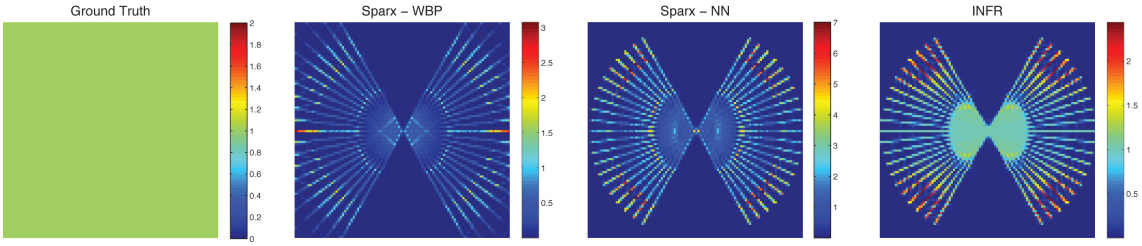

Figure 3.6: Reconstructions of Dirac delta function for limited angle tomography. The power spectra of the reconstructions of the Dirac delta function with projections from $-60°$ to $60°$ ($3°$ angular increment) for the different reconstruction methods.

	SPARX WBP (s)	SPARX NN (s)	INFR (s/iter)
Volume size 50^3	0.46	1.69	0.09
Volume size 100^3	0.85	7.43	0.39
Volume size 200^3	2.68	46.7	1.86

Table 3.1: Elapsed time for various reconstruction methods. The reconstruction time is evaluated for cubic volumes reconstructed from 41 projections of dimensions 50×50, 100×100, and 200×200, respectively. All tests were performed on a machine equipped with Intel Xeon Processor X5570 of clock speed 2.93 GHz and 48 GB main memory.

is also observed in the presence of noise.

To better understand the generation of meaningful coefficients in the missing wedge area, the Dirac delta function (in 3D) was reconstructed from limited angular projections. An empty volume was generated with only one voxel set to 1 at the center and subsequently projected from $-60°$ to $60°$. After reconstruction using different methods, the power spectrums of each method were examined, which are shown in Figure 3.6. It clearly demonstrates the advantage of INFR by having a uniform power spectrum inside the sampled area in Fourier space, i.e., within the limit of the Crowther frequency. Moreover, the "leaking" of information from the sampled area into the missing wedge can be observed, demonstrating an extrapolation of data from the sampled area into the missing wedge.

Furthermore, the speed of various reconstruction methods was evaluated. Due to the iterative scheme of INFR, its computational time depends on the number of computed iterations. The elapsed time of reconstructing volumes of difference sizes were measured and the results are shown in Table 3.1, from which we can see that INFR per iteration is fastest. However, typically many iterations are necessary to achieve better accuracy. A trade-off has to be made.

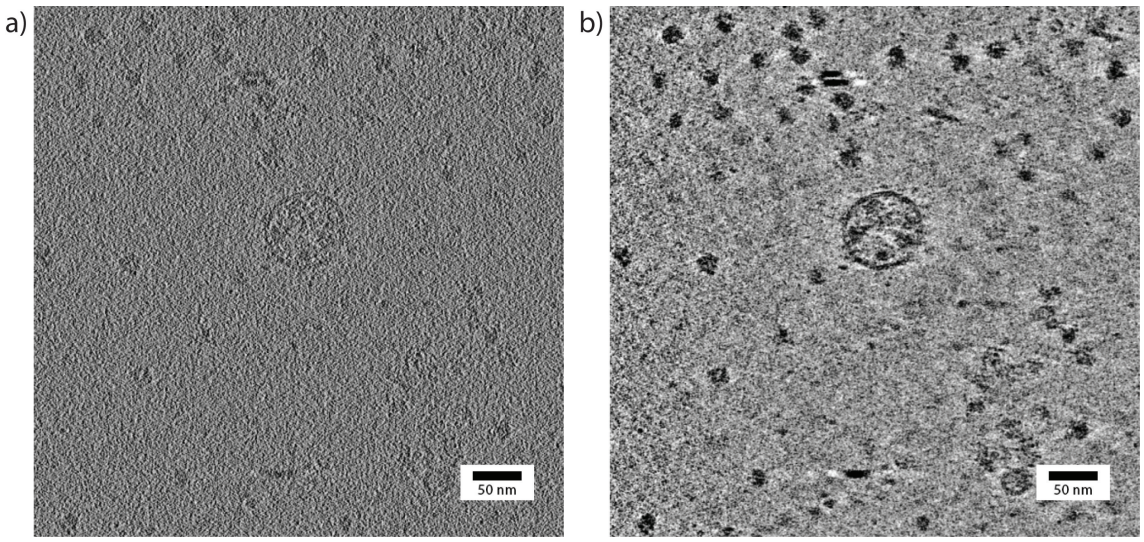

Figure 3.7: Comparison of reconstructions from experimental data. A slice from the XY plane of the reconstructions of *S. cerevisiae* lysate using WBP and INFR, respectively, is shown.

3.7.2 Reconstruction of Experimental Data

INFR was furthermore tested on experimental data of 80S ribosome from yeast lysate (section 3.6), from which one tomogram of 5 µm nominal defocus was chosen. The downsampled tomogram (voxel size 2.3 nm) was reconstructed using INFR and WBP as a comparison. Due to practical reasons the NN method was excluded from further analysis: the implementation is not originally designed for subtomogram reconstruction and it would be difficult to integrate it into the workflow. The simulations indicate that the main difference in the INFR reconstruction is in the low-resolution part of the missing wedge, which is reconstructed similarly in WBP and the NN method. Hence, the comparison of INFR to WBP is sufficient to assess the main consequences of INFR.

One slice of the reconstructed tomogram can been seen in Figure 3.7. INFR was terminated after 20 iterations, which were sufficient for convergence according to the plot of the residual (Figure 3.3). For comparison a reconstruction by WBP was performed.

The performance of INFR was evaluated quantitatively with respect to two aspects: template matching and subtomogram averaging. After reconstruction, template matching was carried out, followed by the determination of the 1,000 highest correlation peaks [Förster et al., 2010]. The plot of the correlation values of these peaks as a histogram typically results in a Gaussian distribution for true positive matches as well as a background distribution that increases strongly towards lower correlation values. Depending on the quality of the tomograms the Gaussian curve

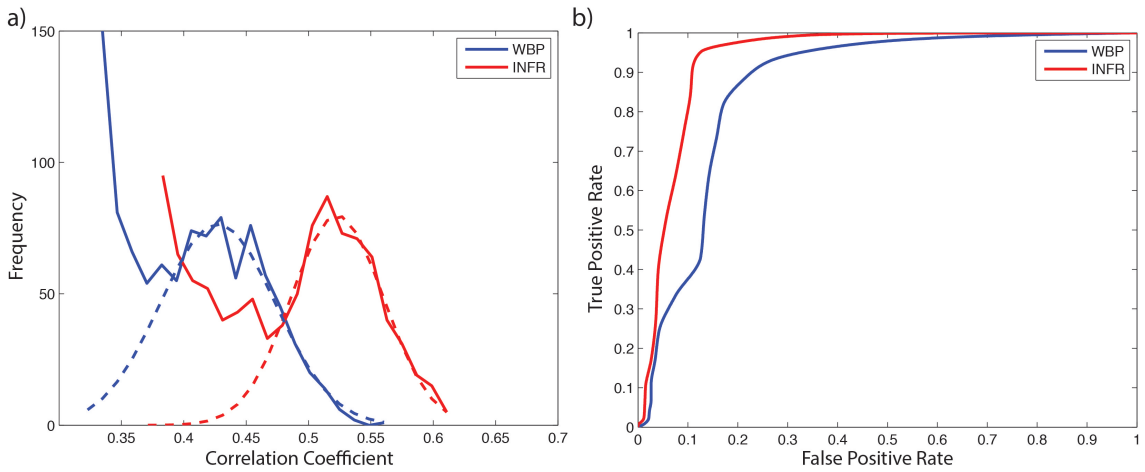

Figure 3.8: Effect of reconstruction method on template matching performance. (a) The histograms of correlation coefficients from 1000 picked candidates using WBP and INFR. The dashed lines correspond to their respective Gaussian fittings. (b) The ROC curves for WBP and INFR. $TPR = TP/P$, $FPR = FP/N$.

corresponding to true positives is well separated from the background or it may not be discernable at all in the worst case.

The histograms of the correlation coefficients are plotted in Figure 3.8a for the tomograms reconstructed with the different methods. The correlation values are substantially higher for INFR. More importantly, if we fit Gaussians to the first halves of the histograms and assume these detections are mostly true positives, the Receiver Operating Characteristic (ROC) curves can be plotted (Figure 3.8b), i.e., the estimated true positive rate against the estimated false positive rate. The ROC curves clearly indicate a better separation of the true positives from false positives for INFR than for WBP.

To compare the reconstruction quality in terms of subtomogram averages, unbinned subtomograms (160^3, voxel size 2.88 Å) of the top 443 candidates were reconstructed, which contain about 70% of all ribosomes in the tomogram according to the histogram analysis. For a fair comparison, the same candidates, the ones from identification in the INFR volume, were used. These subtomograms were then translationally and rotationally aligned with respect to each other using the fast subtomogram alignment algorithm presented in chapter 5. Figure 3.9a shows the pairwise FSC curves of the averages obtained from the subtomograms reconstructed using INFR and WBP. The FSC indicates a slightly higher resolution for the INFR reconstruction. Furthermore, the accuracy of the subtomogram averages was assessed by cross-resolution (Figure 3.9b) with a 7.9 Å resolution single particle reconstruction (EMDB: 1668). The cross-resolution FSC confirms the higher accuracy of the average from INFR compared to that from WBP subtomograms. Note

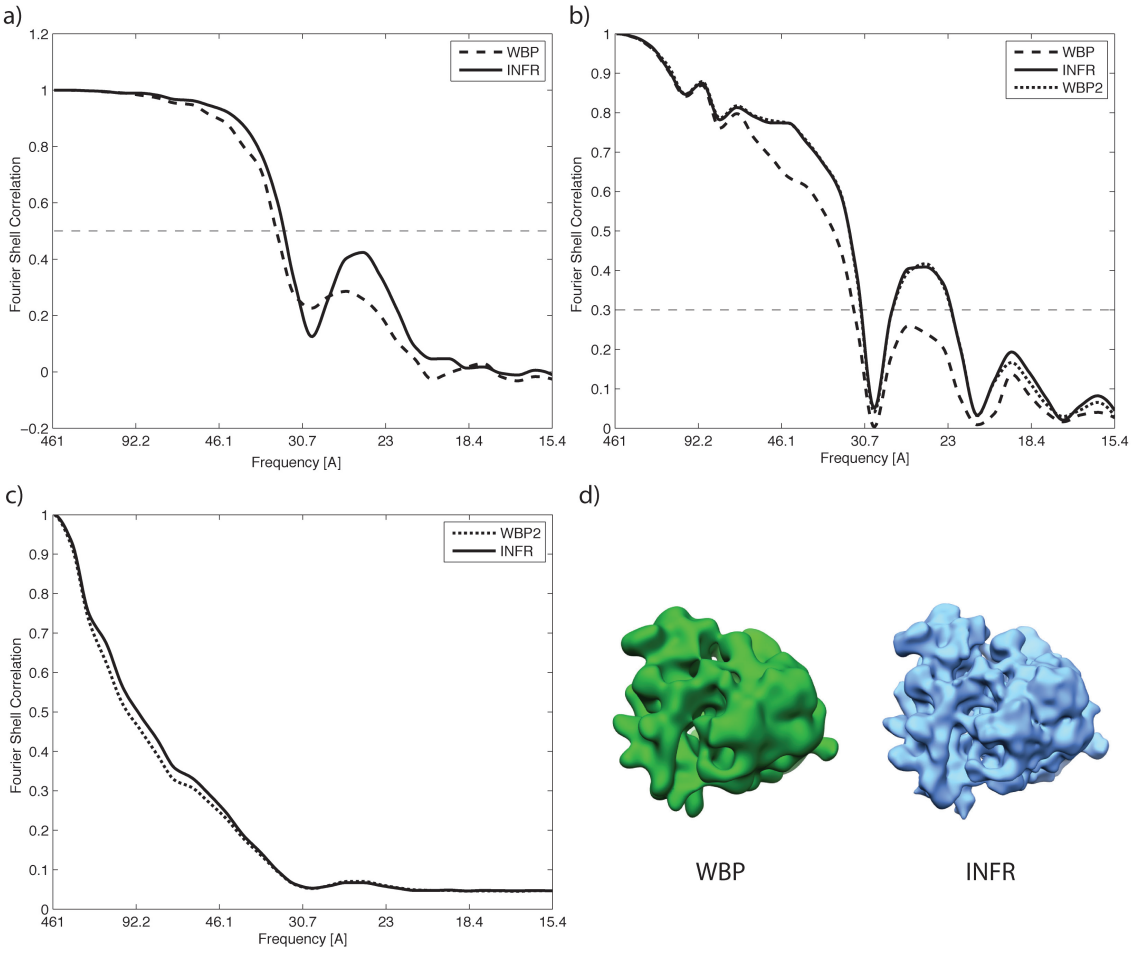

Figure 3.9: Effect of reconstruction method on subtomogram averaging. (a) The pairwise FSC curves of aligned top 443 particles reconstructed using WBP and INFR. (b) Cross-resolution FSC curves for WBP, INFR and WBP2. Here, WBP2 is the average of WBP subtomograms with the alignment parameters determined for the INFR subtomograms. (c) Averaged FSC curves of individual subtomograms against WBP2 and INFR. (d) The subtomogram averages from WBP and INFR, filtered by cross-resolution 0.3 criterion.

that the first zero of the CTF is more pronounced for the average using INFR, which is indicative of less noise and the signal beyond the 1st zero of the CTF is much stronger. Using a FSC threshold of 0.3, which is typically used for cross-resolution [Rosenthal and Henderson, 2003] the INFR subtomogram average is clearly superior to the WBP average (22.8 vs 32.2 Å) (Figure 3.9d).

To only compare the reconstruction quality and exclude the influence of subtomogram alignment, the alignment parameters determined by INFR were used to average the subtomograms reconstructed by WBP. Then the cross-resolution of this

average was measured (we name it WBP2, Figure 3.9b), which is essentially the same as the one of INFR. Thus, the choice of the reconstruction algorithm has little effect on the resolution of the average itself at the sampling chosen in this test (voxel size on the specimen level: 2.88 Å, resolution: ∼22 Å).

To monitor the accuracy of the individual subtomograms, the averaged FSC was calculated (Figure 3.9c) of each individual subtomogram against the averages of WBP2 and INFR, respectively. The averaged FSC of the INFR subtomograms is slightly higher than that of the WBP2 subtomograms in the low frequency part. This suggests the individual subtomogram reconstructed by INFR is more accurate, on average, than the one reconstructed by WBP, which is the reason for the more accurate subtomogram alignment.

3.8 Discussion

Here, an iterative reconstruction algorithm is introduced that makes use of the powerful NUFFT. INFR shows substantial advantages over state-of-the-art approaches, without imposing any prior information or constraint. Specifically, compared to the fast Fourier summation algorithm and gridding method, the major difference is the iterative scheme, which yields superior performance, especially when data sampling is incomplete. For simulations the achieved resolution by INFR is close to the Crowther limit when the sampling angle range is complete. More importantly, for limited angle tomography the missing information in Fourier space is recovered to some extent. The rationale behind this recovery is that the information is spread out in the vicinity of the sampling points with the help of the adjoint nonuniform Fourier transforms during the iterative reconstruction procedure. This extrapolation results in fewer artifacts in real space, especially along the "missing wedge" direction (Figure 3.5d). Due to the benefit of this method, one could not only achieve a better detection rate for template matching, but also a more accurate subtomogram alignment yielding higher resolution subtomogram averages.

Importantly, the generation of meaningful information in the missing wedge by INFR is not due to the use of prior information, which is in contrast to other iterative Fourier-space based reconstruction akin to the Gerchberg-Papoulis algorithm for super-resolution [Gerchberg, 1974, Papoulis, 1975]. In this type of algorithm the finite support of a specimen and its resulting well-defined envelope are used to extrapolate from sampled data into unsampled areas. An implementation of such a method tailored to electron tomography has revealed that the missing wedge can largely be recovered if the object of interest is bounded [Miao et al., 2005]. However, in CET this condition is essentially never met because the objects of interest (e.g., macromolecules, virions, organelles) are embedded in vitreous water, which also gives rise to considerable, nonuniform signal (sometimes referred to as structural noise). INFR does not make such assumption making the method applicable to

CET.

INFR can be easily generalized to reconstruction in single particle analysis or other sampling schemes such as tomography using double-tilt geometry or random conical tilting. In such case, the density compensation has to be calculated in a more complicated fashion, e.g., using Voronoi diagrams. However, the speedup achieved by the 2D×1D sampling pattern for single-axis tomography will not be available anymore. INFR is implemented as an easy-to-use library, which requires only one parameter (the number of iterations) to be specified, or none if the automatic stopping criterion is adopted.

A drawback of INFR compared to WBP and the NN method is its computational speed: it requires a number of iterations to get the final result, which leads to linear growth of the reconstruction time. Due to the efficient computation of the weighting function and the adopt of 2D×1D reconstruction scheme, a speedup of at least one order of magnitude is achieved compared to the NN method per iteration (Table 3.1), which makes the computation time tolerable for routine application in CET. In the tests here, 20 iterations resulted in reconstructions that were more accurate than those obtained using WBP while the computation time was comparable to the NN method.

4

Macromolecule Identification

4.1 Introduction

After tomogram reconstruction an important task is to localize and identify the macromolecule of interest inside the tomogram. This step serves as the basis of the subsequent analysis, subtomogram alignment and averaging. When the identified particles have homogeneous structures, the resulting resolution of the subtomogram average can be greatly improved. Therefore, this step is crucial here and the goal is to produce a set of particles, which contains more true positives and less false positives. However, major challenges are: 1. The identification is hampered by the low SNR of the cryo-tomogram, typically in the range of 0.1-0.01. As a consequence, most of the methods developed in computer vision for 3D object recognition with high SNR cannot be directly applied here. 2. The artefact caused by the "missing wedge" problem creates additional difficulties for the detection. 3. The abundance and spatial information of the macromolecule of interest are normally unknown, which poses a challenge for the computation.

The state-of-the-art method is template matching (section 4.2). Despite its success in CET, it also confronts some challenges: (i) High-contrast features (e.g., membranes and gold fiducials used to align the projections) often generate false positive matches. (ii) Human interaction is required to set the threshold for allowing the discrimination between the true positives and false positives [Ortiz et al., 2006]. To reduce the false positive rate and avoid subjective thresholding, a post-processing step can be appended. For example, linear discriminant analysis of correlation values from different templates in simulations has been used to generate a composite score for classification of the candidates [Beck et al., 2009]. (iii) A template is required as the input. It is typically generated from higher resolution structures obtained by other methods, e.g., x-ray crystallography or SPA. However, they might not be available, especially for the structurally uncharacterized complexes. (iv) The result of template matching might introduce template-bias to the subsequent steps [Subramaniam, 2006], especially when an external high resolution structure is used as the

template. Attention has to be paid to avoid overfitting to the template. (v) Template matching works well for large macromolecular complexes (above 2.5 MDa), such as ribosomes, which have strong contrast in cryo-tomograms. The challenge is to detect complexes with smaller weight, which have less contrast.

Here, an identification protocol is proposed, which uses a supervised learning technique for the binary classification of the candidates from template matching. For each candidate, the features derived from spherical harmonics expansion of subtomograms are used. These features have the advantage of being rotation-invariant and can be computed fast.

A part of this chapter has been published in [Chen et al., 2012].

4.2 Template Matching

In this section, the details of template matching technique are discussed. Template matching is a method to find macromolecular complexes in cryo-tomograms based on structural templates of the macromolecules of interest, i.e., using prior structural knowledge of the target macromolecules. In principle, template matching is a matched filter, which is a commonly used technique in signal processing for detection and localization of patterns corrupted by noise [Turin, 1960]. In CET, the template is compared to all possible same-sized subregions inside the tomogram. In order to take into account of the local imaging variations such as the voxel intensity difference and gray value offset, a normalization has to be employed. Moreover, the macromolecules inside the tomogram typically have unknown orientations, which have to be searched explicitly. In practice, the orientation space is scanned using Euler angles with a specific angular increment. This increment should be chosen according to the Crowther criterion [Crowther et al., 1970], typically oversampled by a factor of at least two. Mathematically, given a object with the diameter d, the relationship between the angular increment $\Delta\alpha$ and the obtainable resolution r can be expressed as [Crowther et al., 1970]:

$$r \approx \Delta\alpha \cdot d. \tag{4.1}$$

A common choice of the angular increment is in the range of 7°-15°. In the end, for each position and orientation, a scoring function is computed, which indicates the similarity between the template and the subvolume of the tomogram. Overall, template matching in CET consists of following steps (Figure 4.1): 1. template generation, 2. scoring function calculation and 3. localization. The details of each step are discussed below.

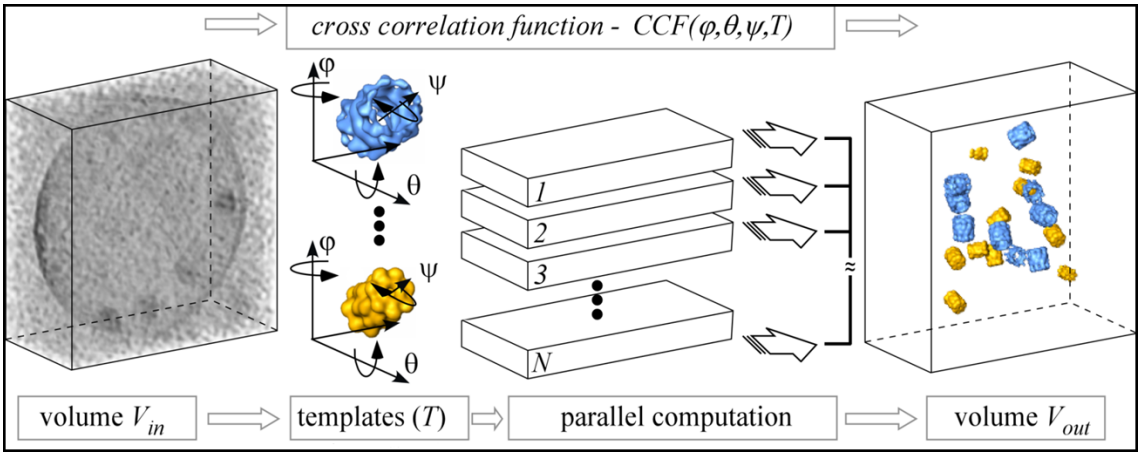

Figure 4.1: Steps of template matching. For a given input volume V_{in}, first the templates T are prepared. Then the cross-correlation function between V_{in} and T under different rotations are calculated. This step can be parallelized to accelerate the computation. Finally, after the peak extraction the macromolecular complexes of interest can be localized as the output V_{out}. Image from [Lucić et al., 2005].

4.2.1 Template Generation

Specifically, the generation of templates from atomic models is presented here. Structure determination of biological macromolecules can be accomplished by different techniques. Arguably the most prolific method is X-ray crystallography, which provides the positions of the atoms that constitute a specific macromolecule. The Protein Data Bank (PDB, `www.rcsb.org`), which serves as an archive for experimentally determined structures of proteins, nucleic acids, and complex assemblies, stores the atomic coordinates of biological macromolecules. These atomic models can be used to generate structural templates for CET.

The key of a successful template is to simulate the imaging process of CET as realistically as possible, which typically requires the following steps (Figure 4.2, [Förster, 2005]): 1. Calculate the electron density according to the coordinates and identities of the atoms specified in the PDB file. For biological materials, the electron optical density is proportional to the electrostatic potential of the macromolecule, which is approximately proportional to the atomic number Z of each element. In this way, the coordinates of the atoms contained in the PDB file are translated into the electron density map and sampled at the pixel size used in the experiment. 2. Convolute the density map with the appropriate point spread function (PSF) of the microscope and low-pass filter it to the corresponding resolution. 3. Rescale the template to the relevant pixel size consistent with the tomogram. The generated template can then be used to calculate the scoring function.

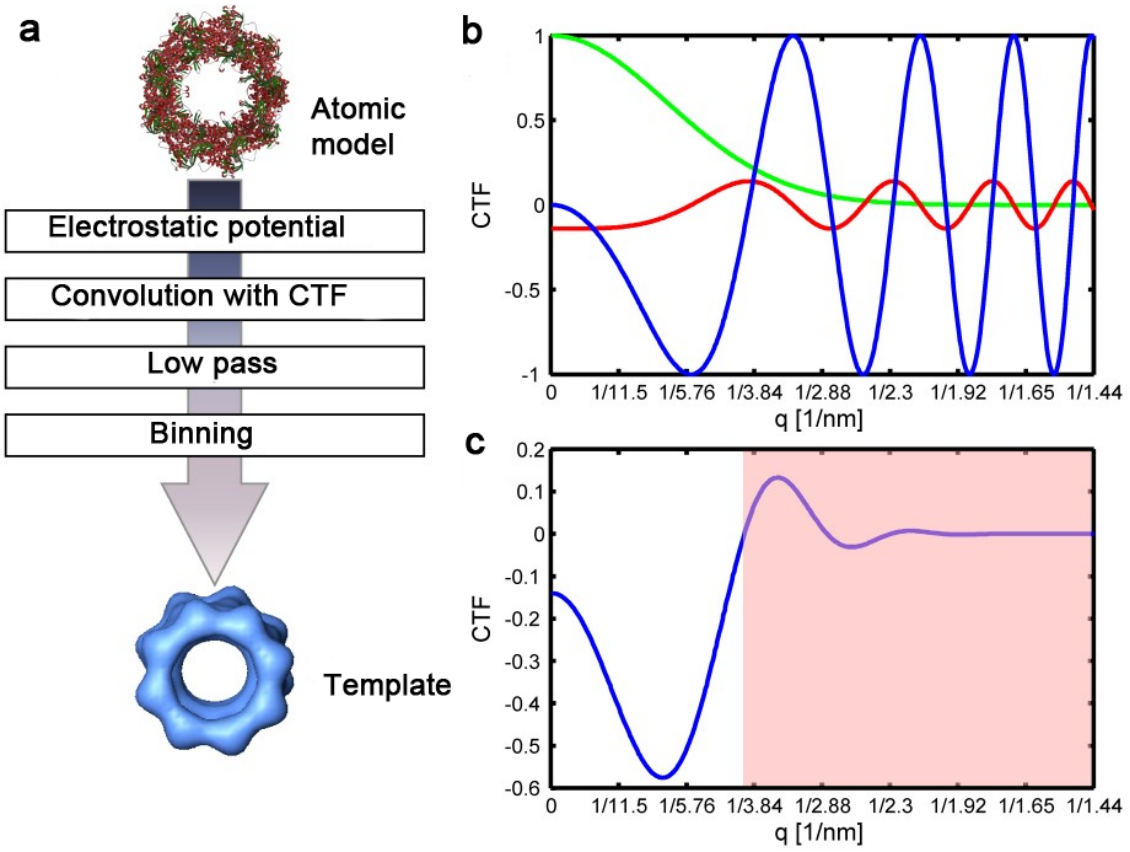

Figure 4.2: Template generation. (a) Template generation workflow, given the atomic model. (b) The CTF function has to be applied to the electrostatic potential map. The CTF consists of two additive contributions, phase contrast (blue) and amplitude contrast (red), which are damped by the modulation transfer function (MTF) of the CCD camera (green). (c) The CTF should be low-pass filtered to a proper resolution to prevent overfitting. All the information in the red-colored region is eliminated. Adapted from [Förster, 2005].

4.2.2 Scoring Function Calculation

In principle, the scoring function is a similarity measure of the template T and the tomogram V. Mathematically, given a 3D volume $V(\boldsymbol{p}) = V(x, y, z)$, we can define the rotation operator $\Lambda_{\boldsymbol{R}}$ as: $\Lambda_{\boldsymbol{R}} V(\boldsymbol{p}) = V(\boldsymbol{R}^T \boldsymbol{p})$ (assuming the origin is at the center). $\boldsymbol{R}$ is an Euler angle and can be parametrized with, e.g., ZYZ convention: $\boldsymbol{R}(\alpha, \beta, \gamma) = \boldsymbol{R}_Z(\alpha)\boldsymbol{R}_Y(\beta)\boldsymbol{R}_Z(\gamma)$. Given a similarity measure Sim, the scoring function can be calculated as:

$$Score = \max_{\boldsymbol{R}} Sim(\Lambda_{\boldsymbol{R}} T, V(\boldsymbol{p})). \tag{4.2}$$

In other words, for every possible position in the target volume V, the best orientation $\boldsymbol{R}$ is determined, such that the similarity Sim is maximized. The final score at position $\boldsymbol{p}$ is then the maximal value of Sim over all sampled rotations. In CET, correlation-based similarity measures are widely used. In the following, different variants of similarity measure are discussed.

Cross-correlation function. For discrete functions, the standard normalized cross-correlation function is defined as:

$$CC(\boldsymbol{\tau}) = \frac{1}{N} \sum_{\boldsymbol{p}} \frac{(T(\boldsymbol{p}) - \overline{T})(V(\boldsymbol{p}+\boldsymbol{\tau}) - \overline{V})}{\sigma_T \sigma_V}, \tag{4.3}$$

where T is the search template and V is the target volume. Both $T(\boldsymbol{p})$ and $V(\boldsymbol{p})$ are two discrete functions defined over N points. $\overline{T}$, $\overline{V}$ are the means, and σ_T, σ_V are the standard deviations. The cross-correlation function CC describes the correlation between T and V for every relative displacement $\boldsymbol{\tau} = (\tau_x, \tau_y, \tau_z)$. Both volumes T and V are firstly normalized and correlated. Equation 4.3 can be efficiently calculated by using correlation theorem and the fast Fourier transform:

$$CC = \mathcal{F}^{-1}(\mathcal{F}^{*}(T)\mathcal{F}(V)), \tag{4.4}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse Fourier transformations, respectively, and $*$ denotes the complex conjugate.

Local cross-correlation function. When the template T has a smaller size than the target V, or the region of interest occurs over a small region. A mask function $M(\boldsymbol{p})$ is introduced (specified by the user) that explicitly defines the extent of the area of interest. This leads to the local cross-correlation function [Roseman, 2003]:

$$LCC(\boldsymbol{\tau}) = \frac{1}{n} \sum_{\boldsymbol{p}} \frac{(T(\boldsymbol{p}) - \overline{T}) \cdot M(\boldsymbol{p}) \cdot (V(\boldsymbol{\tau}+\boldsymbol{p}) - \overline{V})}{\sigma_T \sigma_{MV}(\boldsymbol{p})}, \tag{4.5}$$

where $n = \sum_{\boldsymbol{p}} M$, and $\sigma_{MV}(\boldsymbol{p})$ is the local standard deviation of V under M. This is equivalent to calculating the normalized correlation coefficient between the search object and the masked subregion of the target for every relative position of the mask. The local correlation function has the advantage that nomalization of correlation only occurs locally, rather than the whole area. It should also be noted that the mask M is not necessarily binary, i.e., masks may also have smooth boundaries and pixels contribute with weights smaller than 1 to the correlation.

Local normalized cross-correlation function. To address the incomplete angular sampling of cryo-tomograms ("missing wedge" problem), the local normalized cross-correlation (LNCC) is proposed [Frangakis et al., 2002] to improve the similarity measure. The basic idea is to constrain the correlation calculation in the sampled region in Fourier space. Because the orientation of the searched particle is

unknown, the "missing wedge" effect has to be applied to the template T for every rotation before calculating the correlation. Mathematically, LNCC is calculated as:

$$LNCC(\boldsymbol{\tau}) = \frac{\sum_{\boldsymbol{p}} V(\boldsymbol{\tau}+\boldsymbol{p}) \cdot (T(\boldsymbol{p}) \otimes PSF_{\boldsymbol{R}}(\boldsymbol{p})) \cdot M(\boldsymbol{p})}{\sqrt{\sum_{\boldsymbol{p}}((V-\overline{V}) \cdot M)^2} \cdot \sqrt{\sum_{\boldsymbol{p}}((T-\overline{T}) \otimes PSF_{\boldsymbol{R}} \cdot M)^2}}. \tag{4.6}$$

Here, $PSF_{\boldsymbol{R}}$ is a point spread function that causes the same "missing wedge" effect and it is a function of the orientation $\boldsymbol{R}$ of the template. Assuming the sampled region in Fourier space is ω, $PSF_{\boldsymbol{R}}$ can be calculated as: $PSF_{\boldsymbol{R}} = \mathcal{F}^{-1}(\Lambda_{\boldsymbol{R}}\omega)$. $\otimes$ is the convolution operator. Obviously, LNCC is equivalent to LCC when the whole Fourier space is sampled. LNCC can also be calculated efficiently using fast Fourier transform.

4.2.3 Localization

After the calculation of the scoring function (Equation 4.2) with a proper similarity measure, the result needs to be further processed and interpreted in order to localize the macromolecules of interest. Because the obtained scoring function is a quantitative measure of the similarity of the template and the target volume, the peaks of the scoring function indicate possible occurrences of the template. The higher the coefficient is, the more likely the object of interest is located at the corresponding position. Therefore, a predefined number of peaks can be extracted from the scoring function in descending order. Note that after each peak is extracted, the area within the approximate radius of the template has to be excluded from the further analysis to prevent multiple counting. In the end, each extracted peak indicates a candidate with corresponding position and orientation information.

Due to the low SNR of the tomogram, some peaks will correspond to false positive matches. Another problem is how to determine the number of extracted peaks, because the number of macromolecule of interest inside a tomogram is essentially unknown. To address these issues, Ortiz et al proposed an empirical approach [Ortiz et al., 2006]: 1. Plot the histogram of the scores of the peaks. 2. Choose a threshold and fit all the scores above this threshold to a Gaussian distribution. 3. Use the determined mean μ and standard deviation σ to estimate the detection fidelity. This approach is under the assumption that the scores from the positives form a Gaussian distribution that can be relatively easily separated from the one of negatives. It works well for the macromolecules with high contrast, e.g., ribosomes [Ortiz et al., 2006]. However, for other macromolecules the true positive distribution may have more overlap with the negative distribution making the estimation of the number of false positives based on the histogram difficult.

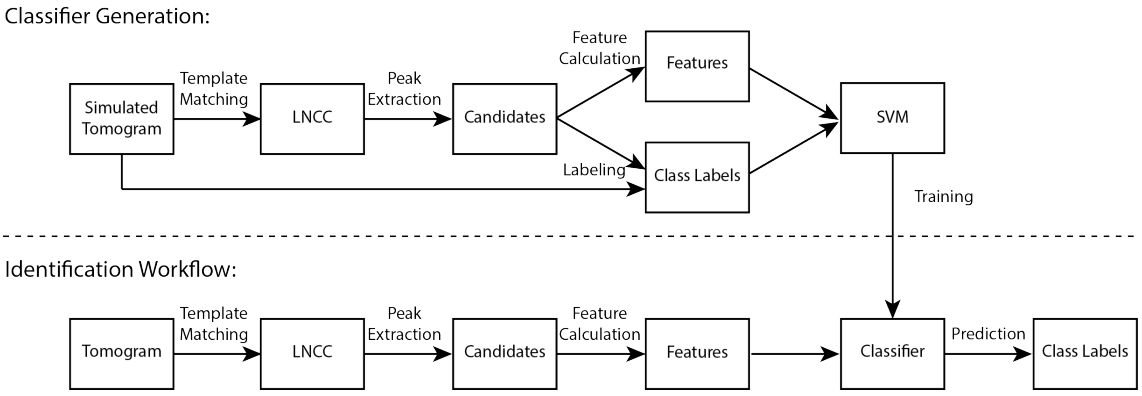

Figure 4.3: Classifier generation and identification workflow. In particular, the identification workflow consists of three steps: 1. Template matching and peak extraction. 2. Feature calculation. 3. Class prediction using the trained classifier, which is obtained from the classifier generation (the part illustrated above the dashed line).

4.3 Identification Workflow

The identification problem is first defined mathematically here. Given an input 3D tomogram V, the objective of identification of a target macromolecular complex in V is to find a set of subtomograms $\{v_1, \ldots, v_n\}$ containing the copies of the molecule, and their corresponding positions $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\}$ (center of mass) and orientations $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n\}$.

Here, a 3-step supervised learning approach is proposed for this problem: (i) candidate generation using template matching, (ii) feature calculation and (iii) final decision using a support vector machine (SVM) [Cortes and Vapnik, 1995]. SVMs are supervised learning models and they belong to the family of generalized linear classifiers, i.e., a SVM tries to find an optimal hyperplane to separate the samples of different classes. The hyperplane is determined by maximizing the margin and minimizing the structural risk. SVMs are well suited for the task here because: 1. SVMs have good out-of-sample generalization abilities. 2. SVMs are capable of handling high-dimensional features and non-linearly separable problems with the help of the soft margin and kernel techniques. 3. The theory of SVMs is well established and there are various efficient implementations. The overall workflow is depicted in Figure 4.3. The details of step (i) and (iii) will be discussed in the following. The feature calculation will be presented in section 4.5.

Candidate generation using template matching. This is a pre-detection step for obtaining the candidates. Given a tomogram V and the structural template T of the target macromolecular complex, the LNCC is calculated using Equation 4.6. Orientations are sampled explicitly with a pre-defined angular step. The candidates

$C = \{v'_1, \ldots, v'_{n'}\}$ can then be found by determining the local maxima of the LNCC (peak extraction), yielding their corresponding positions and orientations:

$$\boldsymbol{p}, \boldsymbol{R} = \arg\max_{\boldsymbol{p},\boldsymbol{R}} LNCC(V, T). \tag{4.7}$$

Final decision using SVM. In this step, a SVM classifier is used to discriminate the true and false positives in the candidates. The input of the classifier (test set $\mathcal{D}'$) consists of the rotation-invariant features (discussed in section 4.5) $\mathcal{D}' = \{\boldsymbol{x}_i\}, \boldsymbol{x}_i = SH(v'_i), i \in [1, n']$ of the candidates. The class labels of candidates will be predicted by the classifier trained below. Finally, all the candidates labelled as the positive class $\{v_1, \ldots, v_n\} \in C, n \leqslant n'$ are the final output.

Classifier training. SVM is a supervised learning algorithm. It has to be trained before it is used for predicting the class labels (Figure 4.3). The training can be conducted either on simulated or experimental tomograms. Given m candidates (training samples) and their features $\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$, the class labels $y = \{y_1, \ldots, y_m\}, y_i \in \{-1, 1\}, i \in [1, m]$ of all candidates are determined by their distances to the ground truth locations. If the candidate is closer than a threshold, it is labelled as positive, otherwise negative. Finally, the labels y and corresponding features $\boldsymbol{x}$ constitute the training set $\mathcal{D} = \{(\boldsymbol{x}_i, y_i)\}, i \in [1, m]$ of the SVM.

LIBSVM [Chang and Lin, 2011] is used as the implementation of SVM. Specifically, the RBF kernel is chosen for training and the best parameters C and γ are determined by a grid search. [1] Additionally, five-fold cross-validation is applied to avoid overfitting. To account for the unbalanced training set, different weights are assigned to the classes according to the quantity of the samples in each class [Chang and Lin, 2011]. Briefly, if the number of positive class in the training set is n_+ and the number of negative class is n_-, the weight of the positive class is set to $W_+ = n_-$ and the weight of the negative class is then $W_- = n_+$. Finally, the obtained classifier can then be used to predict the class labels of the incoming candidates from a new tomogram.

4.4 Spherical Harmonics

Before describing the feature calculation, the basics about spherical harmonics (SH) are introduced here. We firstly define the spherical function $f(\boldsymbol{\eta}) = f(\theta, \phi)$ on the unit sphere S^2, where $\boldsymbol{\eta}$ is a vector of $\theta \in [0, \pi]$, the colatitude (angle from the north pole) and $\phi \in [0, 2\pi)$, the longitude (angle around the z axis) (Figure. 4.4). A square-integrable function $f(\boldsymbol{\eta}) \in \mathcal{L}^2(S^2)$ can be expanded into a series of spherical

[1] Intuitively, C is a trade-off between the misclassification of training samples against the simplicity of the decision surface; γ defines the influence of a single training sample.

harmonics [Hobson, 1931]:

$$f(\boldsymbol{\eta}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \hat{f}_m^l Y_m^l(\boldsymbol{\eta}). \tag{4.8}$$

Here, Y_m^l are the spherical harmonic functions, and l and m are the degree and order, respectively. If the function f is band-limited or can be approximated as such, then we have:

$$f \approx \tilde{f} = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \hat{f}_m^l Y_m^l(\boldsymbol{\eta}), 0 \leq l < B, \tag{4.9}$$

where B is the bandwidth corresponding to the angular sampling of f. Namely, the number of sampling points of θ and ϕ should be at least $2B$, according to the sampling theorem.

The spherical harmonic function $Y_m^l(\boldsymbol{\eta}) : S^2 \to \mathbb{C}$ is defined as:

$$Y_m^l(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_m^l(\cos\theta) e^{im\phi}, \tag{4.10}$$

where $P_m^l(\cos\theta)$ denotes the associated Legendre polynomials. A set of spherical harmonics forms an orthogonal basis system on the unit sphere S^2.

The coefficients $\hat{f}_m^l$ can be calculated by:

$$\hat{f}_m^l = \int_{\boldsymbol{\eta} \in S^2} f(\boldsymbol{\eta}) \overline{Y_m^l(\boldsymbol{\eta})} \mathrm{d}\boldsymbol{\eta}. \tag{4.11}$$

Here, the overline denotes the complex conjugation. This procedure is called the Spherical Fourier Transform (SFT). Analogous to FFT, an algorithm exists for efficient computation of the SFT [Healy et al., 2003]. Its computational complexity is $O(B^2 \log^2 B)$.

4.5 Spherical-Harmonics-Based 3D Rotation-Invariant Features

After the retrieval of the candidates, their corresponding features are calculated. The features play a key role in the whole procedure because they ultimately determine the performance of the identification. These features should be not only noise robust, but also computationally efficient. Here, a set of rotation-invariant features based on spherical harmonics is proposed for each candidate. Rotation-invariant features are of special interest because they are fast to compute: an exhaustive angular search for the orientation is no longer necessary.

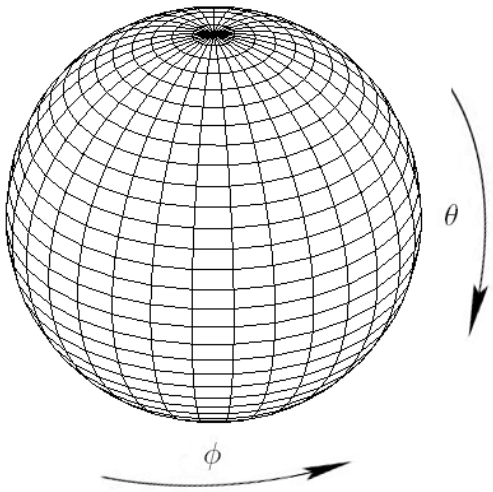

Figure 4.4: Parameterization on a sphere. $\theta \in (0, \pi)$ is the colatitude down from the Z-axis and $\phi \in [0, 2\pi)$ is the longitude counterclockwise off the X-axis.

The first step of calculating the SH features for a subtomogram v centered at the position of a candidate is to convert v to spherical coordinates:

$$v(x, y, z) = f(r, \theta, \phi). \tag{4.12}$$

For the spherical coordinate holds $r \in [0, R]$ (distance to the center of v), $\theta \in (0, \pi)$ (colatitude) and $\phi \in [0, 2\pi)$ (longitude) (see Figure 4.4). For a specific r, $f(r)$ is a spherical function. If $f(r)$ is band-limited to B or can approximated as such, $f(r)$ should be sampled at twice the bandwidth B at $\theta = \frac{(2i+1)\pi}{4B}, \phi = \frac{j\pi}{B}, i, j = 0, \ldots, 2B-1$ according to the Nyquist-Shannon sampling theorem, which yields $4B^2$ samples for each f. Here, the input is discrete 3D volumes in Cartesian space. In order to compute the features defined on spherical functions, the spline interpolation is used here for calculating the values of the points at spherical coordinates $r \cdot (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta) = (x, y, z)$.

One important question in practice is how to choose the bandwidth B. In principle, the higher B is, the more accurate the approximation $\tilde{f}$ will be. On the other hand, the expense is the rapidly increasing computation time. Moreover, it is beneficial to limit B to a low value when the input is noisy. Therefore, a trade-off has to be made for an appropriate setting of B. Mathematically, the surface area of a sphere with the radius r is $4\pi r^2$. According to the sampling theorem, the number of the sampling points $4B^2$ can be approximately determined as:

$$4B^2 \approx 2 \cdot 4\pi r^2 \Rightarrow B \approx \sqrt{2\pi} \cdot r. \tag{4.13}$$

The coefficients $\hat{f}_m^l$ themselves are not rotation-invariant. However, rotating a spherical function will not change the L_2-norm of $\hat{f}_m^l$ on each frequency l, because

the spherical harmonics are orthogonal basis functions of the unit sphere. Mathematically, the following rotation-invariant features can be constructed for a spherical function $f(r)$ [Kazhdan et al., 2003]:

$$SH(f(r)) = (\parallel \hat{f}_0^0(r) \parallel, \sum_{m=-1}^{1} \parallel \hat{f}_m^1(r) \parallel, \ldots, \sum_{m=-(B-1)}^{B-1} \parallel \hat{f}_m^{B-1}(r) \parallel). \tag{4.14}$$

Here, the sums of $\parallel \hat{f}_m^l(r) \parallel$ within each frequency l is used, because they are obviously also rotation-invariant; they provide a more compact feature representation and are more robust to the noise. Compared to [Kazhdan et al., 2003], the energy information $\sum_m \parallel \hat{f}_m^l(r) \parallel$ is used, instead of $\parallel \sum_m \hat{f}_m^l(r) \parallel$.

Equation 4.14 is calculated on a spherical function $f(r, \theta, \phi)$. When it is applied to a 3D data v, the rotation-invariant features are an assembly of $SH(f(r))$ of different radii from the volume center: $SH(v) = \{SH(f(1)), \ldots, SH(f(R))\}$. There are a few notes about these features: (i) These features are based on amplitude information of $\hat{f}_m^l$. The compressed information results in ambiguity of the features. In detail, the ambiguity is manifested in two aspects. First, rotating each spherical function $f(r)$ differently will yield the same feature representation. Second, for two spherical functions f and g, it is not sufficient to conclude there exists a rotation $\boldsymbol{R}$ so that $\Lambda_{\boldsymbol{R}} f = g$ given $SH(f) = SH(g)$, because the phase information is discarded (see Appendix A). (ii) For CET, strictly speaking, these features will differ to some extent for subtomograms depicting the same macromolecules in different orientations because the unsampled regions in Fourier space depend on the orientation of the macromolecules. This will have influence of all the spherical functions defined in real space. Nevertheless, as we shall see below the variations are sufficiently small to allow for discrimination of different macromolecules (Figure 4.5).

4.6 Dataset Preparation

- **Simulated Tomogram.** Ten tomograms ($512 \times 512 \times 512$ voxels) were simulated as described in section 3.5, each of which contained 5 different types of abundant objects (30 copies of each): 80S ribosome (PDB ID: 3IZS, 3IZF, 3IZB and 3IZE), 60S ribosome (PDB ID: 3IZB and 3IZE), 20S proteasome (PDB ID: 1PMA), GroEL (PDB ID: 1SS8) and gold beads of different sizes. All of the tomograms were simulated with a defocus value of 4 µm and a pixel size of 0.47 nm. The resulting tomograms were finally binned twice (pixel size 1.88 nm) to be consistent with the typical processing of experimental tomograms.

  Here, the focus is on the identification of 80S ribosomes (positive class). After template matching (12° as the angular increment for computing the LNCC)

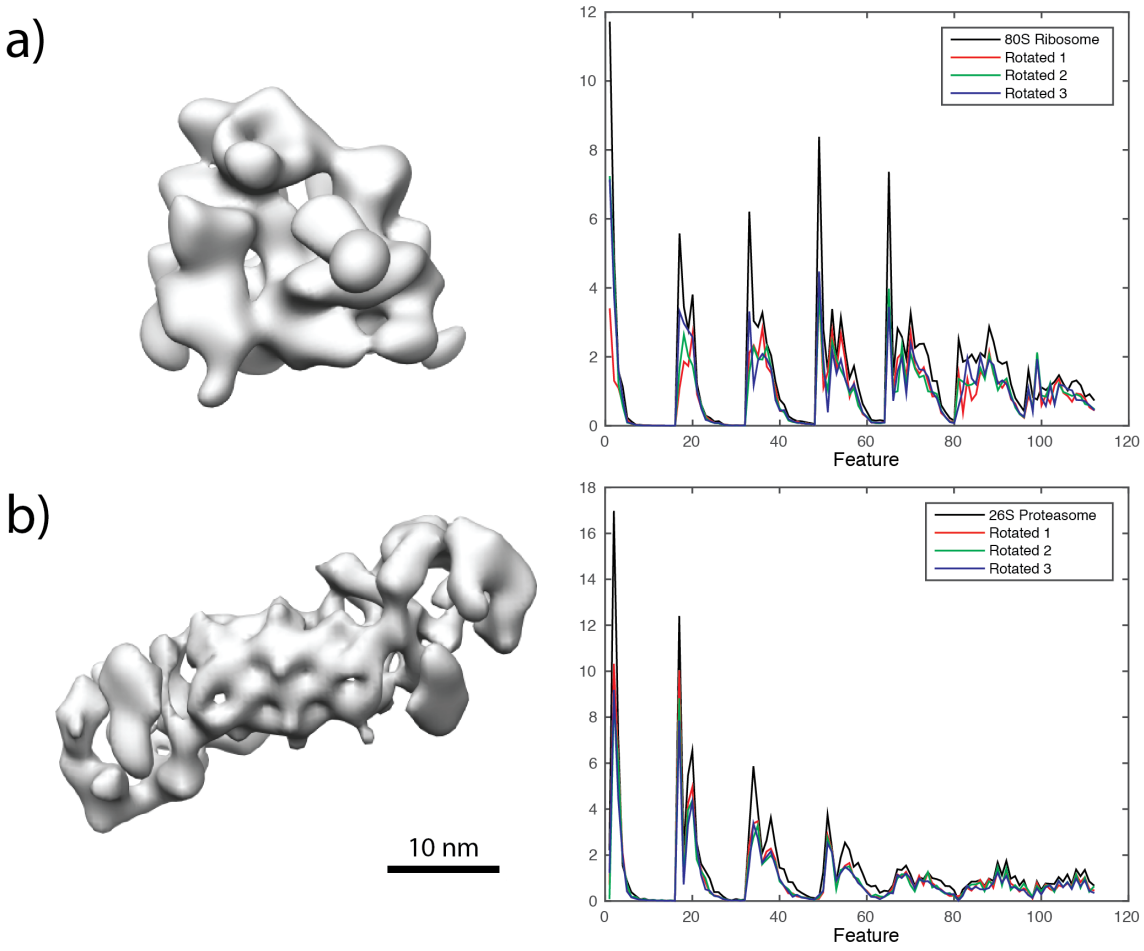

Figure 4.5: The influence of the missing wedge on the features. (a) The features were first calculated on a 80S ribosome map. Then it was randomly rotated three times and the missing wedge effect (tilt range: $-60°$ to $+60°$) was applied to each rotated map. Finally, the features were computed the same way. Left: a 80S ribosome density map. Right: the calculated features. (b) The same protocol as above applied on a 26S proteasome map.

1000 peaks (more than three times the amount of 80S ribosomes) were extracted to ensure a high coverage of the positive class. The class labels of the candidates were then determined. In this case, 258 candidates were labelled as the positive class (86% coverage) and the remaining 742 as the negative class.

The SH features were computed for all candidates. The radii for decomposing the subtomograms were chosen according to the size of the object of interest, 80S ribosome in this case. The size of 80S ribosome is approximately 25 nm and the pixel size is 1.88 nm, which leads to the maximal radius:

$25nm/(1.88 * 2nm/pix) \approx 7pix$. The bandwidth B of the spherical harmonics was determined according to Equation 4.13. Here, $B = \sqrt{2\pi} \cdot 7 \approx 17.5$. For the computational reason, B is chosen to be the closest integer of the power of 2, which is 16 in this case. Consequently, the dimension of the SH feature space was 112. The features and the class labels of the candidates formed the training set for the SVM. After the training, the obtained classifier was evaluated both on simulated and experimental tomograms.

- **Experimental Tomogram of ER-associated Ribosomes.** An experimental tomogram of endoplasmic reticulum (ER) microsomes derived from canine pancreas (Figure 4.7a) was used for the evaluation purpose. The tomogram (tilt range: $-60°$ to $+60°$, $3°$ angular increment) was acquired on a FEI Tecnai Polara TEM equipped with a Gatan GIF 2002 energy filter (300 kV acceleration voltage, 4 µm defocus, object pixel size 0.47 nm). After template matching with the 80S ribosome template ($12°$ as the angular increment for computing the LNCC), 500 peaks were extracted and these candidates were subjected to classification.

4.7 Results

4.7.1 Identification on Simulated Volumes

Ten additional tomograms were simulated as the test set. After template matching the SH features were calculated for each candidate. The class labels of the test set were then predicted by the classifier obtained in section 4.6. As a comparison, the state-of-the-art approach [Ortiz et al., 2006] was also evaluated (template matching with a single template followed by the thresholding, or TM in the following).

The results are shown in Table 4.1, which shows the SH approach performed vastly superior to the TM approach on the simulated data. Furthermore, the Receiver Operating Characteristic (ROC) curves of the classifiers are shown in Figure 4.6. SH clearly overcomes the TM approach by having a higher true positive rate while keeping the false positive rate lower.

4.7.2 Identification on An Experimental Volume

The performances of the proposed approach was further evaluated on the experimental tomogram (section 4.6). Using the classifier trained on simulated dataset (section 4.6), SH approach predicted 224 candidates as the positive class and 276 as the negatives. The classification result is plotted in Figure 4.7b, which shows most of the false positives, such as the gold beads or carbon films, were removed. For a fair comparison, the 224 particles with top scores from TM approach were also assigned to the positive class.

Approach	Accuracy	Precision	Recall
TM	61.4%	32.4%	33.1%
SH	**96%**	**91.8%**	**94.4%**

Table 4.1: Identification results on simulated tomograms. $Accuracy = (TP + TN)/(P + N)$, $Precision = TP/(TP + FP)$ and $Recall = TP/P$ with P: number of positives, N: number of negatives, TP: number of true positives, FP: number of false positives and TN: number of true negatives. For a fair comparison, the peak extraction threshold for the TM approach was set such that the amount of the positive class was the same as the one of the SH approach.

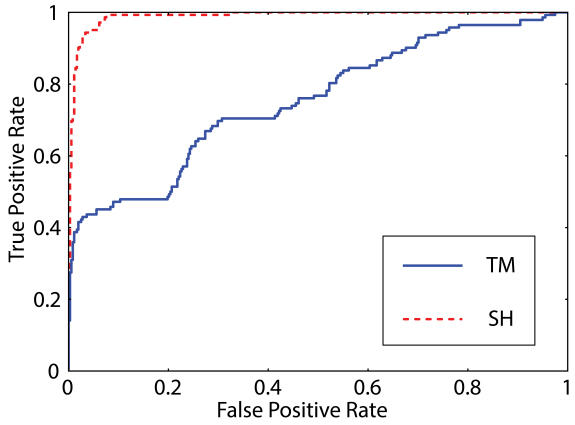

Figure 4.6: ROC curves of classifications for 80S ribosomes on simulated tomograms. The correlation threshold of peak extraction is varied for plotting the ROC curve of TM approach.

Due to the lack of ground truth, the performances were evaluated in two aspects. Firstly, the whole 500 candidates were manually labelled by the experts. Based on these subjective decisions, the results from SH and TM were evaluated in Table 4.2, which also indicate clear improvement of SH over the TM approach, even if the SVM was trained using the simulated data. Secondly, the resulting positives were evaluated based on their averages obtained by the subtomogram alignment and averaging (chapter 5). The average of positives from the SH approach exhibited ribosome-specific features and readily distinguishable ER-membranes, in contrast to the average from the TM approach, which was clearly affected by false positives with strong signals, probably gold beads and carbon films (Figure. 4.8a). This is further confirmed by the pairwise FSC curves and the cross-resolution (section 2.2.9) with a 18.3 Å resolution single particle map (EMDB: 1093) (Figure. 4.8b). Thus, the averages suggest improvements of detection accuracy by SH approach.

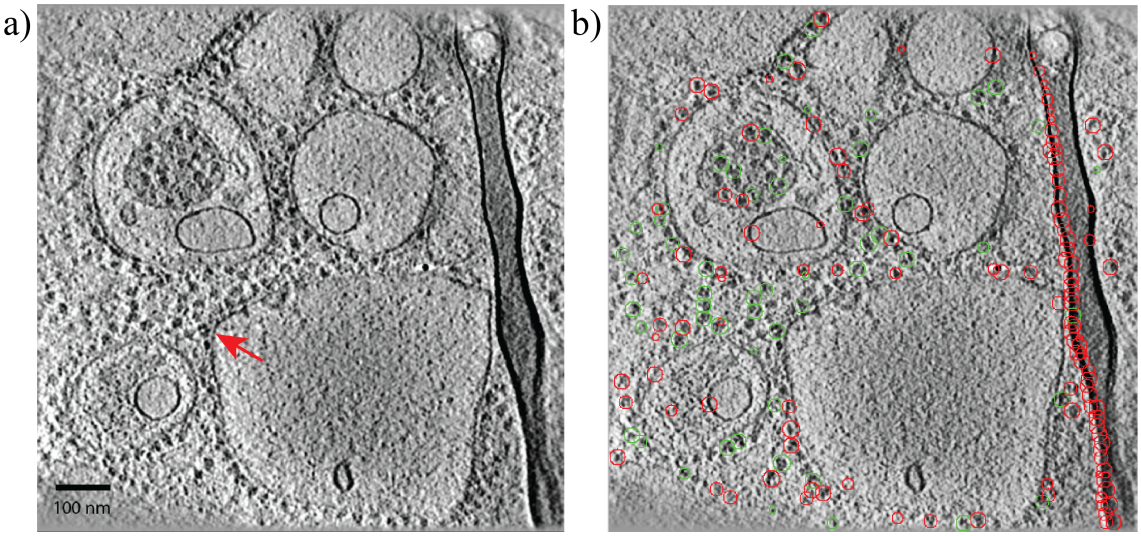

Figure 4.7: A slice view and identification result of an experimental tomogram. (a) A slice view of an experimental tomogram depicted ER-associated ribosomes. The arrow points to a potential ER-associated 80S ribosome. (b) Identification result of the SH approach. The green circles represent the positives, while the red ones are the negatives.

Approach	Accuracy	Precision	Recall
TM	57%	61.2%	51.7%
SH	**79%**	**85.7%**	**72.5%**

Table 4.2: Identification results on an experimental tomogram of ER-associated ribosomes based on the ground truth from manual labeling.

4.8 Discussion

A protocol is presented here to identify macromolecular complexes in cryo-electron tomograms using a 3-step supervised learning approach with rotation-invariant features. These features are calculated in spherical coordinate system and are based on spherical harmonics expansion. Due to the fact that the L_2-norm ("energy") of the spherical harmonic coefficients within the same frequency remain constant under any rotation, SH features are rotation-invariant. Moreover, SH features are noise robust because they characterize the object by its "energy" information. Finally, the existence of efficient algorithm for spherical Fourier transform makes the computation of SH features fast. Using SVMs, SH features associated to a specific complex are learned and trained. The resulting classifier is utilized to identify macromolecules of interest in a new tomogram.

The performance of the SH approach was first assessed on simulated tomograms,

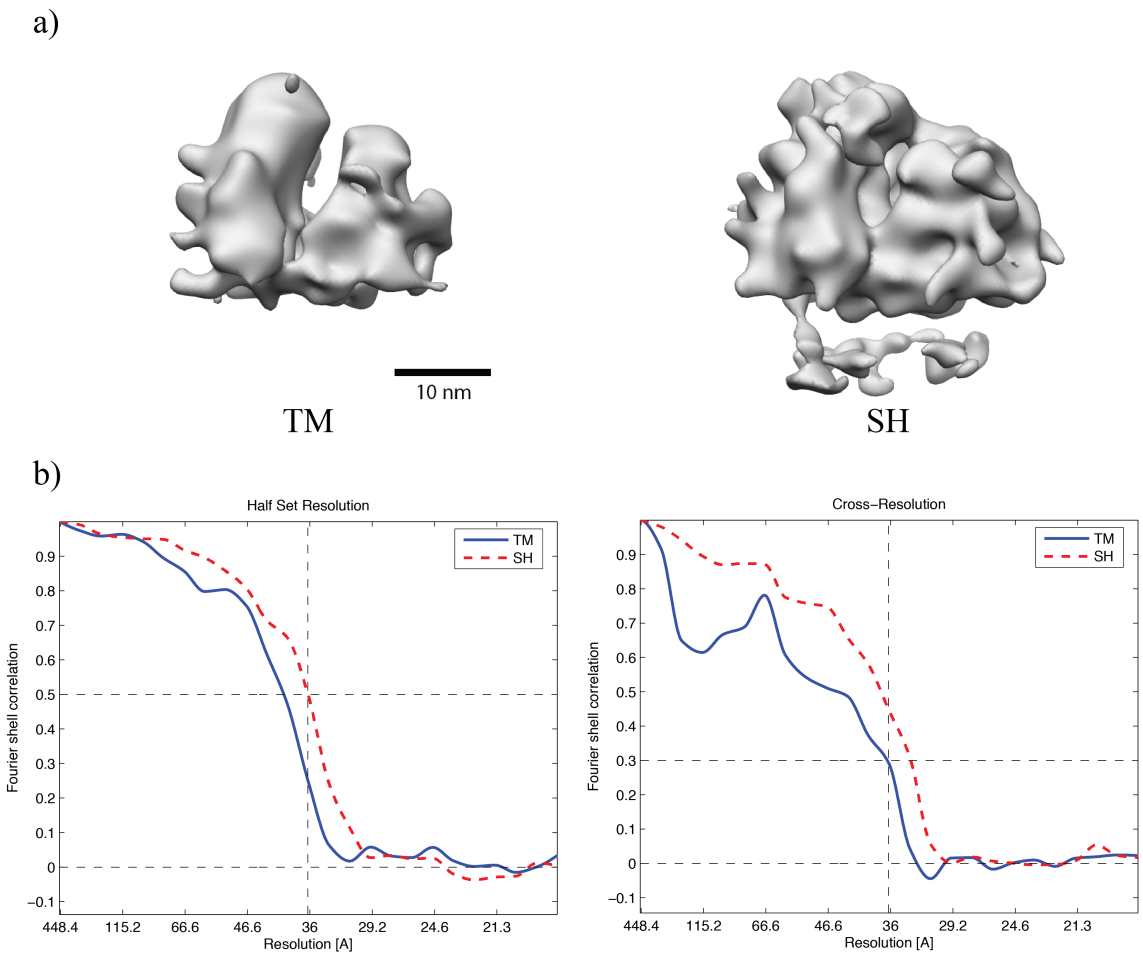

Figure 4.8: Comparison of subtomogram averaging between TM and SH approaches. (a) Subtomogram averages of TM approach (left) and SH approach (right). (b) Left: the pairwise FSC curves, right: cross-resolution FSC curves.

which showed significant advances of the accuracy, precision and recall rate compared to the state-of-the-art approach: template matching followed by thresholding. Furthermore, the SH approach was evaluated on an experimental tomogram of ER-associated ribosomes. It has been shown that SH approach was able to detect a vast majority of the positives while removing most of the false positives such as the gold nanoparticles or carbon films. The improvement was confirmed by the subtomogram averaging, which yielded a final structure of 80S ribosome with higher resolution and faithful ribosome-specific features. Because the classifier was trained on a simulated dataset, further improvements can be expected by improving the quality/quantity of simulation. It is worth mentioning that the classifier can also be trained using experimental dataset, if manual labelling is available.

On the other hand, SH approach presented here faces several challenges: (i)

Strictly speaking, the SH features introduced here are not rotation-invariant due to the "missing wedge" effect of the cryo-tomograms. Because the unsampled regions of each subtomogram in Fourier space are different, the underlying spherical functions are different even for the same macromolecular complex. Nevertheless, as confirmed by both simulation and experimental datasets, the features are robust enough to characterize correctly the different structures. The negative effect of this phenomenon could be reduced by constructing a large dataset with a proper angular coverage of the macromolecule, hence resulting in a stronger classifier. (ii) The preparation of the training set might be cumbersome and time-consuming. For different targets of macromolecules and different imaging condition of the TEM (such as the pixel size and defocus value), the simulation has to be conducted with corresponding configuration and the classifier has to be re-trained. It is a time-consuming task and cannot be generalized. Alternatively, if multiple tomograms with the same imaging condition are present, one can manually label one of the tomograms, train the classifier and use it to predict for the rest of the tomograms. Nevertheless, it is subjective and time-consuming for the operator. (iii) A bottleneck of this approach is that the sensitivity (recall rate) is limited, because the candidate list is still provided by template matching. This might become a problem for the detection of small macromolecules, on which the performance of template matching is normally inferior.

5

Subtomogram Alignment

5.1 Introduction

Once the macromolecule of interest is localized and identified in the tomogram, higher resolution insights can be obtained by aligning and averaging subtomograms, each depicting the macromolecule of interest [Bartesaghi and Subramaniam, 2009, Förster and Hegerl, 2007]. Commonly, thousands of subtomograms can yield resolutions of up to 15-20 Å [Briggs et al., 2009, Eibauer et al., 2012]. More recently, resolutions in the subnanometer regime have been obtained from more than 100,000 subtomograms [Schur et al., 2013]. Recent advances in automated data acquisition make it feasible to record thousands or more subtomograms [Korinek et al., 2011, Suloway et al., 2009, Zheng et al., 2010], which, in principle, should allow achieving a similar level of resolution as in single-particle cryo-electron microscopy reconstructions [Lander et al., 2012]. However, fast and accurate methods are needed for aligning and averaging such large datasets.

As mentioned in section 2.2.7, the expectation-maximization algorithm is a widespread approach for subtomogram alignment. While the translational match of a subtomogram and the reference can be searched efficiently, the rotational match needs explicit sampling of the whole rotational space. Its computational cost is enormous. As a consequence, the alignment procedures are typically started from initial references and rotational sampling is restricted to angles in the vicinity of rotations determined in previous alignment iterations, which may introduce "reference bias" into the resulting average [Subramaniam, 2006].

Here, a computationally efficient and accurate reference-free subtomogram alignment protocol is proposed that makes use of fast rotational matching. In contrast to previously introduced fast rotational matching approaches [Bartesaghi et al., 2008, Xu et al., 2012], the complete (amplitude and phase) information is used for scoring, which yields significantly more accurate alignment results for low SNRs. The alignment method is combined with a CTF correction of subtomograms using Wiener filtering.

This chapter is associated with a previous publication [Chen et al., 2013].

5.2 Correlation in Rotational Space

First, the mathematical basis of correlation in the rotational space using spherical harmonic functions is presented. Then the alignment problem is formulated in Fourier space and is finally addressed using fast translational and rotational matching.

5.2.1 *SO(3)* Fourier Transform

The basics about spherical harmonics and SFT have been explained in section 4.4. Here, the Fourier transform on the rotation group $SO(3)$ is introduced.

For a rotation $\boldsymbol{R}$ defined on the rotation group $SO(3)$, we can parameterize $\boldsymbol{R}$ with ZYZ convention (successive rotations around Z-, Y-, and Z-axis), namely: $\boldsymbol{R}(\alpha,\beta,\gamma) = \boldsymbol{R}_Z(\alpha)\boldsymbol{R}_Y(\beta)\boldsymbol{R}_Z(\gamma)$. For a function defined on the rotation group: $h(\boldsymbol{R}) \in \mathcal{L}^2(SO(3))$, we have [Kostelec, 2008]:

$$h(\boldsymbol{R}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=-l}^{l} \hat{h}^l_{mn} U^l_{mn}(\boldsymbol{R}), \tag{5.1}$$

$$U^l_{mn}(\boldsymbol{R}(\alpha,\beta,\gamma)) = e^{-im\alpha} P^l_{mn}(\cos\beta) e^{-in\gamma}.$$

Here P^l_{mn} are the generalized associated Legendre polynomials. The coefficients $\hat{h}^l_{mn}$ can be obtained by:

$$\hat{h}^l_{mn} = \int_{\boldsymbol{R}\in SO(3)} h(\boldsymbol{R})\overline{U^l_{mn}(\boldsymbol{R})}\mathrm{d}\boldsymbol{R}. \tag{5.2}$$

This is called the $SO(3)$ Fourier Transform (SOFT). If the bandwidth of h is limited to B, it has been shown that it can be efficiently calculated and its computational complexity is $O(B^3 \log^2 B)$ [Kostelec, 2008].

5.2.2 Cross-Correlation of Spherical Functions

With the two transforms (SFT and SOFT) described, the cross-correlation function of two spherical functions $f(\boldsymbol{\eta})$, $g(\boldsymbol{\eta})$ can be computed efficiently. We firstly define the spherical cross-correlation function as

$$SCC(\boldsymbol{R}) = \int_{\boldsymbol{\eta}\in S^2} f(\boldsymbol{\eta})\overline{\Lambda_{\boldsymbol{R}} g(\boldsymbol{\eta})}\mathrm{d}\boldsymbol{\eta}. \tag{5.3}$$

Maximizing $SCC(\boldsymbol{R})$ is equivalent to finding a rotation $\boldsymbol{R}$ to rotate g such that the two functions match best. It can be proven that (Appendix A) [Makadia and

Daniilidis, 2006]:

$$SOFT(SCC) = \hat{f}^l_m \overline{\hat{g}^l_p}. \tag{5.4}$$

In other words, the $SO(3)$ Fourier transform of $SCC(\boldsymbol{R})$ equals to the pointwise product of $SFT(f)$ and the complex conjugate of $SFT(g)$, which is the generalized convolution theorem of spherical functions.

Obviously, the computational complexity of computing $SCC(\boldsymbol{R})$ is the same as that of SOFT. The obtainable angular precision is $(\frac{360}{2B})^\circ$ for α,γ and $(\frac{180}{2B})^\circ$ for β (without considering possible gains by interpolation).

There is another more efficient algorithm [Kovacs and Wriggers, 2002] to compute Equation 5.3 using discrete Fourier transform rather than SOFT which can reduce the computation complexity to $O(B^3 \log B)$. By factorizing the rotation

$$\boldsymbol{R}(\alpha, \beta, \gamma) = \boldsymbol{R}_1(\alpha - \frac{\pi}{2}, \frac{\pi}{2}, 0)\boldsymbol{R}_2(\pi - \beta, \frac{\pi}{2}, \gamma - \frac{\pi}{2}), \tag{5.5}$$

the cross-correlation function can be rewritten as

$$SCC(\boldsymbol{R}) = \sum_{l=0}^{B-1} \sum_{m=-l}^{l} \sum_{p=-l}^{l} \sum_{k=-l}^{l} \hat{f}^l_m \overline{\hat{g}^l_p} P^l_{mk}(0) P^l_{kp}(0) e^{i(m\alpha' + k\beta' + p\gamma')}, \tag{5.6}$$

where $\alpha' = \alpha - \frac{\pi}{2}$, $\beta' = \pi - \beta$, $\gamma' = \gamma - \frac{\pi}{2}$. Therefore, applying a Fourier transform $\mathcal{F}$ on $SCC(\boldsymbol{R})$ results in:

$$\mathcal{F}(SCC(\boldsymbol{R})) = \sum_{l=0}^{B-1} \hat{f}^l_m \overline{\hat{g}^l_p} P^l_{mk}(0) P^l_{kp}(0). \tag{5.7}$$

That means, after an inverse 3D Fourier transform of the coefficients and a transfer of the angular parameters, the original spherical cross-correlation function can be retrieved. Using this approach, the computational complexity is essentially determined by that of FFT, which is $O(B^3 \log B)$. Therefore it is faster than the SOFT approach.

5.2.3 Constrained Cross-Correlation of Spherical Functions

The above approach to compute the correlation of two spherical functions assumes a complete sampling of θ and ϕ. If only parts of the functions are sampled, the correlated area has to be constrained and the correlation coefficients have to be normalized within the common area of the two correlated functions [Förster and Hegerl, 2007]. Two binary mask functions, m_f and m_g, are used to indicate the sampled areas of f and g, respectively. For convenience, the spherical correlation operator $\star$ is first introduced:

$$f \star g(\boldsymbol{R}) := SCC(\boldsymbol{R}). \tag{5.8}$$

The overlapping area of m_f and m_g is denoted by M, which depends on the rotation $\boldsymbol{R}$. Then the normalized spherical constrained cross-correlation function can be calculated as [Huhle et al., 2009]:

$$SCCC(\boldsymbol{R}) = \frac{\int_{S^2}[(f(\boldsymbol{\eta}) - \overline{f_M})m_f][\Lambda_{\boldsymbol{R}}(g(\boldsymbol{\eta}) - \overline{g_M})m_g]\mathrm{d}\boldsymbol{\eta}}{\sqrt{\int_M [f(\boldsymbol{\eta}) - \overline{f_M}]^2\mathrm{d}\boldsymbol{\eta}} \cdot \sqrt{\int_M [g(\boldsymbol{\eta}) - \overline{g_M}]^2\mathrm{d}\boldsymbol{\eta}}}. \tag{5.9}$$

Here, the overline represents the mean value of the respective function. It can be proven that overall six spherical correlation operations are needed for its calculation. Expansion of the numerator of Equation 5.9 yields:

$$\begin{aligned}\int_{S^2}[(f(\boldsymbol{\eta}) - \overline{f_M})m_f][\Lambda_{\boldsymbol{R}}(g(\boldsymbol{\eta}) - \overline{g_M})m_g]\mathrm{d}\boldsymbol{\eta} \\ = fm_f \star gm_g - \overline{f_M}(m_f \star gm_g) - \overline{g_M}(fm_f \star m_g) + \overline{f_M}\,\overline{g_M}(m_f \star m_g).\end{aligned}$$

The first factor in the denominator of Equation 5.9 can be calculated by:

$$\sqrt{\int_M [f(\boldsymbol{\eta}) - \overline{f_M}]^2\mathrm{d}\boldsymbol{\eta}} = \sqrt{f^2 m_f \star m_g + \overline{f_M}^2(m_f \star m_g) - 2\overline{f_M}(fm_f \star m_g)}.$$

The mean of spherical function under the overlap area, $\overline{f_M}$, can be calculated as:

$$\overline{f_M} = \frac{fm_f \star m_g}{m_f \star m_g}.$$

The second factor in the denominator of Equation 5.9 can be determined analogously. Therefore, the computational complexity of the constrained correlation computed using spherical harmonics is the same as the aforementioned unconstrained correlation. The same equation applies for complex spherical functions.

5.3 Fast Volumetric Matching

Finding the best alignment parameters between two 3D subtomograms V_1 and V_2 can be mathematically described as:

$$\arg\max_{\boldsymbol{\tau},\boldsymbol{R}} CCC(V_1, T_{\boldsymbol{\tau}}\Lambda_{\boldsymbol{R}}V_2), \tag{5.10}$$

where $T_{\boldsymbol{\tau}}V(\boldsymbol{x}) := V(\boldsymbol{x} - \boldsymbol{\tau})$ is the translation operator on volumes, $\Lambda_{\boldsymbol{R}}$ is the rotation operator, and CCC is the constrained cross-correlation coefficient of two 3D volumes. Here, the translational and rotational searches are separated, which is in contrast to the previously pursued approach of simultaneously sampling rotation and translation [Förster et al., 2005, Hrabe et al., 2012], due to the fact that

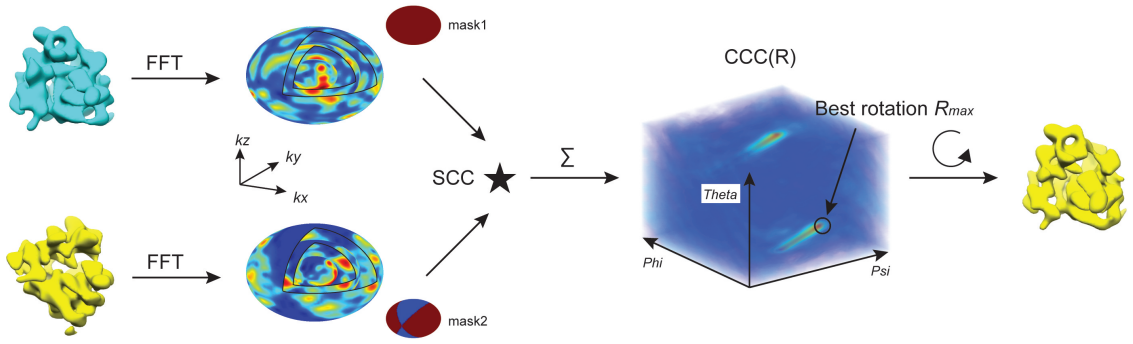

Figure 5.1: Workflow of FRM. The CCC of the Fourier transformed volumes (masks specify the respective missing wedges) as a function of the rotation of the yellow density can be computed using spherical cross-correlation (SCC).

simultaneous acceleration in terms of both, translation and rotation could not be mathematically formulated in this framework. This strategy of separate optimizations is also commonly done in the field of single particle analysis for essentially the same reason. In the following, how to find the best-scoring translation $\boldsymbol{\tau}_{max}$ is first described, and then the best-scoring rotation $\boldsymbol{R}_{max}$. Finally, those two separate optimizations are integrated into one procedure.

Assuming $\boldsymbol{R}_{max}$ is known, the translation $\boldsymbol{\tau}_{max} = \arg\max CCC(V_1, \Lambda_{\boldsymbol{R}_{max}} V_2)$ can be efficiently determined by fast translational matching (FTM): The subtomograms are first constrained in Fourier space according to the common areas [Frangakis et al., 2002]. Then the local correlation of V_1 and V_2 can be computed using the FFT-based formulation [Roseman, 2003].

In order to apply Equation 5.9 for rotational matching, the SCCC has to be extended to 3D volume data and the missing wedge problem has to be addressed. To achieve the latter the CCC is formulated in Fourier space, which allows incorporating the missing wedge as masks of spherical functions in Fourier space as shown in the following (Figure 5.1).

First an equivalent of Equation 5.9 is derived in Fourier space. We define $\hat{V}_1$ and $\hat{V}_2$ as the Fourier transforms of V_1 and V_2, and two mask functions m_1 and m_2, which indicate the respective missing wedges. The Fourier coefficients are converted to spherical coordinates: $\hat{V}(k_x, k_y, k_z) = \hat{V}(k, \theta, \phi)$. Because the mean value of a volume only affects $\hat{V}(0)$, we can set $\hat{V}(0) = 0$. The SCCC between $\hat{V}_1$ and $\hat{V}_2$ in frequency band k can be derived from Equation 5.9 by variable substitution and

setting all the mean terms to 0:

$$
\begin{aligned}
SCCC(\boldsymbol{R};k) &= \frac{SCC_{12}(\boldsymbol{R};k)}{\sqrt{SCC_{11}(\boldsymbol{R};k)}\cdot\sqrt{SCC_{22}(\boldsymbol{R};k)}}, k \geq 1, \\
SCC_{12}(\boldsymbol{R};k) &= \hat{V}_1(k,\theta,\phi)m_1 \star \hat{V}_2(k,\theta,\phi)m_2, \\
SCC_{11}(\boldsymbol{R};k) &= |\hat{V}_1(k,\theta,\phi)|^2 m_1 \star m_2, \\
SCC_{22}(\boldsymbol{R};k) &= m_1 \star |\hat{V}_2(k,\theta,\phi)|^2 m_2.
\end{aligned}
\tag{5.11}
$$

The CCC between $\hat{V}_1$ and $\hat{V}_2$ is then given by [Kovacs and Wriggers, 2002]:

$$
CCC(\boldsymbol{R}) = \frac{\sum_{k=1}^{k_{max}} SCC_{12}(\boldsymbol{R};k)\cdot k^2}{\sqrt{\sum_{k=1}^{k_{max}} SCC_{11}(\boldsymbol{R};k)\cdot k^2}\cdot\sqrt{\sum_{k=1}^{k_{max}} SCC_{22}(\boldsymbol{R};k)\cdot k^2}},
\tag{5.12}
$$

where k_{max} is the maximal frequency band during calculation. The peak of this function indicates the best-scoring rotation $\boldsymbol{R}_{max}$. Equation 5.12 describes how to compute the CCC by combining the $SCC(\boldsymbol{R};k)$ of each frequency k. Here, this method is called fast rotational matching (FRM). There are several notes about this function:

1. Its computation time depends on the chosen maximal Fourier frequency k_{max} and the maximal bandwidth B of the spherical harmonics. Overall, $3k_{max}$ spherical correlation operations are needed to compute the CCC. The larger k_{max} and B are, the more time the computation will consume, but larger k_{max} and B also yield higher accuracy. To avoid unnecessary computation cost, B is chosen according to the sampling theorem such that $B = \sqrt{2\pi}k$ (Equation 4.13) and k_{max} is determined according to the resolution of the reference.

2. If Equation 5.12 is used to calculate the correlations of subtomograms and a reference (e.g., V_1), m_1 and m_2 are typically fixed. Therefore all $SCC_{11}(\boldsymbol{R};k)$ only need to be calculated once, i.e., all the necessary number of spherical correlation operations can be reduced from $3k_{max}$ to $2k_{max}$.

3. Another feature of Equation 5.12 is that the calculation in each frequency can be handled separately. For example, by applying weights on different frequencies according to the FSC (section 2.2.9) the negative impact of noise can be alleviated [Rosenthal and Henderson, 2003, Stewart and Grigorieff, 2004].

Finally, an iterative procedure called Fast Volumetric Matching (FVM) is proposed to retrieve $\boldsymbol{\tau}_{max}$ and $\boldsymbol{R}_{max}$ when aligning V_2 to V_1. In FVM, the 6D optimization problem is approximated by successive solution of two 3D optimization problems ("divide and conquer") in an iterative manner. The overall workflow of FVM is depicted in Figure 5.2a and formally described in Algorithm 5.1.

Algorithm 5.1 Fast volumetric matching

Input: V_1, V_2

Output: $\boldsymbol{\tau}_{max}$, $\boldsymbol{R}_{max}$

1: Estimate an initial set of N candidate rotations $\{\boldsymbol{R}_0^1, \ldots, \boldsymbol{R}_0^N\}$. Because the power spectrum of a subtomogram is translation invariant, we can obtain the top N best rotational estimations between $|\hat{V}_1|$ and $|\hat{V}_2|$ using Equation 5.12 without any initial guess for the translations.
2: $i \leftarrow 0$ {current number of iteration}
3: I {maximal number of iterations allowed}
4: $Scores$ {the scoring function}
5: **for** $\boldsymbol{R}_i^j \in \{\boldsymbol{R}_i^1, \ldots, \boldsymbol{R}_i^N\}$ **do**
6: $\quad \boldsymbol{\tau}_i^j \leftarrow FTM(V_1, \Lambda_{\boldsymbol{R}_i^j} V_2)$ {use FTM to find the best translation parameter}
7: $\quad \boldsymbol{R}_{i+1}^j \leftarrow FRM(V_1, T_{\boldsymbol{\tau}_i^j} V_2)$ {use FRM to find the best rotation parameter}
8: $\quad Scores(\boldsymbol{\tau}_i^j, \boldsymbol{R}_i^j) \leftarrow CCC(V_1, T_{\boldsymbol{\tau}_i^j} \Lambda_{\boldsymbol{R}_i^j} V_2)$ {calculate the scores}
9: $\quad$ **if** $|\boldsymbol{\tau}_i^j - \boldsymbol{\tau}_{i-1}^j| \leq \varepsilon$ or $i + 1 \geq I$ **then**
10: $\quad\quad$ break
11: $\quad$ **end if**
12: $\quad i \leftarrow i + 1$
13: **end for**
14: **return** $\boldsymbol{\tau}_{max}, \boldsymbol{R}_{max} \leftarrow \arg\max Scores$

Here, the key differences between FVM and two state-of-the-art approaches, "Bartesaghi's approach" [Bartesaghi et al., 2008] and "Xu's approach" [Xu et al., 2012], are outlined:

1. In Bartesaghi's approach $|\hat{V}_1|$ and $|\hat{V}_2|$ are summed up along k to yield two real spherical functions and the best-scoring rotation is computed. This compression can cause significant errors at low SNRs. Mathematically, the following approximate scoring function is calculated:

$$CCC'(\boldsymbol{R}) = \frac{(\sum_k |\hat{V}_1|)m_1 \star (\sum_k |\hat{V}_2|)m_2}{\sqrt{(\sum_k |\hat{V}_1|)^2 m_1 \star m_2} \cdot \sqrt{m_1 \star (\sum_k |\hat{V}_2|)^2 m_2}}.$$

2. Xu's approach does not project 3D data to 2D. However, the rotation is optimized using $|\hat{V}_1|$ and $|\hat{V}_2|$ instead of $\hat{V}_1$ and $\hat{V}_2$. Thus, all the phase information is discarded, which is similar to the rotation function used in X-ray crystallography. Mathematically, it computes the following approximate scor-

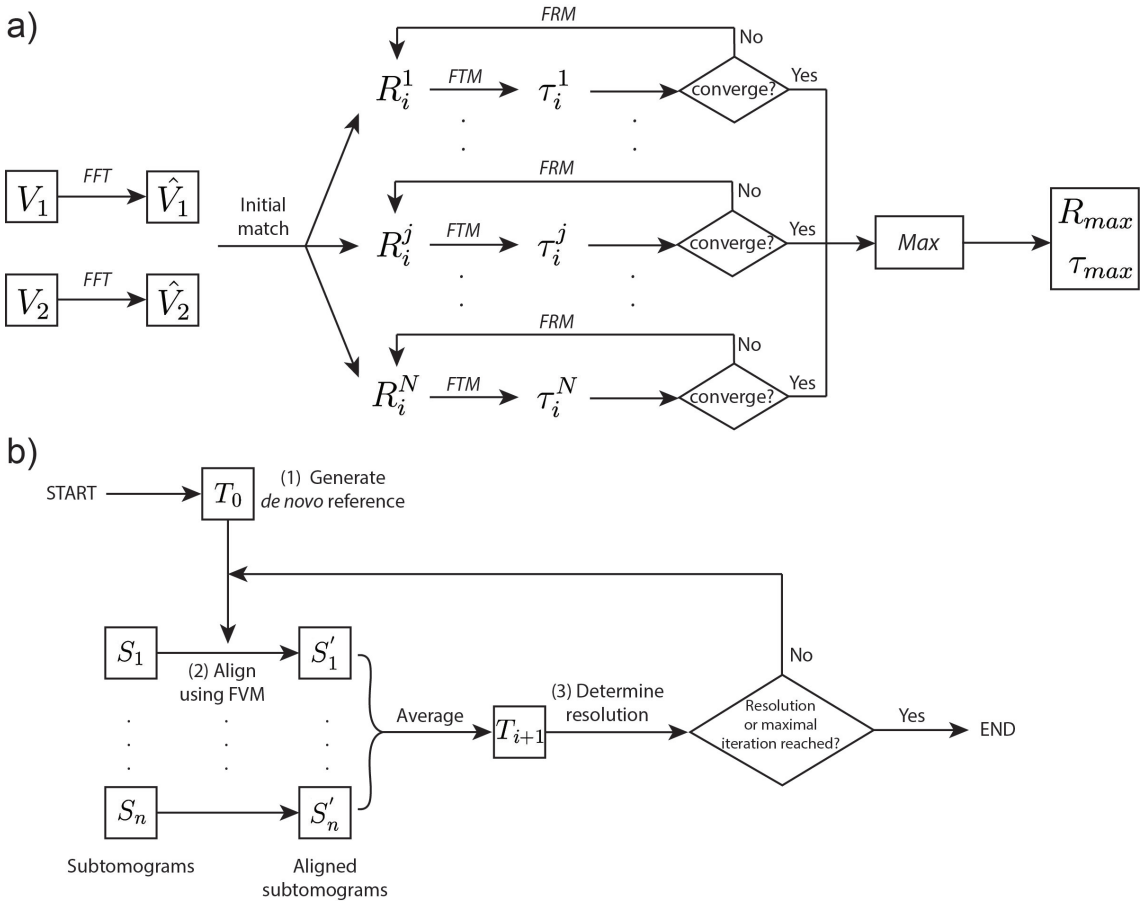

Figure 5.2: Fast volumetric matching and reference-free workflow. (a) FVM workflow. In the iterative procedure fast rotational matching (FRM) and fast translational matching (FTM) are used sequentially to obtain the best rotation $\boldsymbol{R}$ and translation $\boldsymbol{\tau}$ of V_2 compared to V_1. (b) Reference-free alignment workflow. In this quasi-expectation maximization algorithm the initial *de novo* reference T_0 algorithm is the average of the subtomograms with random orientations. The reference T is iteratively refined until the termination criterion is reached.

ing function:

$$CCC''(\boldsymbol{R}) = \frac{\sum_{k=1}^{k_{max}} SCC''_{12}(\boldsymbol{R};k)\cdot k^2}{\sqrt{\sum_{k=1}^{k_{max}} SCC_{11}(\boldsymbol{R};k)\cdot k^2}\cdot\sqrt{\sum_{k=1}^{k_{max}} SCC_{22}(\boldsymbol{R};k)\cdot k^2}},$$
$$SCC''_{12} = |\hat{V}_1|m_1 \star |\hat{V}_2|m_2.$$

3. FVM not only keeps the phase information, but also introduces an iterative procedure to obtain the best rotation and translation from the two separate 3D searches in rotation and translation space.

Algorithm 5.2 Reference-free alignment of subtomograms

Input: To be aligned subtomograms $\{S_1, \ldots, S_n\}$
Output: Aligned subtomograms $\{S'_1, \ldots, S'_n\}$

1: Generate a *de novo* reference T_0 by averaging a random portion of the subtomograms with random orientations.
2: Set the starting frequency k_0.
3: $i \leftarrow 0$ {current number of iteration}
4: I {maximal number of iterations allowed}
5: **while** $i < I$ and $k_i - k_{i-1} < 0$ **do**
6: **for** $S_j \in \{S_1, \ldots, S_n\}$ **do**
7: $S'_j \leftarrow$ use Algorithm 5.1 to align $lowpass(S_j, k_i)$ to $lowpass(T_i, k_i)$.
8: **end for**
9: $T_{i+1} \leftarrow$ average $\{S'_1, \ldots, S'_n\}$
10: $k_{i+1} \leftarrow$ calculate the FSC and determine the new resolution/frequency
11: $i \leftarrow i + 1$
12: **end while**

5.4 Reference-Free Alignment

The fast rotational matching procedure allows exhaustive rotational sampling in reasonable timescales, which is required for Reference-Free Alignment (FRA). The defining feature of RFA is that subtomograms are aligned without an external reference that might bias the alignment procedure. One possible RFA protocol is to use the average of the subtomograms in random orientations as an initial reference [Pfeffer et al., 2012, Scheres et al., 2009, Stölken et al., 2011]. Note that "reference-free" as used here does not mean no reference is used *per se*, rather any external reference is omitted. Thus, reference bias due to arbitrarily chosen starting models is alleviated. With conventional rotational matching exhaustive sampling is extremely time-consuming. The FRA method proposed here profits from the greatly accelerated rotational sampling and the high rotational sampling density. The workflow of the method is illustrated in Figure 5.2b and is described in detail in Algorithm 5.2:

The key points of this approach are: 1. The algorithm is entirely data-driven; an external reference is not required. 2. The iterative alignment algorithm is a quasi-expectation maximization algorithm, which ensures that it continuously improves the similarity score of a reference to the subtomograms and it converges rapidly (typically within ten iterations). 3. During each iteration k_{max} in Equation 5.12, is updated. Thus, the alignment evolves from low (reduce the noise bias) to higher (reveal fine details) frequencies and efficient computation is ensured concomitantly. 4. It cannot be excluded that the alignment might be locked in a local optimum if the SNR is very low or the orientation distribution of the subtomograms is strongly

nonuniform. In such cases, cross-validation methods, such as resolution determination using gold-standard FSC described below, may be beneficial.

5.5 Contrast Transfer Function Correction

A Wiener filter based approach was implemented in RFA procedure to compensate for the effects of the CTF, which allows aligning the subtomograms from tilt series with different defocus values analogous to common protocols in single particle analysis (SPA) (chapter 1) [Frank and Penczek, 1995]. In SPA, CTF correction is typically applied to class averages of many images with identical defocus rather than individual images. However, in CET projections contributing to a subtomogram have different defocus values due to the focus gradient in projections of the tilted specimen, which makes it essentially impossible to work with subtomograms of homogeneous defocus. Therefore, Wiener filtering is performed to each individual projection.

The Wiener filter of a projection t from T tilt angles for subtomogram i (of N) to reconstruct the CTF corrected density has the following form in Fourier space:

$$F_{it}(k) = \frac{CTF_{it}(k)}{\frac{1}{N \cdot T} \sum_{i=1}^{N} \sum_{t=1}^{T} |CTF_{it}(k)|^2 + \frac{1}{SNR}}. \tag{5.13}$$

The term $\frac{1}{SNR}$ is the Wiener filter constant, which is set to 0.1 as commonly done in SPA [Grigorieff, 1998]. In RFA implementation here, each projection was filtered using this filter prior to reconstructing the subtomograms, which were then used to compute the appropriately weighted average [Förster and Hegerl, 2007]. In the CTF, the damping envelope was not considered at this point, which would lead to enhancement of high frequency part in the corresponding filter and may increase noise enhancement. Moreover, the CTF was defined as constant for frequencies below the first maximum because phase contrast is typically not dominant at very low frequencies.

The defocus gradient was incorporated similarly as described in [Fernández et al., 2006, Zanetti et al., 2009]: the average defocus of the entire tilt series was determined based on the Thon rings of a periodogram average and the defocus value at different stripes parallel to the tilt axis was determined based on trigonometry. For each of these stripes the Wiener filtering was performed by adopting the phase-flipping procedure published in [Eibauer et al., 2012]. Overall, two differently filtered versions of each subtomogram were reconstructed: the CTF convoluted subtomogram was used for computing the Wiener-filtered average after each iteration and a phase-flipped subtomogram was used for subtomogram alignment.

5.6 Resolution Determination Based on Gold-Standard FSC

Interpretation and resolution determination of single particle reconstruction according to the gold-standard FSC rather than standard FSC was reported to be able to prevent overfitting and resulting noise enhancement in SPA [Grigorieff, 2000, Scheres and Chen, 2012]. The essence of this measure is that two sets of particles, each containing 50% of the whole dataset, are aligned and averaged independently rather than together, which avoids correlation due to noise. Here, this method was implemented for subtomogram alignment and averaging. The dataset was divided into two random sets and low-pass filtered to a certain resolution for the first iteration (in the ribosome dataset 5.2 nm and proteasome dataset 4 nm). Within each set, the alignment was carried out independently. After each iteration, the average of the second set was aligned to that of the first one and the FSC between the two averages was calculated (gold-standard FSC). The determined resolution was used as the low-pass filter for the next iteration in both sets.

5.7 Dataset Preparation

5.7.1 Simulated Subtomograms

Subtomograms of the *S. cerevisiae* translating 80S ribosome were simulated using atomic models from the Protein Data Bank (ID: 3IZB 3IZE 3IZF 3IZS) with the size of 100^3, a defocus value of 4 µm and a pixel size 0.47 nm. The subtomograms were randomly translated and rotated. Their projections were simulated as described in section 3.5 from $-60°$ to $+60°$ with an angular increment of $3°$ in various noise conditions. Finally, the subtomograms were reconstructed using weighted back-projection. For computation of cross-resolutions (section 2.2.9) of the obtained averages, a tomogram was generated by applying the 3D CTF to the original density map derived from the atomic coordinates. All the simulations were carried out using the TOM toolbox [Nickell et al., 2005].

5.7.2 Ribosome Subtomograms from Yeast Lysate

The sample preparation of *S. cerevisiae* lysate and the tomogram acquisition followed the protocol described in section 3.6. Downsampled tomograms (voxel size: 2.3 nm) were reconstructed by weighted-back projection (section 2.2.5) using the TOM toolbox [Nickell et al., 2005]. A ribosome template was generated (section 4.2.1) from atomic models (PDB ID: 3IZB 3IZE 3IZF 3IZS), convoluted with an appropriate CTF and low-pass filtered to 40 Å as described previously [Förster et al., 2010]. Template matching (section 4.2) was then carried out for each tomogram

to extract the ribosome candidates [Chen et al., 2012, Hrabe et al., 2012], which resulted in 8,000 subtomograms. For reconstruction of the unbinned subtomograms the phases were first corrected (section 2.2.4) [Eibauer et al., 2012] prior to weighted backprojection. The resulting subtomograms were then classified using CPCA (section 2.2.8) [Förster et al., 2008] to remove false positives (e.g., gold beads) from the dataset. Finally, 6,436 subtomograms were retained for further alignment. For these subtomograms, a second set of volumes, convoluted with the CTF, was reconstructed and used for the Wiener filter.

5.7.3 20S Proteasome Subtomograms

20S proteasomes were recombinantly expressed in *Escherichia coli* cells and purified as described in [Witt et al., 2006]. The sample was diluted to a concentration of approximately 0.5 mg/ml and CET grids were prepared as described in section 3.6. Tilt series were acquired with a defocus of 6 µm. The reconstructions and processing steps were performed analogous to the ones of the ribosomes, finally yielding 1,247 subtomograms for the alignment and averaging.

5.8 Results

5.8.1 Implementation

The algorithms proposed here were programmed in C for the sake of efficiency, and a high level interface is also provided in Python for the ease of usability. The whole software package was built based on two libraries: SpharmonicKit [Healy et al., 2003] for computing the SFT, and Situs [Kovacs and Wriggers, 2002] for computing the cross-correlation of spherical functions. Third order spline interpolation was used for transferring the voxel values from the Cartesian to spherical coordinates. For comparison, Bartesaghi's and Xu's approaches [Bartesaghi et al., 2008, Xu et al., 2012] were also implemented. Moreover, a seamless integration to PyTom [Hrabe et al., 2012] was provided for alignment procedures, which are parallelized by message passing interface (MPI) to exploit further speedup.

5.8.2 Speedup compared to Real Space Rotational Searches

The speedup of FVM was first evaluated and compared to rotational sampling in Cartesian space, as currently used in most subtomogram averaging packages using simulated data. 100 ribosome subtomograms were simulated (SNR=0.1) with random translations (in the range of 10 pixels) and rotations. The low-pass filtered original model (5.2 nm resolution) was used as the starting reference for FVM. After each iteration, the averages from different iterations were compared to the original

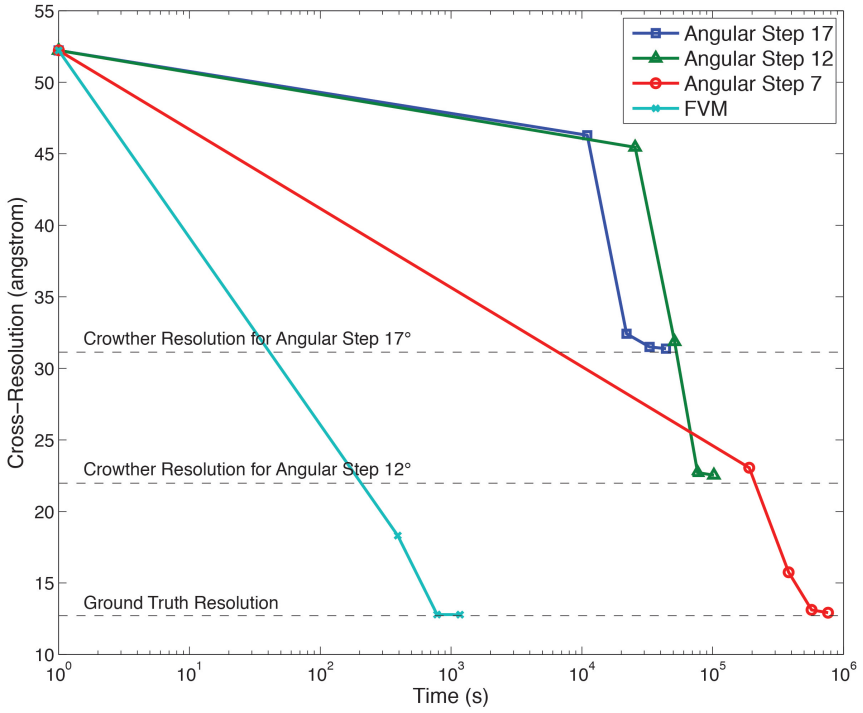

Figure 5.3: Speedup of FVM compared to real space rotational search. The computation time for refining the alignment of simulated subtomograms was evaluated for FVM and exhaustive rotational sampling with different angular intervals, as well as the cross-resolutions of the averages from alignment iterations with respect to the original density.

model by means of FSC to measure the cross-resolutions and recorded the elapsed time. Finally, the protocol was repeated using exhaustive real space rotational searches with different angular steps.

The results of this comparison indicate a large speedup of FVM compared to the real space rotational search (Figure 5.3). For the best possible resolution (12.7 Å), the speedup is approximately 1,000. The resolution of the average resulting from the real space search approximately corresponds to that indicated by the Crowther criterion [Crowther et al., 1970], whereas the resolution from the FVM approach matches the ground truth (average of subtomograms using their true translations and rotations). Thus, the simulations suggest that FVM accelerates subtomogram alignment by approximately three orders of magnitude at resolutions beyond 2 nm without notable loss of accuracy.

If the orientation of subtomograms were known approximately, only a small angular range would need to be searched (albeit the expectation maximization algorithm, strictly speaking, requires global sampling). In such a case, there is a limit when the direct angular search is faster than FVM, depending on several parameters, such

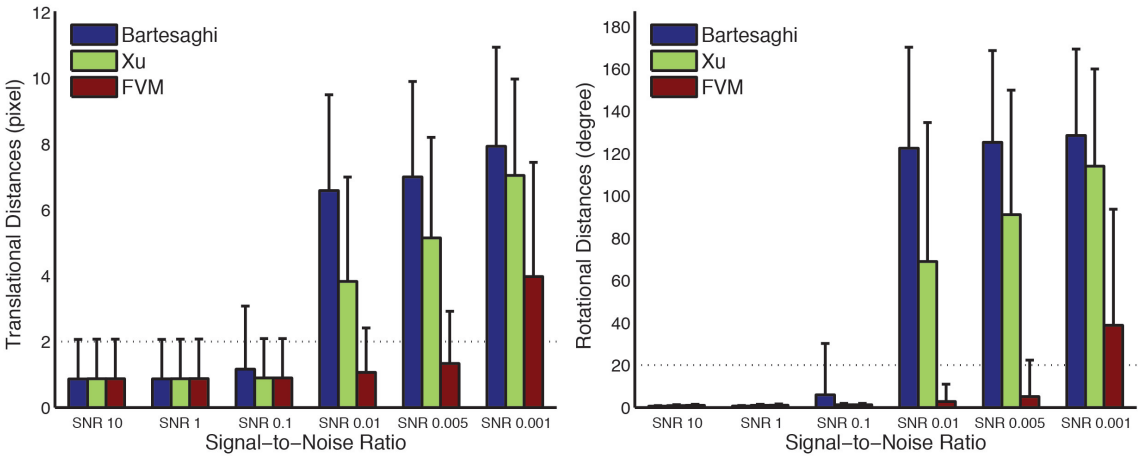

Figure 5.4: Alignment errors of different spherical-harmonics-based matching algorithms for different SNRs. The alignments below the dashed line is considered as correct.

as maximal frequency and the bandwidth of spherical harmonics. For example in this case, for $k_{max} = (2nm)^{-1}$, the elapsed time of FVM equals to the one of direct angular searches of approximately 108 angles. The comparison was run with the same data on a machine equipped with the Intel Xeon Processor X5570 2.93GHz (4 cores) and 48 GB main memory.

5.8.3 Comparison of Alignment Accuracy

Next, FVM was compared with other spherical-harmonics-based approaches: Bartesaghi's approach [Bartesaghi et al., 2008] and Xu's approach [Xu et al., 2012] in terms of accuracy. 1,000 ribosome subtomograms were simulated with different SNRs, which were randomly translated (in the range of 10 pixels) and rotated. The three different approaches were then used to determine the best translational and rotational parameters of the simulated subtomograms when matched against the original volume. The translational and rotational errors are shown in Figure 5.4. Here, an alignment is considered correct if the translational error is smaller than 2 pixels and the rotational error is smaller than 20°.

All three approaches resulted in correct alignment for high SNRs (SNR ≥ 0.1). When the SNR was below 0.01, Bartesaghi's and Xu's approaches yielded incorrect alignments. Consistent with [Xu et al., 2012], it is observed that Xu's approach achieves higher accuracy than Bartesaghi's approach, but the error of both approaches is larger than the tolerance. On the contrary, FVM yielded correct alignment even for SNR=0.005. In summary, the simulations show that FVM yields significantly higher alignment accuracy at typical SNRs of cryo-tomograms (SNR < 0.1).

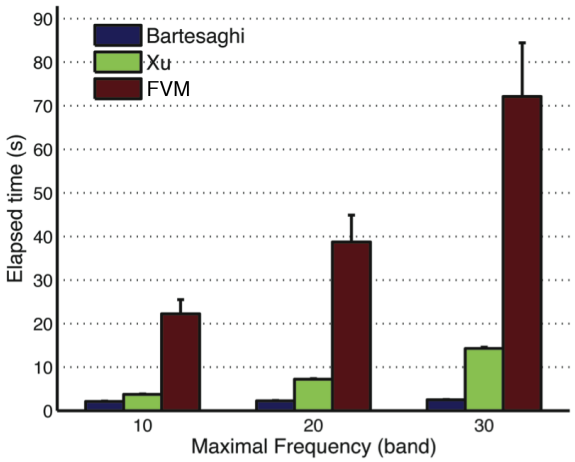

Figure 5.5: Comparison of computation time for subtomogram alignment of simulated 80S ribosomes using three spherical-harmonics-based alignment algorithms.

The computational speed was also compared using different spherical-harmonics-based alignment methods (Figure 5.5). This comparison reveals that Bartesaghi's and Xu's approaches surpassed FVM approximately by a factor of ten and five, respectively. Thus, the gain in alignment accuracy resulted in a reduction of computational speed.

5.8.4 Reference-Free Alignment on Simulated Dataset

To test the RFA, it was first applied to a simulated dataset with various SNRs (Figure 5.6). All resulting averages were analyzed by cross-resolution with the noise-free reference. For comparison, the subtomograms were averaged according to the ground truth translation and rotation parameters and the cross-resolution FSC curves were computed. The FSC curves suggest that the resolutions of the averages from RFA are close to the best possible values. Only for the lowest SNR tested (SNR=0.001) the FSC is worse than the ground truth FSC. This decreased performance is consistent with the accuracy assessment of FVM.

5.8.5 Reference-Free Alignment on Experimental Dataset

RFA was then applied to an experimental ribosome dataset from *Saccharomyces cerevisiae* cell lysate (Figure 5.7a). Throughout the iterative procedure the average evolves from a blob to a structure with canonical ribosomal features (Figure 5.8a). Because the ground truth was unavailable here, the results were evaluated in three aspects (section 2.2.9): the half-set FSC, the gold standard FSC and the cross-resolution FSC. According to the half-set FSC, the resolution (0.5 criterion) of the final average is 20.4 Å (Figure 5.9a). Using the gold standard FSC, i.e., the alignment

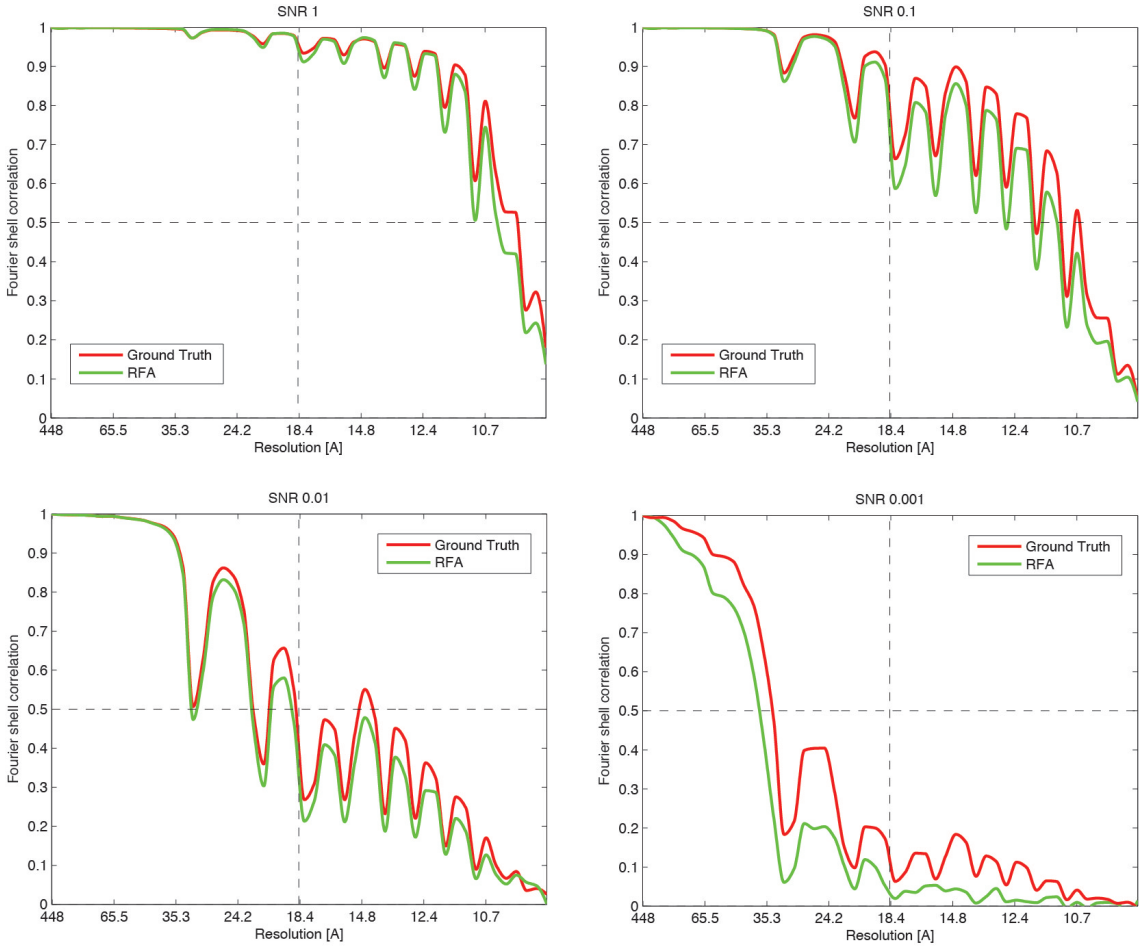

Figure 5.6: Accuracy of averages obtained from RFA. For simulations with SNR=1, 0.1, 0.01 and 0.001 the cross-resolutions of the averages and true CTF-convoluted model were computed. For comparison, the cross-resolutions of averages using the ground truth alignment are shown.

was conducted with the data split into two independent sets, the resulting density was essentially identical to that obtained by RFA. The corresponding resolution was even slightly better (18.9 Å) according to the 0.143 criterion that is recommended for the gold-standard FSC [Scheres and Chen, 2012]. Furthermore, the accuracy of the obtained averages was assessed by cross-resolution with a 7.9 Å resolution single particle reconstruction (EMDB: 1668), which yielded 19.4 Å resolution (0.3 criterion). Thus, the average obtained by RFA is faithful to at least 20 Å resolution.

In a second experimental test case RFA was applied to subtomograms from purified *Thermoplasma acidophilum* 20S proteasomes (Figure 5.7b), which are significantly smaller than the 80S ribosomes (700 kDa vs. 3.3 MDa). The result

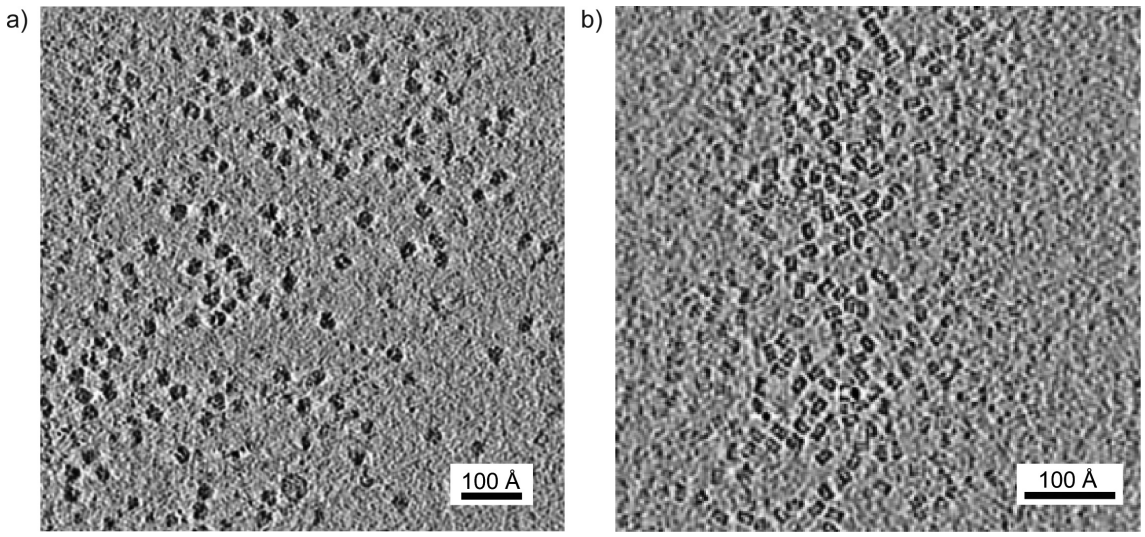

Figure 5.7: Slice view of experimental tomograms. (a) One slice of an experimental tomogram depicting *S. cerevisiae* ribosomes from yeast lysate. (b) One slice of an experimental tomogram depicting *Thermoplasma acidophilum* 20S proteasomes.

converges into a trustworthy average (Figure 5.8b). When the full D_{14} symmetry[1] is exploited, the resolution of the final average, according to different measures, is: 15.1 Å (half-set FSC 0.5), 14.7 Å (gold standard FSC 0.143), and 16.2 Å (cross-resolution) (Figure 5.9b).

5.9 Discussion

An algorithm (FVM) is presented here to find the best translational and rotational match of two subtomograms using CCC as the scoring function. For acceleration of rotational matching, the computation of the CCC is formulated using spherical harmonics in Fourier space. The CCC uses the full information (including phases) and constrains the correlation to Fourier space sectors that are sampled in both volumes. It is shown that FVM is approximately 3 orders of magnitude faster than methods using exhaustive rotation sampling in Cartesian space, as implemented in many packages for subtomogram averaging [Amat et al., 2010b, Castaño Díez et al., 2012, Förster et al., 2005, Hrabe et al., 2012, Schmid and Booth, 2008, Walz et al., 1997, Winkler et al., 2009]. Compared to previously published spherical-harmonics-based approaches [Bartesaghi et al., 2008, Xu et al., 2012], FVM is significantly more accurate because it uses the full information (including phases) and the translational match is iteratively refined. Thus, FVM provides significant computational speedup of subtomogram alignment by correlation methods essentially without loss

[1] 7-fold symmetry around one axis and 7 2-fold axes orthogonal to that one.

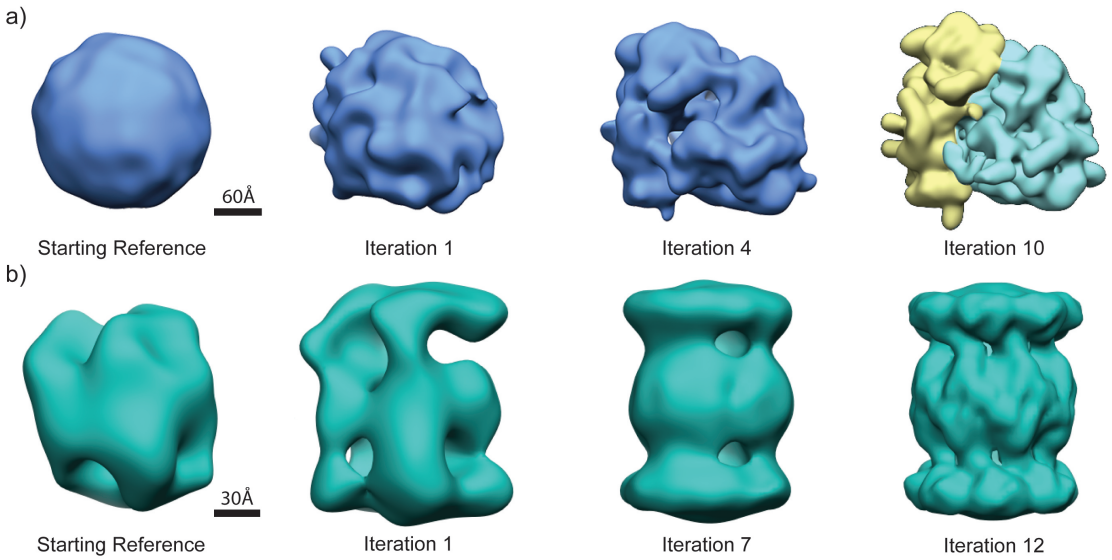

Figure 5.8: Evolution of subtomogram averages during alignment. (a) Evolution of subtomogram averages (80S ribosome from yeast lysate) from the starting reference to the final 20 Å structure. In the final average, the large ribosomal subunit is colored in blue and small subunit in yellow. (b) Evolution of subtomogram averages (20S proteasome) from the starting reference to the final 15 Å structure. The D_{14} symmetry is applied after iteration 7.

of accuracy. This method will therefore become the method of choice in the field of subtomogram averaging.

The FVM framework presented here can, in principle, be used for other scoring schemes. For example, application of different spectral weighting schemes as suggested in the single particle field [Stewart and Grigorieff, 2004] is straightforward. The proposed algorithm would also accelerate alignment and classification by the maximum likelihood approach [Scheres et al., 2009, Stölken et al., 2011] without loss of accuracy.

Based on FVM, a reference-free alignment procedure (RFA) was developed, which does not require external references. RFA is a quasi-expectation maximization procedure (Algorithm 5.2). The problem here is to find the maximum a posteriori estimates of the alignment parameters, given a set of observations (subtomograms). In detail, the current model in the E-step is calculated by the subtomogram averaging (Equation 2.16) and the parameter estimation in the M-step is provided by the proposed subtomogram alignment algorithm, i.e., FVM. The rest of RFA is basically the same as the expectation maximization, which is proven to converge [Wu, 1983]. Nevertheless, a local optimum might be reached. The RFA starts from an average of the subtomograms in random orientations. The reference and its determined resolution are updated during each iteration, which ensures efficient computation of the

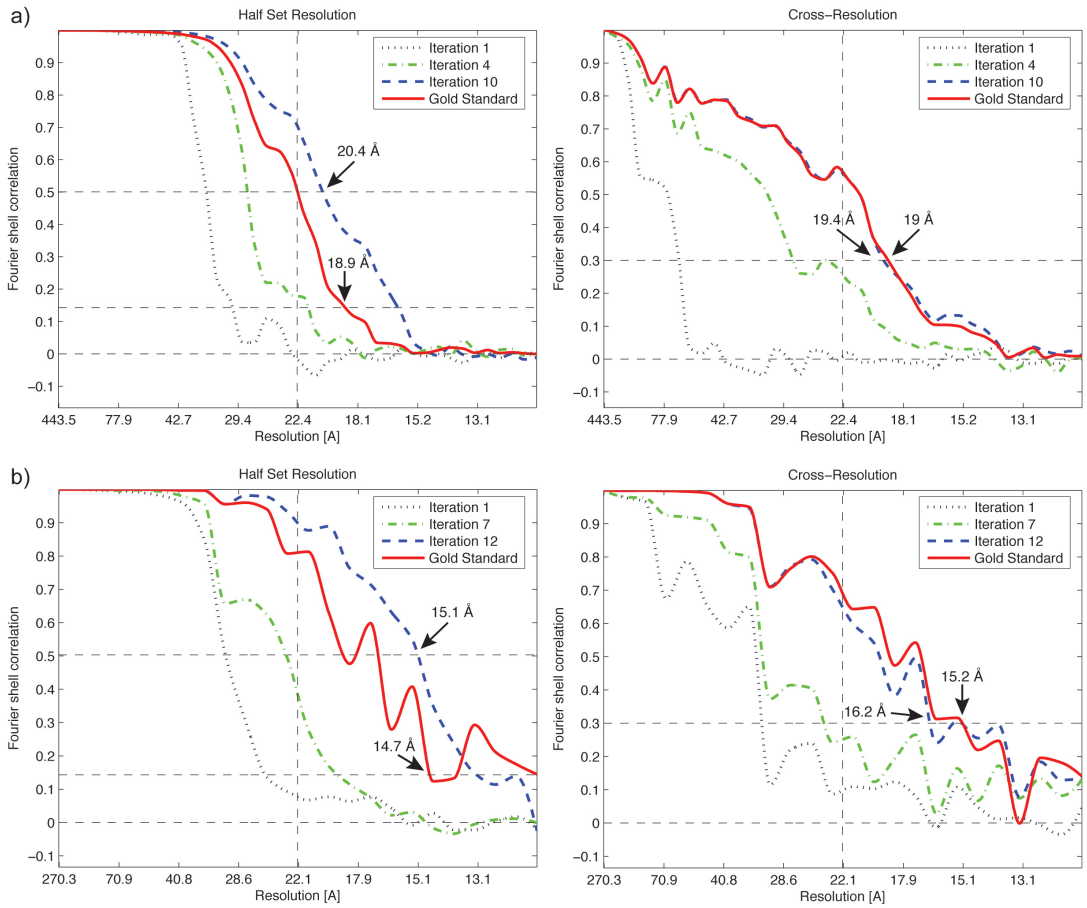

Figure 5.9: (a) The pairwise FSC and cross-resolution FSC curves of 80S ribosome averages from various iterations of the RFA procedure, as well as alignment of gold-standard FSC. (b) Same for the subtomogram averages of 20S proteasomes.

result. The protocol was tested on subtomograms of *S. cerevisiae* ribosomes from cell lysate and purified *Thermoplasma acidophilum* 20S proteasomes. The results suggest that the proposed alignment procedure is suitable to resolve macromolecular complexes well beyond 20 Å.

It is anticipated that even higher resolution averages can be obtained in the future using larger datasets and improved methods for CTF correction, which will greatly benefit from the new generation of direct electron detectors allowing more accurate defocus determination [Glaeser et al., 2011, Guerrini et al., 2011]. It will be interesting to see whether subtomogram averaging will, for example, allow resolving transmembrane helices in lipid membranes in future studies. In any case, the computational framework proposed here provides a solid basis for fast and accurate subtomogram alignment and enables processing of massive datasets in the future.

6

Subtomogram Classification

6.1 Introduction

In this chapter, the focus is on the subtomogram classification, which is particularly challenging due to several reasons: (i) The signal-to-noise ratio (SNR) of cryo-electron tomograms is poor (typically in the range of 0.1 - 0.01). (ii) The tilt range for data acquisition is limited, typically from $-60°$ to $60°$, which results in an incomplete sampling in Fourier space (missing wedge problem). (iii) The number of classes is typically unknown beforehand. (iv) The classes of subtomograms can be unbalanced (strongly differing populations). (v) The structural differences between the class averages can be subtle.

It is worth clarifying here that the term classification in the"subtomogram classification" is irrelevant to the one used in the machine learning field, which typically means the supervised learning. In this chapter, to avoid the confusion, the term classification means to classify the subtomograms.

Despite the successes of various classification methods (section 2.2.8) when applied to respective datasets, their performances tend to be limited in particular for unbalanced classes and subtle structural differences. Here, an unsupervised learning approach named AC3D (Autofocused Classification of 3D cryo-electron subtomograms) is proposed that can automatically focus the classification on the most variable parts of 3D structures. This similarity metric can capture subtle differences and does not involve any human intervention, thus alleviating bias. Based on this metric, an iterative multi-reference clustering scheme is introduced that makes use of the FVM algorithm for subtomogram alignment (chapter 5) to achieve a substantial speedup. Moreover, k-means++ is adapted as the initialization strategy for the clustering procedure to avoid being trapped in local optima and to accelerate the convergence. Comparisons of AC3D against the Constrained Principle Component Analysis (CPCA) approach [Förster et al., 2008] and the maximum likelihood approach MLTOMO [Scheres et al., 2009] on a simulated dataset show significant

Algorithm 6.1 AC3D

Input: SS: A set of input subtomograms, k: Number of classes.
Output: CS: Class-labeled subtomograms.

1: Pre-align SS
2: Initialize k class centers $SV = \{V_1, \ldots, V_k\}$
3: **while** #class changes $> 0.5\%$ **do**
4: Align SS to SV and obtain the corresponding scores SCS
5: Determine the noise class so that $SS = SS' \cup SS_{noise}$
6: Calculate the focused scores (FSS) of SS' with respect to SV
7: Determine the class labels according to FSS, which results in CS
8: Update the alignment of CS according to the class assignment
9: Average classes in CS to get the new class centers SV
10: **end while**
11: **return** CS

improvements of classification accuracy. Application of AC3D on experimental cryo-tomograms of ER-associated ribosomes yields clearly distinct conformations including established ribosome states without any human intervention or prior knowledge.

This chapter is based on a previous publication in [Chen et al., 2014].

6.2 Overall Classification Workflow

The overall workflow is first briefly described in Algorithm 6.1 and some important components will be explained in the following sections. The iterative optimization procedure of AC3D is a multi-reference scheme, which is closely related to k-means clustering. However, a more efficient initialization is used (see section 6.3). The basic workflow goes as follows: Firstly, the subtomograms are pre-aligned using the FVM algorithm described in chapter 5. After initialization the class centers (the subtomogram averages) are computed. During each iteration, subtomograms are aligned and assigned to the "closest" class center. All the class centers get updated subsequently using the assigned class members and their respective alignments. The whole procedure iterates until it converges or the maximal number of iterations is reached.

There are a few challenges when implementing this algorithm for CET. Firstly, an appropriate similarity metric is required to measure the "distance" of each subtomogram to the class average. Constrained cross-correlation (CCC) is used here, which constrains the correlation to the commonly sampled region in Fourier space [Förster et al., 2008]. However, computing the CCCs is time-consuming because each subtomogram has to be optimally aligned to the class centers prior to computing the CCC. The alignment is a problem of 6 DoFs, 3 for translation and 3

for rotation. The speed of subtomogram alignment is the bottleneck of the entire procedure and thus limits its practical use. This problem is tackled by the FVM algorithm presented in chapter 5. Secondly, the SNR of cryo-electron tomograms is relatively low making it difficult to identify outliers/noise that may deteriorate the clustering performance. This problem is explicitly handled here using the score distribution functions. Thirdly, it is difficult to classify subtle structural differences in CET data. The CCC quantifies the similarity between two volumes globally or within a subjectively chosen real-space mask of interest [Förster et al., 2008]. An objective and robust way has to be found to define the mask where significant differences are located because the noise may otherwise deteriorate the classification performance. Here, an algorithm is introduced to automatically focus the clustering on the variable parts of the macromolecule of interest and calculate the so-called focused score as the similarity measure. Some features will be discussed in detail in the following sections.

6.3 Initialization of Class Assignment

K-means clustering normally starts with a random initialization of the class assignment. Nevertheless, it is known that the performance of k-means strongly depends on the starting condition. There is no guarantee that the global optimum can be achieved. Moreover, a bad initialization decelerates the convergence of k-means. A common strategy is to run k-means multiple times with different seeds and then to choose the result with the best score as the final output. However, this strategy is not applicable here because each iteration is computationally intensive.

Arthur and Vassilvitskii proposed an algorithm named k-means++ to improve the initialization step [Arthur and Vassilvitskii, 2007]. The basic idea is to choose k cluster centers successively, each of which is randomly picked with a probability proportional to its squared distance from the closest existing center. It is shown that k-means++ converges faster than k-means with random initialization and guarantees it is $O(\log k)$-competitive with the optimal clustering. In contrast, the performance of k-means with random initialization can be arbitrarily worse than the optimum [Kanungo et al., 2004].

Here, k-means++ was implemented with a few important modifications for application to CET (Figure 6.1): 1. The class center is not a single subtomogram, but rather an average of a certain portion of the whole dataset containing N subtomograms. The reason is that one single subtomogram has low SNR and is affected by the missing wedge. 2. The first class center is the average of the aligned subtomograms with top $\lfloor N/k \rfloor$ scores, which are obtained by the CCCs from the prealignment. This class is thus similar to the average of the whole dataset. 3. The subsequent class centers are the averages of $\lfloor N/k \rfloor$ subtomograms from the whole dataset. These subtomograms are chosen at random with probabilities proportional

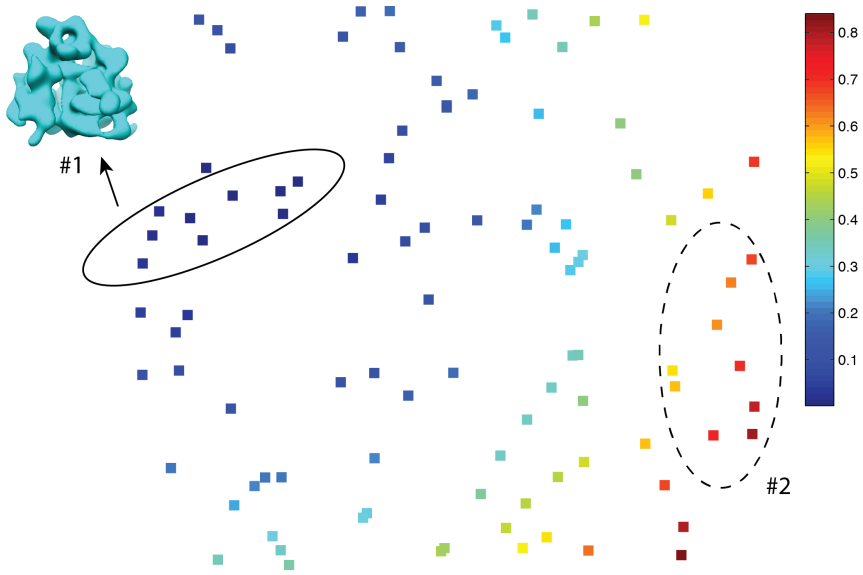

Figure 6.1: K-means++ is adapted as the initialization strategy. In the two-dimensional simplification each square represents a subtomogram. The distance between the squares is the distance measure given by Equation 6.1. For convenience of the illustration, the 2D simplification is drawn here. Assuming a subset of subtomograms (upper left, outlined with a solid line) is already chosen yielding the first class center (#1), the next class center (#2) is then the subtomogram average of a new subset (e.g., bottom right, outlined with a dashed line), in which each subtomogram is randomly picked with a probability proportional to the squared distance function (indicated by the colors of the squares and the scalebar).

to the squared distance functions. 4. The distance function D used here is the normalized Euclidean distance, which can be derived from the CCC. Mathematically, given a set of class centers $SV = \{V_1, \ldots, V_k\}$ and a subtomogram S, D can be calculated as:

$$D(S, SV) = \min_{V \in SV} \sqrt{2 - 2 \cdot CCC(V, S)}. \tag{6.1}$$

The final initialization algorithm is presented in Algorithm 6.2. The computational cost of this step is marginal compared to the others in Algorithm 6.1 and the whole clustering procedure normally converges faster with this new strategy.

6.4 Noise Class Handling

The subtomograms under investigation often include outliers, typically false positives from the automated or manual detection or subtomograms that are too noisy to be aligned accurately. These outliers tend to degrade the clustering performance.

Algorithm 6.2 Initialization of AC3D

Input: SS: A set of input aligned subtomograms, k: Number of desired classes.
Output: SV: A set of initial class centers.

$n = \lfloor N/k \rfloor$
Sort SS according to the scores and average the top n subtomograms to get V_1
$SV = \{V_1\}$
for $i = 2 : k$ **do**
  $SS' = \{\}$
  for $j = 1 : n$ **do**
    $\forall S \in SS$, calculate $P \propto D^2(S, SV)$
    Pick $S_j \in SS$ without replacement at random with probability P_j
    $SS' \leftarrow SS' \cup \{S_j\}$
  end for
  Average SS' to get V_i and $SV \leftarrow SV \cup \{V_i\}$
end for
return SV

To ensure the robustness of the classification with respect to such outliers, a certain percentage of all the subtomograms is assigned to a "noise class" during each iteration. This step is conducted before the class label determination step. If a subtomogram is assigned to the noise class it will be excluded from the remaining steps of that iteration. Importantly, the subtomogram will be included again in the subsequent iterations and may be assigned to a different class.

To decide which subtomogram belongs to the noise class, the probability is calculated using the score distributions. Given a set of subtomograms $SS = \{S_1, \ldots, S_N\}$ and a set of class centers $SV = \{V_1, \ldots, V_k\}$, SS is first aligned to SV using the FVM algorithm. For each $V_j \in SV$, we will have a set of similarity scores (score distribution function): $SCS_j = \{SC_1, \ldots, SC_N\}$. Assuming the noise class has low score and it is statistically independent of all class centers SV, the probability of a subtomogram $S_i \in SS$ not belonging to a class V_j is proportional to the distance of the subtomogram to the class center (Equation 6.1): $P_{i,j}\{SC_i < SC\}, \forall SC \in SCS_j$. Finally, the overall probability of S_i being noise can be calculated by the joint probabilities of not belonging to any class: $P_i = \prod_{j=1}^{k} P_{i,j}$. Sorting the probabilities and setting a threshold of the list will then yield the noise class.

6.5 Focus Mask and Focused Score

Another critical step is the automatic calculation of the focus mask FM and the corresponding focused score FS. Given two volumes (class centers) V_1 and V_2, FM is calculated as follows (Figure 6.2): 1. Low-pass filter V_1 and V_2 (according to the

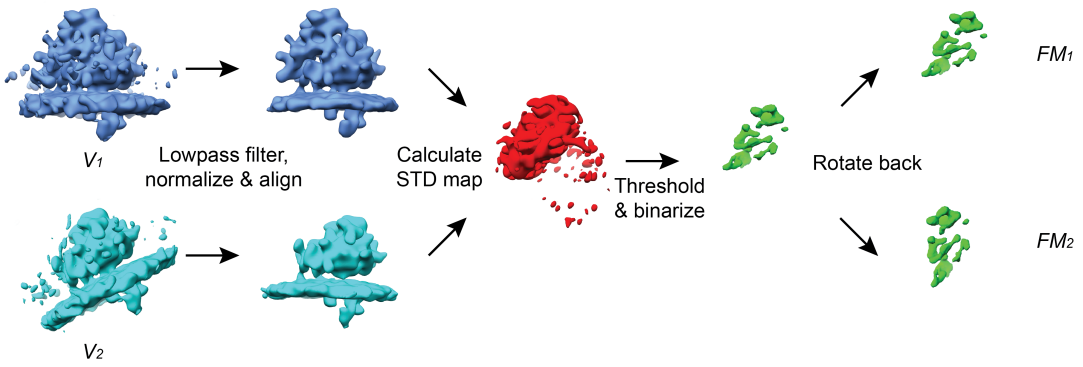

Figure 6.2: Steps of calculating the focus masks FM_1 and FM_2 of two class centers V_1 and V_2.

corresponding resolution) to reduce noise influence and normalize them (mean=0, std=1) so that they have approximately the same intensity scale. 2. Align V_1 and V_2 to make sure they have the highest correlation. 3. Multiply the aligned V_1 and V_2 with a mask M, if provided, to enforce FM to be computed inside M. Note this step is optional and M is only used for explicitly constraining the classification, e.g., filtering out hypervariable areas. 4. Calculate the standard deviation map STD of the aligned V_1 and V_2. In this case STD of two volumes is essentially the same as their absolute difference map. 5. Threshold STD according to a specified value and binarize it by setting the areas above the threshold to 1 and those below it to 0 resulting in FM. 6. Transform FM back to the respective orientations and positions of V_1 and V_2, which results in a pair FM_1 and FM_2. Note for each pair (V_1, V_2), their focus masks are also a pair (FM_1, FM_2).

Finally, $FS_{j,i}$ of a subtomogram S_i and V_j can be obtained by first aligning S_i to V_j and then calculating the local CCC [Förster et al., 2008]:

$$FS_{j,i} = \sum_{x,y,z} S'_i(x,y,z) \cdot V'_j(x,y,z), \tag{6.2}$$

$$S'_i = \frac{FM_j(x,y,z) \cdot (FT^{-1}(\hat{S}_i \cdot \omega) - \overline{S'_i})}{\sqrt{\sum_{x,y,z}(FM_j(x,y,z) \cdot (FT^{-1}(\hat{S}_i \cdot \omega) - \overline{S'_i}))^2}},$$

$$\overline{S'_i} = \frac{1}{\sum_{x,y,z} FM_j(x,y,z)} \sum_{x,y,z} FT^{-1}(\hat{S}_i \cdot \omega).$$

Herein, $\hat{S}_i$ is the Fourier transform of S_i and ω is the corresponding sampling region in Fourier space. V'_j can be computed analogously. Note that if FM_j is a unit volume, $FS_{j,i}$ is identical to $CCC_{j,i}$.

The most important factor involved in calculating the focused score is the thresh-

old. For the convenience of the specification, a parameter t is introduced here: $threshold = mean(STD) + t * (max(STD) - mean(STD))$. It is obvious that if $t = 0$, the threshold equals to the mean value of the STD map; if $t = 1$, the maximum value of the STD map is used as the threshold. The parameter t indirectly determines the size of the focus mask. For the applications in CET, $0.2 - 0.4$ is normally a good range.

6.6 Multiclass Label Determination

Binary class label determination is straightforward. Given a subtomogram S_i and two class centers (V_1, V_2), we first calculate $(FM_{1,i}, FM_{2,i})$ and their corresponding $FS_{1,i}, FS_{2,i}$. The class label of S_i will then correspond to the class average with the larger value between $FS_{1,i}$ and $FS_{2,i}$.

Multiclass label determination, i.e., class assignment to more than two classes, is not trivial because FM is defined pairwise. Focus masks that incorporate the structural discrepancies of more than two volumes are less discriminative than those pinpointing pairwise differences because the focus mask of multiple volumes will involve more voxels than any pairwise FM. In order to use the pairwise FM for classification a voting strategy is used for the multiclass label assignment (Figure 6.3). The focus score is defined with respect to a pair of class centers and can be therefore considered as a binary classifier, which can produce a vote to one of the classes between the pair. For each comparison of a subtomogram S_i with a pair of class centers, the binary class label is determined according to the vote. The final class label of S_i is determined by a voting of all the pairwise comparisons.

6.7 Dataset Preparation

6.7.1 Simulation of Ribosome Subtomograms

Three different states of ribosomes were simulated using atomic models from the Protein Data Bank (PDB) (Figure 6.4a): the *Saccharomyces cerevisiae* 80S ribosome (IDs: 3IZB, 3IZE, 3IZF and 3IZS), the *S. cerevisiae* 80S ribosome bound to the Sec61 translocon (ID: 2WWB) and the *S. cerevisiae* 80S ribosome bound to the Signal Recognition Particle (SRP) (ID: 1RY1). The simulations were conducted as described in section 3.5 for SNR=0.01. For testing the performance on an unbalanced dataset, the number of particles for each class was 150, 100, 50, respectively. Furthermore, 100 noise particles were added into the dataset to test the robustness. They were spheres with diameters ranging from 15 to 30 nm. They had similar mean values as the 80S ribosome and the same SNR. In total 400 subtomograms of size 100^3 voxels were simulated with a defocus of 4 µm and voxel size 0.47 nm. The tilt angles ranged from $-60°$ to $60°$ with $3°$ as the angular increment. The tomograms

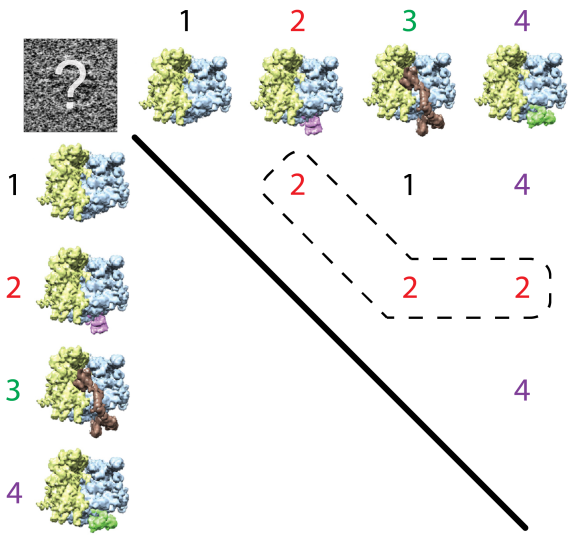

Figure 6.3: Multiclass label determination. A voting strategy is used for multiclass label determination. The subtomogram under investigation with unknown class label (top left) will be assigned to the class with the most votes from pairwise comparisons, i.e., class 2 in this case.

were randomly translated with respect to the center within the range of 10 voxels and randomly rotated.

6.7.2 Experimental Dataset of ER-Associated Ribosomes

Rough microsomes were prepared from dog pancreas and vitrified on lacey carbon molybdenum EM grids (Ted Pella, USA) as described in [Pfeffer et al., 2012]. Tilt series were acquired using a FEI Titan Krios TEM equipped with a Gatan "K2 summit" direct electron detector, operated in frame mode with 5-7 frames per projection image. The TEM was operated at an acceleration voltage of 300 kV. Single-axis tilt series were recorded from $-60°$ to $+60°$ with an angular increment of $2°$ at a nominal defocus of 4 µm and an object pixel size of 2.62 Å using the Serial EM acquisition software [Mastronarde, 2005]. The cumulative electron dose did not exceed 60 electrons/Å^2.

Frames from the K2 DDD were aligned using quasi-expectation maximization implemented in the MATLAB toolbox AV3 [Förster et al., 2005]. Phase correction of single projections was performed using the MATLAB scripts described in [Eibauer et al., 2012], and tomogram reconstruction (object pixel: 2.1 nm) and template matching were accomplished using PyTom [Hrabe et al., 2012] as described in [Pfeffer et al., 2012], followed by extraction of ribosome candidates. No Wiener filtering was applied in this case. A preliminary classification [Förster et al., 2008] was carried

out to remove most of the false positives, e.g., gold markers, ER membranes or carbon edges. Finally, 2,584 subtomograms (200^3 voxels, object pixel: 0.262 nm) were retained and reconstructed for further processing.

6.8 Results

6.8.1 Classification of Simulated Ribosome Subtomograms

AC3D was first assessed on a simulated dataset of *S. cerevisiae* 80S ribosomes bound to different cofactors involved in SRP mediated protein translocation into the Endoplasmic Reticulum (ER) (Figure 6.4a): the 80S ribosome alone, the 80S ribosome bound to the Sec61 translocon, and the 80S ribosome bound to the SRP. For convenience, we name the 80S ribosome as class #1, the 80S ribosome bound to the Sec61 channel as class #2, the 80S ribosome bound to the SRP as class #3, and noise particles as class #0. The parameter t (section 6.5) was set to 0.2. For comparison, this dataset was also classified into four classes using CPCA in combination with k-means clustering [Förster et al., 2008] and the maximum likelihood approach MLTOMO implemented in Xmipp [Scheres et al., 2009]. For CPCA 5 eigenvectors were retained for k-means and for MLTOMO 20 iterations were executed with reg0=5, regF=0 and reg_steps=5.

The confusion matrices are shown in Table 6.1, in which also the true positive rates (TPR) and false positive rates (FPR) are listed. Table 6.1 indicates a significantly better performance of AC3D compared to CPCA and MLTOMO in terms of both, TPR and FPR. Moreover, the classification results of AC3D (class centers) are shown in Figure 6.4b, in which the 3D densities are colored by the STD map (prior to threshold) to illustrate the autofocus ability of AC3D. All the tested methods exhibited excellent performances of distinguishing the noise class. This is due to the property of the noise simulation here. Currently, the assumption is that there is no correlation between the noise and the structure signal. The Gaussian white noise was added into the projections according to the specified SNR. As a result, the simulated noise particles were relatively easy to be classified out. However, the noise situation in reality is much more complicated and not completely understood. Thus, the real case scenario is usually more difficult. The noise was added here as a control.

To demonstrate the benefits of two key components of AC3D, i.e., the advanced initialization (k-means++) and the focused score, the classification results of AC3D were evaluated with each of these two features turned off (Table 6.2). When the random class assignment was used in the initialization step, the obtained accuracies were essentially identical in this case, but the convergence was slower (2 more iterations) compared to AC3D with k-means++. Thus, k-means++ increases the classification speed. When the conventional CCC was used as the similarity metric

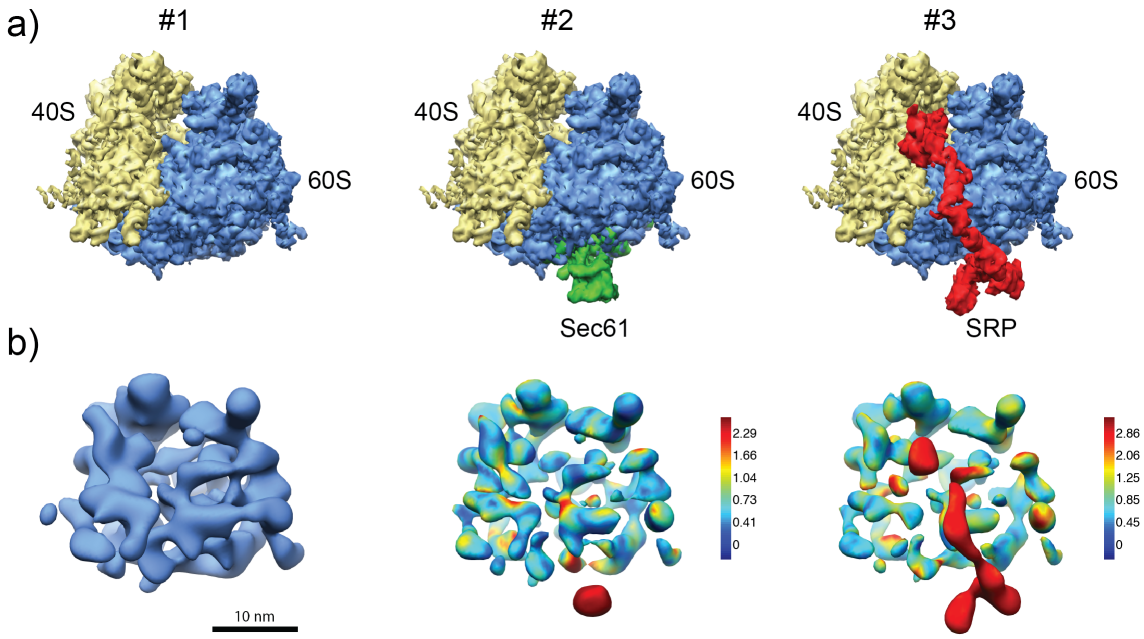

Figure 6.4: Classification result on simulated dataset of 80S ribosomes. (a) Densities simulated from atomic models of three ribosome states. From left to right: *S. cerevisiae* 80S ribosome (model #1), 80S ribosome bound to the Sec61 translocon (model #2), and 80S ribosome bound to the SRP (model #3). (b) Classification result. From left to right: subtomogram average of class #1, subtomogram average of class #2 colored by the STD against #1 and subtomogram average of class #3 colored by the STD against #1.

CPCA					
		Predicted			
		#0	#1	#2	#3
Actual	#0	100	0	0	0
	#1	15	76	59	0
	#2	8	56	36	0
	#3	7	0	0	43

	TPR	FPR
#0	100%	10%
#1	50.7%	22.4%
#2	36%	19.7%
#3	86%	0%

MLTOMO					
		Predicted			
		#0	#1	#2	#3
Actual	#0	100	0	0	0
	#1	2	106	42	0
	#2	1	68	31	0
	#3	0	28	15	7

	TPR	FPR
#0	100%	1%
#1	70.7%	38.4%
#2	31%	19%
#3	14%	0%

AC3D					
		Predicted			
		#0	#1	#2	#3
Actual	#0	93	7	0	0
	#1	4	125	21	0
	#2	2	0	98	0
	#3	1	0	0	49

	TPR	FPR
#0	93%	2.3%
#1	83.3%	2.8%
#2	98%	7%
#3	98%	0%

Table 6.1: Results of compared classification approaches for simulated ribosome dataset. Classes #1-#3 are shown in Figure 6.4 and class #0 corresponds to the noise class. From the class assignments the true positive rate (TPR) and false positive rate (FPR) were computed.

AC3D with random initialization					
		Predicted			
Actual		#0	#1	#2	#3
	#0	91	6	2	1
	#1	6	123	21	0
	#2	3	0	97	0
	#3	0	0	0	50

	TPR	FPR
#0	91%	3%
#1	82%	2.4%
#2	97%	7.7%
#3	100%	0.3%

Convergence: 8 iterations

AC3D without focused score					
		Predicted			
Actual		#0	#1	#2	#3
	#0	93	1	0	6
	#1	2	74	54	20
	#2	3	49	37	11
	#3	2	8	33	7

	TPR	FPR
#0	93%	2.3%
#1	49.3%	23.3%
#2	37%	29%
#3	14%	10.6%

Convergence: 7 iterations

AC3D					
		Predicted			
Actual		#0	#1	#2	#3
	#0	93	7	0	0
	#1	4	125	21	0
	#2	2	0	98	0
	#3	1	0	0	49

	TPR	FPR
#0	93%	2.3%
#1	83.3%	2.8%
#2	98%	7%
#3	98%	0%

Convergence: 6 iterations

Table 6.2: Influence of AC3D's initialization and focused score on classification accuracy. Classification of the simulated ribosome dataset was performed by AC3D with random initialization and with a uniform FM for comparison with the AC3D enabling all the features.

in AC3D, the classification accuracy degraded dramatically. Thus, the superior classification performance of AC3D compared to CPCA and MLTOMO can be almost exclusively attributed to the focused score.

Finally, the parameter t in AC3D was evaluated at different levels (0, 0.2, 0.4, 0.6) to demonstrate its influence. The confusion matrices are shown in Table 6.3, which shows the best performance was achieved when $t = 0.2$. The classifications were inferior when $t = 0$ and $t = 0.4$, which indicates the underlying focus masks were either too big or small. Lastly, $t = 0.6$ yielded the worst accuracy.

6.8.2 Classification of ER-Associated Ribosomes

AC3D was further tested on an experimental dataset of mammalian ribosomes bound to the ER protein translocon. In previous studies of the same sample, the membrane-bound 80S ribosome and two complexes with prominent lumenal domains were resolved [Pfeffer et al., 2012, Pfeffer et al., 2014]: the translocon-associated protein complex (TRAP) and the oligosaccharyl-transferase complex (OST). The acquired subtomograms depict ribosomes bound to ER-derived microsomes. Due to the highly variable diameters of the microsomes the curvature of the membrane would dominate the classification; to prevent classification according to membrane curvature the classification was constrained on the ribosome and the ER lumenal region. This was achieved by constructing a mask M (section 6.5) covering the corresponding region and excluding the membrane part.

The whole dataset was firstly classified into 4 classes by AC3D with $t = 0.4$. No noise class was specified here in order to be a fair comparison with [Pfeffer et al., 2014]. The resulting four classes are depicted in Figure 6.5a: class #1 clearly

AC3D: $t = 0$					
		Predicted			
Actual		#0	#1	#2	#3
	#0	95	5	0	0
	#1	4	130	14	2
	#2	1	7	60	32
	#3	0	1	4	45

	TPR	FPR
#0	95%	1.7%
#1	86.7%	5.2%
#2	60%	6%
#3	90%	9.7%

AC3D: $t = 0.2$					
		Predicted			
Actual		#0	#1	#2	#3
	#0	93	7	0	0
	#1	4	125	21	0
	#2	2	0	98	0
	#3	1	0	0	49

	TPR	FPR
#0	93%	2.3%
#1	83.3%	2.8%
#2	98%	7%
#3	98%	0%

AC3D: $t = 0.4$					
		Predicted			
Actual		#0	#1	#2	#3
	#0	92	6	0	2
	#1	6	75	47	22
	#2	1	10	89	0
	#3	1	0	0	49

	TPR	FPR
#0	92%	2.7%
#1	50%	6.4%
#2	89%	15.7%
#3	98%	6.9%

AC3D: $t = 0.6$					
		Predicted			
Actual		#0	#1	#2	#3
	#0	95	2	1	2
	#1	2	69	23	56
	#2	2	30	68	0
	#3	1	0	0	49

	TPR	FPR
#0	95%	1.7%
#1	46%	12.8%
#2	68%	8%
#3	98%	16.6%

Table 6.3: Influence of AC3D's parameter t on classification accuracy. Classification of the simulated ribosome dataset was performed by AC3D with different parameters t: 0, 0.2, 0.4 and 0.6.

captures 80S ribosomes bound to a translocon population with only TRAP, class #2 80S ribosomes bound to a translocon population with TRAP and OST, class #3 60S large ribosomal subunits with only TRAP, and class #4 60S ribosomal subunits associated with TRAP- and OST-containing translocons. The populations of the four classes are 564 (21.8%), 970 (37.5%), 737 (28.5%) and 313 (12.1%) particles, respectively. The classification agrees well with the results in [Pfeffer et al., 2014], where the foci for classification were chosen based on biological prior knowledge. In detail, CPCA classification was conducted on the same dataset with firstly a sphere mask focusing on the entire ribosome and then another sphere mask covering the ER-lumenal region [Pfeffer et al., 2012, Pfeffer et al., 2014]. The resulting class averages are essentially the same as depicted in Figure 6.5a. The confusion matrix

		AC3D			
		#1	#2	#3	#4
CPCA	#1	299	381	250	68
	#2	253	575	21	14
	#3	9	8	327	49
	#4	3	6	139	182

Table 6.4: Confusion matrix of classification results from CPCA and AC3D on ER-associated ribosomes.

of the classification results from CPCA and AC3D is shown in Table 6.4.

Moreover, a further classification round was conducted on the particles included in classes #1 and #2, focusing on the 80S ribosome part only. The number of classes was set to 3 and the obtained class averages are shown in Figure 6.5b. Consistent with previous studies using cryo-electron microscopy single particle analysis [Frank and Gonzalez, 2010, Melnikov et al., 2012, Wilson and Doudna Cate, 2012], a highly flexible ribosomal L1-stalk is observed (Figure 6.5b, right panel). Furthermore, we find a non-ribosomal density of approximately 100 kDa bound to the ribosomal stalk base in classes C1 and C2, but not C3, which likely corresponds to canonical translation elongation or termination factors. The number of subtomograms assigned to class C1 was 637 (41.5%), class C2 507 (33%) and class C3 390 (25.4%). The classification result was furthermore quantitatively assessed by the Fourier Shell Correlation (FSC) curves. Three types of FSC curves were calculated for each class: intra-class FSC, inter-class FSC and FSC of a random, same-sized portion of subtomograms (Figure 6.5c), from which we can see the intra-class FSCs are generally better or similar than the random FSCs. Because the FSC measures the global similarity, which is dominated by the structurally invariant core ribosome, the superiority of intra-class FSCs is more obvious when compared to inter-class FSCs, which indicate the level of similarity between the different classes. Taken together, these classification results suggest that AC3D is capable of separating different conformations of ER-associated ribosomes, which all agree with previous studies relying on much larger datasets.

6.9 Discussion

Here, a multi-reference clustering algorithm (AC3D) was introduced for subtomogram classification and simultaneous alignment. For large datasets AC3D is computationally more efficient than clustering approaches requiring pairwise correlations of all subtomograms, such as PCA-based approaches [Bartesaghi et al., 2008, Förster et al., 2008]. The main distinguishing feature of AC3D among multi-reference approaches is the ability to automatically focus the similarity measurement to regions

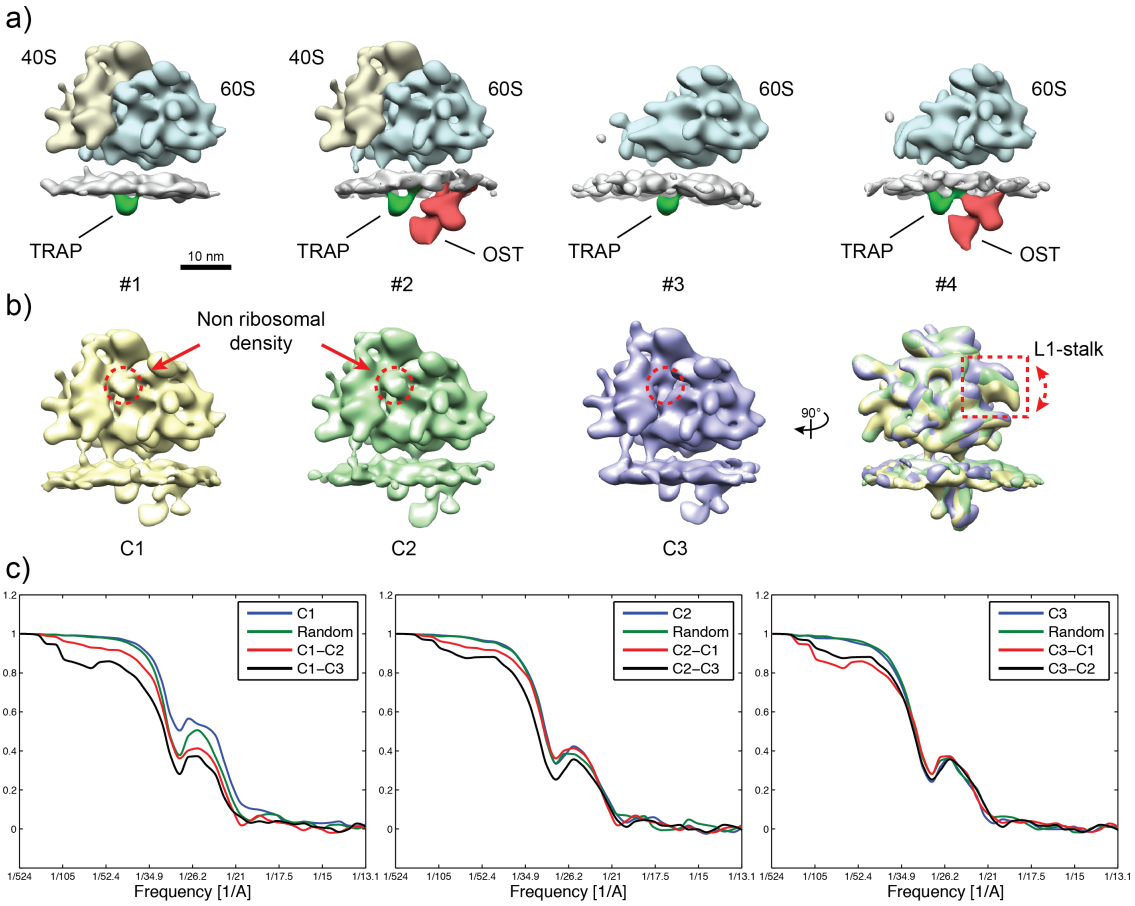

Figure 6.5: Classification result for mammalian ribosomes bound to the native ER protein translocon. (a) The whole dataset was first classified into 4 classes that apparently corresponded to the following assemblies: 80S ribosomes bound to a translocon population with only TRAP (class #1), 80S ribosomes bound to a translocon population with TRAP and OST (class #2), 60S ribosomes with only TRAP (class #3), and 60S ribosomes with TRAP and OST (class #4). (b) Classes #1 and #2 were merged and further classified into 3 classes (class C1, C2 and C3) with the focus on the ribosome density. The dotted circles mark the presence/absence of a non-ribosomal density bound to the ribosomal stalk base, which likely corresponds to canonical translation elongation or termination factors. The 3 class averages are overlaid on the rightmost side to show the high flexibility of the ribosomal L1-stalk (outlined with a dotted rectangle). (c) The FSC curves of the class averages in the panel b. For each class, three types of FSC curves are plotted: the intra-class FSC, the inter-class FSC and the FSC of a random portion with the same number of subtomograms.

of significant structural discrepancies. This autofocus ability does not require any prior knowledge or human intervention, which avoids hypothesis-driven bias of classification results. Moreover, k-means++ was adapted for the initialization of the iterative clustering algorithm, which improves the convergence speed and makes the procedure less vulnerable to local optima. Last but not least, the integration of the FVM algorithm for subtomogram alignment makes AC3D computationally highly efficient compared to other state-of-the-art approaches without compromising on accuracy.

When evaluated on a realistic simulated dataset of 80S ribosomes bound to different cofactors, AC3D achieved a nearly perfect classification of the different states while two other tested state-of-the-art classification approaches, CPCA and MLTOMO, yielded significantly less accurate class assignments. The dataset was designed such that it encapsulated three challenges of subtomogram classification: (i) In particular between two classes, the bare 80S ribosome and the 80S ribosome bound to the Sec61 channel, the structural difference arose from only a 60 kDa density, indicating that AC3D can identify highly subtle structural heterogeneity in low SNR data. (ii) The populations of different classes were unbalanced by a factor of up to three. The ability of handling the unbalance situation comes from two aspects. First, the class centers are generated by weighted averaging of the subtomograms (Equation 2.16). The weighting ensures a relatively fair comparison of the similarity, regardless of the size of different classes. Secondly, the focused score (section 6.5) and the voting strategy (section 6.6) allow a subtomogram to be assigned to a small class, even though the subtomogram's overall similarity with a larger class might be higher. This prevents the large class "absorbs" more and more entities from the others. (iii) A considerable amount of outliers was present. It is highly encouraging that AC3D yielded a near-to-perfect classification result under these challenging conditions, which often occur in experimental data from physiological samples.

AC3D was then applied to an experimental dataset of ER-associated ribosomes. For the ER-lumenal part of the complex, essentially the same classes were retrieved as the ones obtained previously using biological knowledge-based classifications [Pfeffer et al., 2014]: the OST complex was present in the translocon holocomplex in substoichiometric amounts. The most prominent classes for the cytosolic ribosomal density were assembled 80S ribosomes and 60S ribosomal subunits. Thus, the smallest structural difference detected in the initial classification was the presence/absence of the 250 kDa lumenal OST density. The significant enrichment of OST in translocon complexes bound to fully assembled 80S ribosomes (62.8% occupancy) compared to 60S ribosomal subunits (29.8% occupancy) suggests that OST has a higher affinity to translocon complexes engaged in cotranslational translocation of a nascent peptide across the ER membrane. This affinity variation would imply that the ER protein translocon is not a temporally invariant complex, but rather undergoes compositional dynamics according to the translational state of the

associated ribosome.

More subtle structural differences were detected when 80S ribosomal densities were classified, revealing well-established flexibility of the L1-stalk and cofactor binding to the ribosomal stalk base. The approximate mass of 100 kDa of the cofactor would be consistent for example with the 95 kDa eukaryotic elongation factor 2 (eEF2). Previously, different conformational states of the ribosome during translation could only be observed in CEM single particle data of purified ribosome particles. The classification results presented here for ribosomes in their native membrane suggest that CET in conjunction with subtomogram classification by AC3D will become a powerful method to study the mechanics of large macromolecular machines in their physiological environment.

7

Conclusion

In this thesis, a typical processing pipeline of CET is covered (Figure 1.3), which includes four main aspects: tomogram reconstruction, macromolecule identification, subtomogram alignment, and subtomogram classification. They are the key to high-resolution structural studies of macromolecules *in situ*. Several algorithms were presented here for each of these four steps. They were implemented as part of the open-source software package PyTom, with the aim of assisting and unifying the processing of CET data. Specifically, each individual aspect is briefly summarized in the following and an outlook for possible further developments is provided.

Tomogram Reconstruction. The first step after acquiring the projections is to reconstruct the 3D tomogram. Reconstruction itself is a vibrant research topic which has a long history, e.g., in medical imaging field, and attracts the attention of many mathematicians. The problem is well understood if the sampling is complete, making the underlying equation system unique to solve. This is, however, rarely the case in practice. Especially for CET, the specimen is imaged within a limited angular range. As a consequence, no unique solution exists and the reconstruction procedure becomes unstable. What makes the reconstruction in CET even more challenging is the low SNR, leading to an unknown amount of errors in the result.

NUFFT is of particular interest in such case. Previously, it has been shown that NUFFT exhibits excellent interpolation accuracy (forward problem) [Keiner et al., 2009]. It is also natural to formulate the reconstruction (backward problem) in such framework because the sampling points during image acquisition lie on a nonuniform grid in Fourier space. Based on these considerations, INFR was introduced in chapter 3 as an alternative Fourier-based reconstruction method. Its power is due to the NUFFT and iterative refinement procedure. It has superior reconstruction accuracy compared to the gridding method and yields tomograms with better contrast. More importantly, with help of the adjoint NUFFT operations in the iterations, INFR is able to recover certain missing information, resulting in a density map with less artifacts. However, the expense is the increased computational time and required

memory.

It will be interesting in the future to see if it is possible to find an appropriate regularization strategy to reduce the noise level or the artifacts caused by the "missing wedge" problem, without losing the high frequency information or introducing any structure bias.

One additional challenge of reconstruction in CET that has emerged in recent years is the rapid improvement of the detector resolution, which results in a large projection size of 4k×4k or even 8k×8k. The computer memory that would be required for reconstruction of full-sized tomograms is typically not available. For example, to reconstruct the full 4k×4k×4k tomogram from projections of size 4k×4k, at least 256 GB memory is needed for only single precision format. Generalization of the implementation of INFR to non-cubic volumes would reduce the memory requirement for thin specimens, but the memory requirement is nevertheless huge. While tomograms are often reconstructed at decreased resolution to provide an overview of the imaged sample, specific features of interest are required to be reconstructed at full resolution, for example for subtomogram averaging. Such reconstruction of specific subregions is a principal problem for reconstruction methods because the underlying equation systems become inconsistent due to the signal from the surrounding that will influence all the subprojections of interest. This is called "interior problem" [Kudo et al., 2008, Natterer, 2001], which mathematically cannot be solved uniquely. It raises challenges to the method described here, as well as to almost all reconstruction methods available until now. To what extent this problem will influence the final structure obtained by subtomogram averaging requires further studies. It would be extremely useful to develop an algorithm that can reconstruct the subtomograms at the original scale with affordable computing resources.

Another obstacle towards high-resolution reconstruction is the accurate correction of the CTF. As discussed in section 2.1.2, the Fourier transform of a projection obtained using a TEM is multiplied by a CTF, which depends on the settings of the microscope, e.g., the defocus, the accelerating voltage, the spherical aberration, etc. The oscillatory nature of CTF makes the image features of certain frequencies weak or even inverted. Therefore, a CTF correction procedure of the projections prior to reconstruction is critical, in order to resolve structural information faithfully in the high frequency region. Currently, the CTF correction in CET is still a challenging task and lacks of powerful tools for automatic correction of large amounts of data.

Macromolecule Identification. After the tomogram is reconstructed, it is ready for further analysis, such as macromolecule identification. If a specific type of macromolecule is of interest, its occurrences in the tomogram have to be extracted from the background. While large structures like membranes can be relatively easily segmented, it still remains infeasible to segment smaller structures (macromolecules). The major difficulties arise from the noise condition and the "missing wedge" phenomenon. The low SNR of the tomogram makes the structural features of macro-

molecules difficult to recognize, which explains why few of the advanced object detection methods from computer vision are applied in CET. This is made even worse by the deformation created by the "missing wedge" problem: the structural information along certain directions are unavailable.

A supervised learning approach was presented in chapter 4 to identify macromolecule in three steps. The major contribution is the development of rotation-invariant features for 3D data. They are based on spherical harmonics, which form an orthogonal basis system on the unit sphere. Importantly, there is an efficient algorithm for the spherical harmonics expansion, which is an analogue of the FFT in Cartesian space. Similar to the Fourier shift theorem, the rotation of spherical functions will not change the amplitudes of the coefficients from spherical harmonics expansions. As a consequence, the features can be computed fast and used as the rotation-invariant signatures for 3D data in SVM to improve the identification performance. The method was applied to an experimental tomogram of ER-associated ribosomes, where template matching suffered from a high percentage of false positives. In contrast, the proposed approach was able to reduce the false positive rate significantly by yielding an average of clear ribosome-specific features and of higher resolution.

Regardless of the advances demonstrated here, the macromolecule identification remains the most challenging task among all the processing steps and there is still room for improvement. For future work, it will be interesting to see if the proposed approach can be generalized to detect other type of macromolecular complexes. While the method was applied to detect ribosomes here, it is expected to provide similar advantages for the detection of other complexes. Hence, it is of special interest to determine the minimum mass of complexes that can be detected in tomograms with satisfactory specificity and sensitivity.

Another direction worth investigating is the unsupervised learning approaches, i.e., clustering of the features without the training phase. Although the supervised learning approach has been demonstrated here to improve the identification performance, the prerequisite for an accurate detection is the training on a large dataset. This becomes an obstacle when applied to CET, because the training dataset is either simulated, which is an approximation of the reality, or is constructed by manual labelling of experimental tomograms, which is subjective and time-consuming. Based on this consideration, the unsupervised learning approach might be an alternative. It is also intriguing to develop advanced features for 3D structures, which should be noise robust and take the "missing wedge" effect into account. These features could be applied either in a supervised learning approach, or in an unsupervised approach, to improve the identification performance.

One additional challenge arises from detecting small macromolecules due to their poor contrast and high noise level. With the advance of the microscope hardware (e.g., DDD, phase plate, etc), a future direction is segmentation. It would be interesting to develop a segmentation algorithm tailored to CET, to see if it is possible to

detect small objects with the help of improved image qualities and denoising techniques. Segmentation is of special interest, because it is an alternative approach to template matching, which rules out any potential template bias.

Subtomogram Alignment. For *in situ* structural studies of macromolecules at high resolution, it typically requires aligning the subtomograms extracted based on the identification step to a common coordinate system and average them to enhance the SNR. This is, however, not a trivial task. In addition to the "missing wedge" problem, the contrast of each individual subtomogram is even lower than the whole tomogram, making the registration algorithms from the medical imaging field not applicable here. The alignment problem can be simplified under the assumption that the structures contained in the subtomograms are rigid and homogeneous.[1] Nevertheless, the alignment is still a complicated problem considering the complexity of the solution space: every subtomogram has 6 alignment parameters to be determined, making it extremely time-consuming to compute.

The main contribution of chapter 5 is to generalize the convolution theorem to the rotational space based on spherical harmonic analysis introduced previously in chapter 4. It is shown that the $SO(3)$ Fourier transform of the convolution between two spherical functions is equal to the pointwise multiplication of the spherical Fourier transforms of the two functions. Based on this theorem, a fast rotational matching (FRM) algorithm was proposed, in which the alignment problem is formulated in Fourier space. The "missing wedge" can thus be conveniently handled by introducing two mask functions. Afterwards, the FRM is integrated with the fast translational matching into a unified alignment framework (FVM) via quasi-expectation-maximization. As tested on a simulated dataset, a speedup of approximately three orders of magnitude was achieved by FVM without decrease of accuracy. Subsequently, it was shown that FVM was able to resolve structures beyond 20 Å on experimental datasets.

Several aspects can be improved in the future: (i) Without considering interpolation, the angular precision is $(\frac{360}{2B})^\circ$. For the current computing power, a descent maximum choice of B is 64, resulting in an angular precision of about 2.8°. This suffices for most scenarios. If higher precision is expected, it has to be combined with a local sampling of finer angular search. (ii) FRM searches the best angular match in the whole rotational space. In some cases, this is unnecessary. For instance, when there is prior knowledge about the orientation, it is preferred to constrain the angular search in the vicinity of a certain angle or around a certain rotation axis. These constraints can be easily imposed in FRM by multiplying $CCC(\boldsymbol{R})$ in Equation 5.12 with a mask function specifying the angular constraints. (iii) FVM is designed for homogenous populations of subtomograms. However, it can be readily incorporated into multi-reference alignment procedures. For example, the demonstrated

[1] The homogeneity assumption can be weakened by the subtomogram classification.

advantages of FVM should enable better alignment and hence better separation of particle populations [Xu et al., 2012]. FVM could also be integrated into maximum likelihood approaches [Scheres et al., 2009, Stölken et al., 2011], which would be an elegant solution to deal with structurally heterogeneous subtomograms. Thus, it is anticipated that the FVM algorithm described here will also be instrumental for dealing with heterogeneous data.

Subtomogram Classification. Structure analysis by subtomogram alignment and averaging assumes all the subtomograms depict the macromolecule with the same conformational state. This is, however, often not the case due to the heterogeneity of the dataset. The heterogeneity arises from either false positives of the identification step, or conformational changes when the macromolecules fulfill their biological functions *in situ*. Subtomogram classification is therefore of particular importance to group different states together, in order to improve the resolution or to study the conformations of macromolecules in their native environment.

Most of the state-of-the-art classification methods quantify the similarity between two subtomograms globally. A more intuitive way is to focus the comparison on the region of significant difference, which provides a stronger discriminative ability than the global comparison. This is the basic idea of AC3D presented in chapter 6. This autofocus does not require any prior knowledge and is calculated automatically, reducing the human bias. Furthermore, the FVM algorithm proposed in chapter 5 is used in the classification framework, providing a significant speedup without sacrificing the accuracy. AC3D was compared to two state-of-the-art approaches on a simulated dataset, where it showed a superior performance to classify the subtle structure variations and the dataset with unbalanced classes. AC3D was further evaluated on an experimental dataset of ER-associated ribosomes, where AC3D was able to classify distinct conformations of well agreement with the previous result [Pfeffer et al., 2014] obtained based on biological prior knowledge.

Whereas AC3D performs excellently for figuring out whether cofactors are absent/present in complexes as shown here, for the future work it is interesting to test how well AC3D works in the situations that the structure variations are not confined to small regions, e.g., to classify datasets consisting of different species (sizes and weights) of macromolecules.

Another challenge that AC3D shares with essentially all multi-reference classification approaches is that the user must specify the number of classes k, which is not straightforward. Too few classes will mix up different conformations, whereas an overestimated number of classes might result in an insufficient amount of subtomograms per class, increasing the negative influence of the noise. A common guideline is to oversample k properly, because this makes it more likely to discover small classes and the clustering result will become more stable. In a subsequent step, the classes can be either manually examined and aggregated, or automatically merged using hierarchical clustering of the class averages [Hrabe et al., 2012]. Nevertheless,

it would be beneficial to find a criterion to automatically determine the number of classes.

Software Platform. All the methods presented in this thesis have been implemented and included in the software PyTom (`www.pytom.org`) [Hrabe et al., 2012], an open-source tool aimed at providing a unified platform of data analysis for CET. PyTom covers a typical processing pipeline described above, as well as numerous handy scripts to assist the 3D image analysis. PyTom adopts a two-layer architecture: the low level is implemented in C/C++ and the high level in Python. The two layers are bridged using SWIG, a software tool designed for connecting programs or libraries written with different programming languages. The benefit of this architecture is to keep the computational efficiency, while providing a flexible and easily extendable interface for the developers and users. New functionalities and customized methods can be added in this platform rapidly without the pain of learning obscure programming languages or taking care of the computational efficiency. Moreover, considering the dramatic increase of the data size nowadays and computational demand thereby, Message Passing Interface (MPI) is employed in PyTom to parallelize the computations in, e.g., template matching, subtomogram alignment and classification steps. Parallelization becomes an indispensable tool enabling the step towards subnanometer resolution using CET. Finally, PyTom provides a web-based, user-friendly interface and a rich documentation, which allow the users to process their data smoothly and unambiguously.

In summary, CET is becoming an important imaging tool nowadays to study the 3D structures of macromolecular complexes in their near-to-native conditions. While there are still continuous (hardware or methodology) developments of CET technique, the processing of the CET data is also crucial and an active research field. The approaches presented in this thesis show a step towards establishing an automatic, accurate, efficient and unbiased analysing framework of CET data.

A

Generalized Convolution Theorem of Spherical Functions

The proof of Equation 5.4 is provided here. To rotate a spherical function $g(\boldsymbol{\eta}) \in \mathcal{L}^2(S^2)$, we have

$$\Lambda_{\boldsymbol{R}} g(\boldsymbol{\eta}) = g[\boldsymbol{R}^T \boldsymbol{\eta}] = \sum_l \sum_{m=-l}^{l} \hat{g}_m^l Y_m^l [\boldsymbol{R}^T \boldsymbol{\eta}]. \tag{A.1}$$

The spherical harmonics are rotated by U_{mn}^l the matrix elements of the irreducible representation of $SO(3)$ [Brink and Satchler, 1993]:

$$Y_m^l [\boldsymbol{R}^T \boldsymbol{\eta}] = \sum_n U_{mn}^l Y_n^l (\boldsymbol{\eta}). \tag{A.2}$$

Therefore, the rotated $g(\boldsymbol{\eta})$ can be expanded as:

$$\Lambda_{\boldsymbol{R}} g(\boldsymbol{\eta}) = \sum_l \sum_{m=-l}^{l} \sum_n \hat{g}_m^l U_{mn}^l Y_n^l (\boldsymbol{\eta}). \tag{A.3}$$

It can be seen that the rotation of a spherical function will not change its energy among degrees l:

$$\| \hat{g}_m^l \| = \| \hat{g}_m^l U_{mn}^l \|, \forall \boldsymbol{R} \in SO(3). \tag{A.4}$$

The correlation $SCC(\boldsymbol{R})$ in Equation 5.3 can then be calculated by using the Fourier expansions of $f(\boldsymbol{\eta})$ and $\Lambda_{\boldsymbol{R}} g(\boldsymbol{\eta})$:

$$\begin{aligned} SCC(\boldsymbol{R}) &= \int_{\boldsymbol{\eta} \in S^2} f(\boldsymbol{\eta}) \overline{\Lambda_{\boldsymbol{R}} g(\boldsymbol{\eta})} \mathrm{d}\boldsymbol{\eta} \\ &= \sum_l \sum_{m=-l}^{l} \sum_n \sum_{p=-n}^{n} \sum_{k=-n}^{n} \hat{f}_m^l \overline{\hat{g}_p^n U_{pk}^n (\boldsymbol{R})} \int_{\boldsymbol{\eta} \in S^2} \overline{Y_k^n (\boldsymbol{\eta})} Y_m^l (\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta}. \end{aligned} \tag{A.5}$$

Due to the orthogonality of the spherical harmonics function, we have:

$$\int_{\boldsymbol{\eta}\in S^2} \overline{Y_m^l(\boldsymbol{\eta})} Y_{m'}^{l'}(\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta} = \delta_{mm'}\delta_{ll'}, \tag{A.6}$$

where δ is the Kronecker delta function. The nonzero terms only appear when $m = m'$ and $l = l'$. Thus, Equation A.5 can be reduced to:

$$SCC(\boldsymbol{R}) = \sum_l \sum_{m=-l}^{l} \sum_{p=-l}^{l} \hat{f}_m^l \overline{\hat{g}_p^l U_{pm}^l(\boldsymbol{R})}. \tag{A.7}$$

By applying SOFT (Equation 5.2) to the above equation, we obtain:

$$SOFT(SCC(\boldsymbol{R})) = \hat{h}_{qr}^n = \sum_l \sum_{m=-l}^{l} \sum_{p=-l}^{l} \hat{f}_m^l \overline{\hat{g}_p^l} \int_{\boldsymbol{R}\in SO(3)} \overline{U_{pm}^l(\boldsymbol{R})} U_{qr}^n(\boldsymbol{R}) \mathrm{d}\boldsymbol{R}. \tag{A.8}$$

Due to the orthogonality:

$$\int_{\boldsymbol{R}\in SO(3)} U_{mp}^l(\boldsymbol{R}) \overline{U_{qr}^n(\boldsymbol{R})} \mathrm{d}\boldsymbol{R} = \delta_{ln}\delta_{mq}\delta_{pr}, \tag{A.9}$$

Equation A.8 have nonzero terms only when $l = n$, $m = q$ and $p = r$, which leads to:

$$SOFT(SCC) = \hat{h}_{mp}^l = \hat{f}_m^l \overline{\hat{g}_p^l}. \tag{A.10}$$

Thus, the proof of Equation 5.4 is completed.

Publication List

Chen, Y., Pfeffer, S., Fernández, J. J., Sorzano, C. O. S., Förster, F. (2014). Autofocused 3D Classification of Cryo-electron Subtomograms. *Structure*, 22(10), 1528-37.

Kuhn-Cuellar, L., Pfeffer, S., Chen, Y., Förster, F. (2014). Automated detection of polysomes in cryoelectron tomography. In *2014 IEEE International Conference on Image Processing (ICIP)*, (pp. 2085–2089)

Chen, X., Chen, Y., Schuller, J. M., Navab, N., Förster, F. (2014). Automatic particle picking and multi-class classification in cryo-electron tomograms. In *2014 11th IEEE International Symposium on Biomedical Imaging (ISBI)*, (pp. 838-841).

Chen, Y., Förster, F. (2013). Iterative reconstruction of cryo-electron tomograms using nonuniform fast Fourier transforms. *Journal of Structural Biology*, 185(3), 309-316.

Chen, Y., Pfeffer, S., Hrabe, T., Schuller, J. M., Förster, F. (2013). Fast and accurate reference-free alignment of subtomograms. *Journal of Structural Biology*, 182(3), 235-245.

Chen, Y., Hrabe, T., Pfeffer, S., Pauly, O., Mateus, D., Navab, N., Förster, F. (2012). Detection and identification of macromolecular complexes in cryo-electron tomograms using support vector machines. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, (pp. 1373-1376).

Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L. K., Mangold, A.-V., Förster, F. (2012). PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. *Journal of Structural Biology*, 178(2), 177-88.

Abbreviations

2D ..... Two Dimensional
3D ..... Three Dimensional
ART ..... Algebraic Reconstruction Technique
CC ..... Cross-Correlation
CCC ..... Constrained Correlation Coefficient
CCD ..... Charged Coupled Device
CEM ..... Cryo-Electron Microscopy
CET ..... Cryo-Electron Tomography
CPCA ..... Constrained Principal Component Analysis
CTF ..... Contrast Transfer Function
DDD ..... Direct Detection Device
DoF ..... Degree of Freedom
EM ..... Electron Microscopy
ER ..... Endoplasmic Reticulum
FEG ..... Field Emission Gun
FFT ..... Fast Fourier Transform
FPR ..... False Positive Rate
FRM ..... Fast Rotational Matching
FSC ..... Fourier Shell Correlation
FTM ..... Fast Translational Matching
FVM ..... Fast Volumetric Matching

LCC ..... Local Cross-Correlation
LNCC ..... Local Normalized Cross-Correlation
ML ..... Maximum Likelihood
MTF ..... Modulation Transfer Function
MPI ..... Message Passing Interface
NMR ..... Nuclear Magnetic Resonance
NUFFT ..... Nonuniform Fast Fourier Transform
PCA ..... Principal Component Analysis
PDB ..... Protein Data Bank
PSF ..... Point Spread Function
RFA ..... Reference-Free Alignment
ROC ..... Receiver Operating Characteristics
SART ..... Simultaneous Algebraic Reconstruction Technique
SCC ..... Spherical Cross-Correlation
SFT ..... Spherical Fourier Transform
SIRT ..... Simultaneous Iterative Reconstruction Technique
SH ..... Spherical Harmonics
SNR ..... Signal-to-Noise Ratio
SOFT ..... $SO(3)$ Fourier Transform
SPA ..... Single Particle Analysis
SRP ..... Signal Recognition Particle
SVM ..... Support Vector Machine
TEM ..... Transmission Electron Microscope
TM ..... Template Matching
TPR ..... True Positive Rate
WBP ..... Weighted Backprojection

Bibliography

[Amat et al., 2010a] Amat, F., Castaño Diez, D., Lawrence, A., Moussavi, F., Winkler, H., and Horowitz, M. (2010a). Alignment of cryo-electron tomography datasets. *Methods in enzymology*, 482:343–67.

[Amat et al., 2010b] Amat, F., Comolli, L. R., Moussavi, F., Smit, J., Downing, K. H., and Horowitz, M. (2010b). Subtomogram alignment by adaptive Fourier coefficient thresholding. *Journal of structural biology*, 171(3):332–44.

[Amat et al., 2008] Amat, F., Moussavi, F., Comolli, L. R., Elidan, G., Downing, K. H., and Horowitz, M. (2008). Markov random field based automatic image alignment for electron tomography. *Journal of structural biology*, 161(3):260–75.

[Armache et al., 2010] Armache, J.-P., Jarasch, A., Anger, A. M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Söding, J., Westhof, E., Wilson, D. N., and Beckmann, R. (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46):19748–53.

[Arthur and Vassilvitskii, 2007] Arthur, D. and Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, Philadelphia, PA, USA. Society for Industrial and Applied Mathematics.

[Barber et al., 1996] Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*, 22(4):469–483.

[Bartesaghi et al., 2005] Bartesaghi, A., Sapiro, G., and Subramaniam, S. (2005). An energy-based three-dimensional segmentation approach for the quantitative interpretation of electron tomograms. *IEEE Transactions on Image Processing*, 14(9):1314–23.

[Bartesaghi et al., 2008] Bartesaghi, A., Sprechmann, P., Liu, J., Randall, G., Sapiro, G., and Subramaniam, S. (2008). Classification and 3D averaging with missing wedge correction in biological electron tomography. *Journal of structural biology*, 162(3):436–50.

[Bartesaghi and Subramaniam, 2009] Bartesaghi, A. and Subramaniam, S. (2009). Membrane protein structure determination using cryo-electron tomography and 3D image averaging. *Current opinion in structural biology*, 19(4):402–407.

[Beck et al., 2009] Beck, M., Malmström, J. A., Lange, V., Schmidt, A., Deutsch, E. W., and Aebersold, R. (2009). Visual proteomics of the human pathogen Leptospira interrogans. *Nature methods*, 6(11):817–23.

[Bracewell, 1986] Bracewell, R. N. (1986). *The Fourier Transform and Its Applications*. McGraw-Hill, 2nd edition.

[Brandt et al., 2009] Brandt, F., Etchells, S. A., Ortiz, J. O., Elcock, A. H., Hartl, F. U., and Baumeister, W. (2009). The native 3D organization of bacterial polysomes. *Cell*, 136(2):261–71.

[Brandt et al., 2001a] Brandt, S., Heikkonen, J., and Engelhardt, P. (2001a). Automatic alignment of transmission electron microscope tilt series without fiducial markers. *Journal of structural biology*, 136(3):201–13.

[Brandt et al., 2001b] Brandt, S., Heikkonen, J., and Engelhardt, P. (2001b). Multiphase method for automatic alignment of transmission electron microscope images using markers. *Journal of structural biology*, 133(1):10–22.

[Brandt and Ziese, 2006] Brandt, S. S. and Ziese, U. (2006). Automatic TEM image alignment by trifocal geometry. *Journal of microscopy*, 222(Pt 1):1–14.

[Briggs et al., 2009] Briggs, J. A. G., Riches, J. D., Glass, B., Bartonova, V., Zanetti, G., and Kräusslich, H.-G. (2009). Structure and assembly of immature HIV. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27):11090–5.

[Brink and Satchler, 1993] Brink, D. M. and Satchler, G. R. (1993). *Angular Momentum*. Clarendon Press, Oxford.

[Brodsky, 2010] Brodsky, J. L. (2010). The use of in vitro assays to measure endoplasmic reticulum-associated degradation. *Methods in enzymology*, 470:661–79.

[Bronstein et al., 2002] Bronstein, M. M., Bronstein, A. M., Zibulevsky, M., and Azhari, H. (2002). Reconstruction in diffraction ultrasound tomography using nonuniform FFT. *IEEE transactions on medical imaging*, 21(11):1395–401.

[Castaño Díez et al., 2007] Castaño Díez, D., Al-Amoudi, A., Glynn, A.-M., Seybert, A., and Frangakis, A. S. (2007). Fiducial-less alignment of cryo-sections. *Journal of structural biology*, 159(3):413–23.

[Castaño Díez et al., 2012] Castaño Díez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *Journal of structural biology*, 178(2):139–51.

[Castaño Díez et al., 2010] Castaño Díez, D., Scheffer, M., Al-Amoudi, A., and Frangakis, A. S. (2010). Alignator: a GPU powered software package for robust fiducial-less alignment of cryo tilt-series. *Journal of structural biology*, 170(1):117–26.

[Chang and Lin, 2011] Chang, C. and Lin, C. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27.

[Chen and Förster, 2013] Chen, Y. and Förster, F. (2013). Iterative reconstruction of cryo-electron tomograms using nonuniform fast Fourier transforms. *Journal of structural biology*, 185(3):309–316.

[Chen et al., 2012] Chen, Y., Hrabe, T., Pfeffer, S., Pauly, O., Mateus, D., Navab, N., and Förster, F. (2012). Detection and identification of macromolecular complexes in cryo-electron tomograms using support vector machines. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1373–1376. IEEE.

[Chen et al., 2014] Chen, Y., Pfeffer, S., Fernández, J. J., Sorzano, C. O. S., and Förster, F. (2014). Autofocused 3D Classification of Cryoelectron Subtomograms. *Structure*, 22(10):1528–37.

[Chen et al., 2013] Chen, Y., Pfeffer, S., Hrabe, T., Schuller, J. M., and Förster, F. (2013). Fast and accurate reference-free alignment of subtomograms. *Journal of Structural Biology*, 182(3):235–245.

[Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3):273–297.

[Crowther et al., 1970] Crowther, R. A., DeRosier, D. J., and Klug, A. (1970). The Reconstruction of a Three-Dimensional Structure from Projections and its Application to Electron Microscopy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 317(1530):319–340.

[Cyrklaff et al., 2005] Cyrklaff, M., Risco, C., Fernández, J. J., Jiménez, M. V., Estéban, M., Baumeister, W., and Carrascosa, J. L. (2005). Cryo-electron tomography of vaccinia virus. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8):2772–7.

[Danev et al., 2010] Danev, R., Kanamaru, S., Marko, M., and Nagayama, K. (2010). Zernike phase contrast cryo-electron tomography. *Journal of structural biology*, 171(2):174–81.

[Davison, 1983] Davison, M. E. (1983). The Ill-Conditioned Nature of the Limited Angle Tomography Problem. *SIAM Journal on Applied Mathematics*, 43(2):428–448.

[DeRosier and Klug, 1968] DeRosier, D. J. and Klug, A. (1968). Reconstruction of three dimensional structures from electron micrographs. *Nature*, 217:130–134.

[Dierksen et al., 1993] Dierksen, K., Typke, D., Hegerl, R., and Baumeister, W. (1993). Towards automatic electron tomography II. Implementation of autofocus and low-dose procedures. *Ultramicroscopy*, 49(1-4):109–120.

[Dierksen et al., 1992] Dierksen, K., Typke, D., Hegerl, R., Koster, A., and Baumeister, W. (1992). Towards automatic electron tomography. *Ultramicroscopy*, 40(1):71–87.

[Dubochet et al., 1988] Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly reviews of biophysics*, 21(2):129–228.

[Eibauer et al., 2012] Eibauer, M., Hoffmann, C., Plitzko, J. M., Baumeister, W., Nickell, S., and Engelhardt, H. (2012). Unraveling the structure of membrane proteins in situ by transfer function corrected cryo-electron tomography. *Journal of structural biology*, 180(3):488–96.

[Erickson and Klug, 1970] Erickson, H. P. and Klug, A. (1970). The Fourier transform of an electron micrograph: effects of defocussing and aberrations, and implications for the use of underfocus contrast enhancement. *Berichte der Bunsengesellschaft für physikalische Chemie*, 74(11):1129–1137.

[Faruqi and McMullan, 2011] Faruqi, A. R. and McMullan, G. (2011). Electronic detectors for electron microscopy. *Quarterly reviews of biophysics*, 44(3):357–90.

[Fernández et al., 2006] Fernández, J. J., Li, S., and Crowther, R. A. (2006). CTF determination and correction in electron cryotomography. *Ultramicroscopy*, 106(7):587–96.

[Fernández et al., 1997] Fernández, J.-J., Sanjurjo, J., and Carazo, J.-M. (1997). A spectral estimation approach to contrast transfer function detection in electron microscopy. *Ultramicroscopy*, 68(4):267–295.

[Fessler and Sutton, 2003] Fessler, J. a. and Sutton, B. (2003). Nonuniform fast fourier transforms using min-max interpolation. *IEEE Transactions on Signal Processing*, 51(2):560–574.

[Förster et al., 2003] Förster, A., Whitby, F. G., and Hill, C. P. (2003). The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. *The EMBO journal*, 22(17):4356–64.

[Förster, 2005] Förster, F. (2005). *Quantitative Analyse von Kryoelektronentomogrammen mittels Korrelationsmethoden*. PhD thesis, Technische Universität München.

[Förster et al., 2010] Förster, F., Han, B.-G., and Beck, M. (2010). Visual proteomics. *Methods in enzymology*, 483:215–43.

[Förster and Hegerl, 2007] Förster, F. and Hegerl, R. (2007). Structure determination in situ by averaging of tomograms. *Methods in cell biology*, 79:741–67.

[Förster et al., 2005] Förster, F., Medalia, O., Zauberman, N., Baumeister, W., and Fass, D. (2005). Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proceedings of the National Academy of Sciences of the United States of America*, 102(13):4729–34.

[Förster et al., 2008] Förster, F., Pruggnaller, S., Seybert, A., and Frangakis, A. S. (2008). Classification of cryo-electron sub-tomograms using constrained correlation. *Journal of structural biology*, 161(3):276–86.

[Frangakis et al., 2002] Frangakis, A. S., Böhm, J., Förster, F., Nickell, S., Nicastro, D., Typke, D., Hegerl, R., and Baumeister, W. (2002). Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22):14153–8.

[Frangakis and Hegerl, 2002] Frangakis, A. S. and Hegerl, R. (2002). Segmentation of two- and three-dimensional data from electron microscopy using eigenvector analysis. *Journal of Structural Biology*, 138(1-2):105–113.

[Frank et al., 2012] Frank, G. A., Bartesaghi, A., Kuybeda, O., Borgnia, M. J., White, T. A., Sapiro, G., and Subramaniam, S. (2012). Computational separation of conformational heterogeneity using cryo-electron tomography and 3D sub-volume averaging. *Journal of structural biology*, 178(2):165–76.

[Frank, 2002] Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. *Annual review of biophysics and biomolecular structure*, 31:303–19.

[Frank, 2006a] Frank, J. (2006a). *Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. Springer, New York, 2nd edition.

[Frank, 2006b] Frank, J. (2006b). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*. Oxford University Press, New York.

[Frank and Gonzalez, 2010] Frank, J. and Gonzalez, R. L. (2010). Structure and dynamics of a processive Brownian motor: the translating ribosome. *Annual review of biochemistry*, 79:381–412.

[Frank and Penczek, 1995] Frank, J. and Penczek, P. A. (1995). On the correction of the contrast transfer function in biological electron microscopy. *Optik*, 98(3):125–129.

[Frank et al., 1996] Frank, J., Radermacher, M., Penczek, P. A., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of structural biology*, 116(1):190–9.

[Frikel, 2013] Frikel, J. (2013). Sparse regularization in limited angle tomography. *Applied and Computational Harmonic Analysis*, 34(1):117–141.

[Gerchberg, 1974] Gerchberg, R. (1974). Super-resolution through Error Energy Reduction. *Optica Acta: International Journal of Optics*, 21(9):709–720.

[Gilbert, 1972] Gilbert, P. (1972). Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117.

[Glaeser et al., 2011] Glaeser, R. M., McMullan, G., Faruqi, A. R., and Henderson, R. (2011). Images of paraffin monolayer crystals with perfect contrast: minimization of beam-induced specimen motion. *Ultramicroscopy*, 111(2):90–100.

[Gordon et al., 1970] Gordon, R., Bender, R., and Herman, G. T. (1970). Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology*, 29(3):471–481.

[Grigorieff, 1998] Grigorieff, N. (1998). Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 22 A in ice. *Journal of molecular biology*, 277(5):1033–46.

[Grigorieff, 2000] Grigorieff, N. (2000). Resolution measurement in structures derived from single particles. *Acta crystallographica. Section D, Biological crystallography*, 56(Pt 10):1270–7.

[Grünewald et al., 2003] Grünewald, K., Desai, P., Winkler, D. C., Heymann, J. B., Belnap, D. M., Baumeister, W., and Steven, A. C. (2003). Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science*, 302(5649):1396–8.

[Guerrini et al., 2011] Guerrini, N., Turchetta, R., Hoften, G. V., Henderson, R., McMullan, G., and Faruqi, A. R. (2011). A high frame rate, 16 million pixels, radiation hard CMOS sensor. *Journal of Instrumentation*, 6(03):C03003.

[Hall and Patwardhan, 2004] Hall, R. J. and Patwardhan, A. (2004). A two step approach for semi-automated particle selection from low contrast cryo-electron micrographs. *Journal of structural biology*, 145(1-2):19–28.

[Hansen, 1998] Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion.* Society for Industrial and Applied Mathematics.

[Harauz and van Heel, 1986] Harauz, G. and van Heel, M. (1986). Exact filters for general geometry three dimensional reconstruction. *Optik*, 73:146–156.

[Hart, 1968] Hart, R. G. (1968). Electron Microscopy of Unstained Biological Material: The Polytropic Montage. *Science*, 159(3822):1464–1467.

[Hawkes, 2006] Hawkes, P. (2006). The Electron Microscope as a Structure Projector. In Frank, J., editor, *Electron Tomography*, pages 83–111. Springer New York.

[Healy et al., 2003] Healy, D., Rockmore, D., Kostelec, P., and Moore, S. (2003). FFTs for the 2-Sphere Improvements and Variations. *Journal of Fourier Analysis and Applications*, 9(4):341–385.

[Hegerl and Hoppe, 1976] Hegerl, R. and Hoppe, W. (1976). Influence of electron noise on three-dimensional image reconstruction. *Zeitschrift Naturforschung*, 31a:1717–1721.

[Henderson, 2004] Henderson, R. (2004). Realizing the potential of electron cryomicroscopy. *Quarterly reviews of biophysics*, 37(1):3–13.

[Heumann et al., 2011] Heumann, J. M., Hoenger, A., and Mastronarde, D. N. (2011). Clustering and variance maps for cryo-electron tomography using wedge-masked differences. *Journal of structural biology*, 175(3):288–99.

[Heymann and Belnap, 2007] Heymann, J. B. and Belnap, D. M. (2007). Bsoft: image processing and molecular modeling for electron microscopy. *Journal of structural biology*, 157(1):3–18.

[Hobson, 1931] Hobson, E. W. (1931). *The Theory of Spherical and Ellipsoidal Harmonics*. Cambridge University Press.

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

[Hrabe et al., 2012] Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L. K., Mangold, A.-V., and Förster, F. (2012). PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. *Journal of structural biology*, 178(2):177–88.

[Hrabe and Förster, 2011] Hrabe, T. and Förster, F. (2011). Structure Determination by Single Particle Tomography. *Encyclopedia of Life Sciences*.

[Huang and Penczek, 2004] Huang, Z. and Penczek, P. A. (2004). Application of template matching technique to particle detection in electron micrographs. *Journal of structural biology*, 145(1-2):29–40.

[Huhle et al., 2009] Huhle, B., Schairer, T., and Strasser, W. (2009). Normalized Cross-Correlation using SOFT. In *Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on*, pages 82–86. IEEE.

[Jackson et al., 1991] Jackson, J. I., Meyer, C. H., Nishimura, D. G., and Macovski, A. (1991). Selection of a convolution function for Fourier inversion using gridding [computerised tomography application]. *IEEE transactions on medical imaging*, 10(3):473–8.

[Kanungo et al., 2004] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2004). A local search approximation algorithm for k-means clustering. *Computational Geometry*, 28(2-3):89–112.

[Kazhdan et al., 2003] Kazhdan, M., Funkhouser, T., and Rusinkiewicz, S. (2003). Rotation invariant spherical harmonic representation of 3D shape descriptors. In *Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 156–164. Eurographics Association.

[Keiner et al., 2009] Keiner, J., Kunis, S., and Potts, D. (2009). Using NFFT 3 - a software library for various nonequispaced fast Fourier transforms. *ACM Transactions on Mathematical Software*, V:1–23.

[Knopp et al., 2007] Knopp, T., Kunis, S., and Potts, D. (2007). A note on the iterative MRI reconstruction from nonuniform k-space data. *International journal of biomedical imaging*, 2007:24727.

[Korinek et al., 2011] Korinek, A., Beck, F., Baumeister, W., Nickell, S., and Plitzko, J. M. (2011). Computer controlled cryo-electron microscopy–TOM a software package for high-throughput applications. *Journal of structural biology*, 175(3):394–405.

[Kostelec, 2008] Kostelec, P. J. (2008). FFTs on the Rotation Group. *Journal of Fourier Analysis and Applications*, 14(2):145–179.

[Koster et al., 1992] Koster, A. J., Chen, H., Sedat, J. W., and Agard, D. A. (1992). Automated microscopy for electron tomography. *Ultramicroscopy*, 46(1-4):207–27.

[Kovacs and Wriggers, 2002] Kovacs, J. A. and Wriggers, W. (2002). Fast rotational matching. *Acta crystallographica. Section D, Biological crystallography*, 58(Pt 8):1282–6.

[Kremer et al., 1996] Kremer, J. R., Mastronarde, D. N., and McIntosh, J. R. (1996). Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology*, 116(1):71–6.

[Krivanek et al., 1995] Krivanek, O., Friedman, S., Gubbens, A., and Kraus, B. (1995). An imaging filter for biological applications. *Ultramicroscopy*, 59(1-4):267–282.

[Kudo et al., 2008] Kudo, H., Courdurier, M., Noo, F., and Defrise, M. (2008). Tiny a priori knowledge solves the interior problem in computed tomography. *Physics in medicine and biology*, 53(9):2207–31.

[Lander et al., 2012] Lander, G. C., Saibil, H. R., and Nogales, E. (2012). Go hybrid: EM, crystallography, and beyond. *Current opinion in structural biology*, 22(5):627–35.

[Lawrence et al., 2006] Lawrence, A., Bouwer, J. C., Perkins, G., and Ellisman, M. H. (2006). Transform-based backprojection for volume reconstruction of large format electron microscope tilt series. *Journal of structural biology*, 154(2):144–67.

[Lucić et al., 2005] Lucić, V., Förster, F., and Baumeister, W. (2005). Structural studies by electron tomography: from cells to molecules. *Annual review of biochemistry*, 74:833–65.

[Ludtke et al., 1999] Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of structural biology*, 128(1):82–97.

[Makadia and Daniilidis, 2006] Makadia, A. and Daniilidis, K. (2006). Rotation recovery from spherical images without correspondences. *IEEE transactions on pattern analysis and machine intelligence*, 28(7):1170–5.

[Mallick et al., 2004] Mallick, S. P., Zhu, Y., and Kriegman, D. (2004). Detecting particles in cryo-EM micrographs using learned features. *Journal of structural biology*, 145(1-2):52–62.

[Marabini et al., 1998] Marabini, R., Herman, G. T., and Carazo, J. M. (1998). 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy*, 72(1-2):53–65.

[Martinez-Sanchez et al., 2011] Martinez-Sanchez, A., Garcia, I., and Fernandez, J.-J. (2011). A differential structure approach to membrane segmentation in electron tomography. *Journal of structural biology*, 175(3):372–83.

[Martinez-Sanchez et al., 2013] Martinez-Sanchez, A., Garcia, I., and Fernandez, J.-J. (2013). A ridge-based framework for segmentation of 3D electron microscopy datasets. *Journal of structural biology*, 181(1):61–70.

[Mastronarde, 1997] Mastronarde, D. N. (1997). Dual-axis tomography: an approach with alignment methods that preserve resolution. *Journal of structural biology*, 120(3):343–52.

[Mastronarde, 2005] Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of structural biology*, 152(1):36–51.

[Matej et al., 2004] Matej, S., Fessler, J., and Kazantsev, I. (2004). Iterative Tomographic Image Reconstruction Using Fourier-Based Forward and Back-Projectors. *IEEE transactions on medical imaging*, 23(4):401–412.

[McEwen et al., 1995] McEwen, B. F., Downing, K. H., and Glaeser, R. M. (1995). The relevance of dose-fractionation in tomography of radiation-sensitive specimens. *Ultramicroscopy*, 60(3):357–73.

[McMullan et al., 2009] McMullan, G., Chen, S., Henderson, R., and Faruqi, A. R. (2009). Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy*, 109(9):1126–43.

[Melnikov et al., 2012] Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nature structural & molecular biology*, 19(6):560–7.

[Miao et al., 2005] Miao, J., Förster, F., and Levi, O. (2005). Equally sloped tomography with oversampling reconstruction. *Physical Review B*, 72(5):052103.

[Moussavi et al., 2010] Moussavi, F., Heitz, G., Amat, F., Comolli, L. R., Koller, D., and Horowitz, M. (2010). 3D segmentation of cell boundaries from whole cell cryogenic electron tomography volumes. *Journal of structural biology*, 170(1):134–45.

[Natterer, 2001] Natterer, F. (2001). *The mathematics of computerized tomography*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[Nicastro et al., 2006] Nicastro, D., Schwartz, C., Pierson, J., Gaudette, R., Porter, M. E., and McIntosh, J. R. (2006). The molecular architecture of axonemes revealed by cryoelectron tomography. *Science*, 313(5789):944–8.

[Nickell et al., 2005] Nickell, S., Förster, F., Linaroudis, A., Net, W. D., Beck, F., Hegerl, R., Baumeister, W., and Plitzko, J. M. (2005). TOM software toolbox: acquisition and analysis for electron tomography. *Journal of structural biology*, 149(3):227–34.

[Nickell et al., 2006] Nickell, S., Kofler, C., Leis, A., and Baumeister, W. (2006). A visual approach to proteomics. *Nature Reviews Molecular Cell Biology*, 7(3):225–230.

[Ogura and Sato, 2004] Ogura, T. and Sato, C. (2004). Automatic particle pickup method using a neural network has high accuracy by applying an initial weight derived from eigenimages: a new reference free method for single-particle analysis. *Journal of structural biology*, 145(1-2):63–75.

[Orlova and Saibil, 2011] Orlova, E. V. and Saibil, H. R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. *Chemical reviews*, 111(12):7710–48.

[Ortiz et al., 2006] Ortiz, J. O., Förster, F., Kürner, J., Linaroudis, A. a., and Baumeister, W. (2006). Mapping 70S ribosomes in intact cells by cryoelectron tomography and pattern recognition. *Journal of structural biology*, 156(2):334–41.

[Papoulis, 1975] Papoulis, A. (1975). A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and Systems*, 22(9):735–742.

[Penczek et al., 1995] Penczek, P., Marko, M., Buttle, K., and Frank, J. (1995). Double-tilt electron tomography. *Ultramicroscopy*, 60(3):393–410.

[Penczek, 2010] Penczek, P. A. (2010). Resolution measures in molecular electron microscopy. In *Methods in Enzymology*, volume 482 of *Methods in Enzymology*, pages 73–100. Elsevier.

[Penczek et al., 1992] Penczek, P. A., Radermacher, M., and Frank, J. (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy*, 40(1):33–53.

[Penczek et al., 2004] Penczek, P. A., Renka, R., and Schomberg, H. (2004). Gridding-based direct Fourier inversion of the three-dimensional ray transform. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 21(4):499–509.

[Pfeffer et al., 2012] Pfeffer, S., Brandt, F., Hrabe, T., Lang, S., Eibauer, M., Zimmermann, R., and Förster, F. (2012). Structure and 3D Arrangement of Endoplasmic Reticulum Membrane-Associated Ribosomes. *Structure*, 20(9):1508–18.

[Pfeffer et al., 2014] Pfeffer, S., Dudek, J., Gogala, M., Schorr, S., Linxweiler, J., Lang, S., Becker, T., Beckmann, R., Zimmermann, R., and Förster, F. (2014). Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. *Nature communications*, 5:3072.

[Pipe and Menon, 1999] Pipe, J. G. and Menon, P. (1999). Sampling density compensation in MRI: rationale and an iterative numerical solution. *Magnetic resonance in medicine*, 41(1):179–86.

[Potts et al., 2001] Potts, D., Steidl, G., and Tasche, M. (2001). Fast Fourier transforms for nonequispaced data: A tutorial. *Modern sampling theory*.

[Qu et al., 2005] Qu, P., Zhong, K., Zhang, B., Wang, J., and Shen, G. X. (2005). Convergence behavior of iterative SENSE reconstruction with non-Cartesian trajectories. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 54(4):1040–5.

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

[Radermacher et al., 1986] Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1986). A new 3-D reconstruction scheme applied to the 50S ribosomal subunit of E. coli. *Journal of microscopy*, 141(Pt 1):RP1–2.

[Radon, 1917] Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Math. Phys. Klasse*, 69:262–277.

[Ramachandran, 1971] Ramachandran, G. N. (1971). Three-Dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions Instead of Fourier Transforms. *Proceedings of the National Academy of Sciences*, 68(9):2236–2240.

[Reimer and Kohl, 2008] Reimer, L. and Kohl, H. (2008). *Transmission Electron Microscopy: Physics of Image Formation*. Springer Series in Optical Sciences, Springer, 5th edition.

[Rigort et al., 2012] Rigort, A., Günther, D., Hegerl, R., Baum, D., Weber, B., Prohaska, S., Medalia, O., Baumeister, W., and Hege, H.-C. (2012). Automated segmentation of electron tomograms for a quantitative description of actin filament networks. *Journal of structural biology*, 177(1):135–44.

[Roseman, 2003] Roseman, A. M. (2003). Particle finding in electron micrographs using a fast local correlation algorithm. *Ultramicroscopy*, 94(3-4):225–36.

[Rosenthal and Henderson, 2003] Rosenthal, P. B. and Henderson, R. (2003). Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology*, 333(4):721–745.

[Saad, 2003] Saad, Y. (2003). *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition.

[Sandberg and Brega, 2007] Sandberg, K. and Brega, M. (2007). Segmentation of thin structures in electron micrographs using orientation fields. *Journal of structural biology*, 157(2):403–15.

[Sandberg et al., 2003] Sandberg, K., Mastronarde, D. N., and Beylkin, G. (2003). A fast reconstruction algorithm for electron microscope tomography. *Journal of structural biology*, 144(1-2):61–72.

[Saxton and Baumeister, 1982] Saxton, W. O. and Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of Microscopy*, 127(2):127–138.

[Scheres and Chen, 2012] Scheres, S. H. W. and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature methods*, 9(9):853–4.

[Scheres et al., 2009] Scheres, S. H. W., Melero, R., Valle, M., and Carazo, J.-M. (2009). Averaging of electron subtomograms and random conical tilt reconstructions through likelihood optimization. *Structure*, 17(12):1563–72.

[Schmid and Booth, 2008] Schmid, M. F. and Booth, C. R. (2008). Methods for aligning and for averaging 3D volumes with missing data. *Journal of structural biology*, 161(3):243–8.

[Schur et al., 2013] Schur, F. K. M., Hagen, W. J. H., de Marco, A., and Briggs, J. A. G. (2013). Determination of protein structure at 8.5Åresolution using cryo-electron tomography and sub-tomogram averaging. *Journal of structural biology*, 184(3):394–400.

[Schweikert, 2004] Schweikert, G. (2004). Quantitativer vergleich der strahlschädigung biologischer proben im transmissions-elektronenmikroskop bei stickstoff und helium temperatur. Master's thesis, Technische Universität München.

[Sorzano et al., 2004] Sorzano, C. O. S., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H. W., Carazo, J. M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *Journal of structural biology*, 148(2):194–204.

[Sorzano et al., 2009] Sorzano, C. O. S., Messaoudi, C., Eibauer, M., Bilbao-Castro, J. R., Hegerl, R., Nickell, S., Marco, S., and Carazo, J. M. (2009). Marker-free image registration of electron tomography tilt-series. *BMC bioinformatics*, 10(1):124.

[Stewart and Grigorieff, 2004] Stewart, A. and Grigorieff, N. (2004). Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy*, 102(1):67–84.

[Stölken et al., 2011] Stölken, M., Beck, F., Haller, T., Hegerl, R., Gutsche, I., Carazo, J.-M., Baumeister, W., Scheres, S. H. W., and Nickell, S. (2011). Maximum likelihood based classification of electron tomographic data. *Journal of structural biology*, 173(1):77–85.

[Subramaniam, 2006] Subramaniam, S. (2006). The SIV surface spike imaged by electron tomography: one leg or three? *PLoS pathogens*, 2(8):e91.

[Suloway et al., 2009] Suloway, C., Shi, J., Cheng, A., Pulokas, J., Carragher, B., Potter, C. S., Zheng, S. Q., Agard, D. A., and Jensen, G. J. (2009). Fully automated, sequential tilt-series acquisition with Leginon. *Journal of structural biology*, 167(1):11–8.

[Tang et al., 2007] Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007). EMAN2: an extensible image processing suite for electron microscopy. *Journal of structural biology*, 157(1):38–46.

[Turin, 1960] Turin, G. (1960). An introduction to matched filters. *IEEE Transactions on Information Theory*, 6(3):311–329.

[van Heel et al., 2000] van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., and Patwardhan, A. (2000). Single-particle electron cryo-microscopy: towards atomic resolution. *Quarterly reviews of biophysics*, 33(4):307–69.

[van Heel and Schatz, 2005] van Heel, M. and Schatz, M. (2005). Fourier shell correlation threshold criteria. *Journal of structural biology*, 151(3):250–62.

[Verhaeghe et al., 2008] Verhaeghe, J., Van de Ville, D., Khalidov, I., D'Asseler, Y., Lemahieu, I., and Unser, M. (2008). Dynamic PET reconstruction using wavelet regularization with adapted basis functions. *IEEE transactions on medical imaging*, 27(7):943–59.

[Volkmann, 2002] Volkmann, N. (2002). A novel three-dimensional variant of the watershed transform for segmentation of electron density maps. *Journal of structural biology*, 138(1-2):123–9.

[Volkmann, 2004] Volkmann, N. (2004). An approach to automated particle picking from electron micrographs based on reduced representation templates. *Journal of Structural Biology*, 145(1-2):152–156.

[Volkmann, 2010] Volkmann, N. (2010). Methods for segmentation and interpretation of electron tomographic reconstructions. *Methods in enzymology*, 483:31–46.

[Walz et al., 1997] Walz, J., Typke, D., Nitsch, M., Koster, A. J., Hegerl, R., and Baumeister, W. (1997). Electron Tomography of Single Ice-Embedded Macromolecules: Three-Dimensional Alignment and Classification. *Journal of Structural Biology*, 120(3):387–395.

[Wan et al., 2011] Wan, X., Zhang, F., Chu, Q., Zhang, K., Sun, F., Yuan, B., and Liu, Z. (2011). Three-dimensional reconstruction using an adaptive simultaneous algebraic reconstruction technique in electron tomography. *Journal of structural biology*, 175(3):277–87.

[Wang et al., 2008] Wang, Y., Yang, J., Yin, W., and Zhang, Y. (2008). A New Alternating Minimization Algorithm for Total Variation Image Reconstruction. *SIAM Journal on Imaging Sciences*, 1(3):248–272.

[Weber et al., 2012] Weber, B., Greenan, G., Prohaska, S., Baum, D., Hege, H.-C., Müller-Reichert, T., Hyman, A. A., and Verbavatz, J.-M. (2012). Automated tracing of microtubules in electron tomograms of plastic embedded samples of Caenorhabditis elegans embryos. *Journal of structural biology*, 178(2):129–38.

[Whitaker and Elangovan, 2002] Whitaker, R. T. and Elangovan, V. (2002). A direct approach to estimating surfaces in tomographic data. *Medical image analysis*, 6(3):235–49.

[Wilson and Doudna Cate, 2012] Wilson, D. N. and Doudna Cate, J. H. (2012). The structure and function of the eukaryotic ribosome. *Cold Spring Harbor perspectives in biology*, 4(5).

[Winkler, 2007] Winkler, H. (2007). 3D reconstruction and processing of volumetric data in cryo-electron tomography. *Journal of structural biology*, 157(1):126–37.

[Winkler et al., 2009] Winkler, H., Zhu, P., Liu, J., Ye, F., Roux, K. H., and Taylor, K. A. (2009). Tomographic subvolume alignment and subvolume classification applied to myosin V and SIV envelope spikes. *Journal of Structural Biology*, 165(2):64–77.

[Witt et al., 2006] Witt, S., Kwon, Y. D., Sharon, M., Felderer, K., Beuttler, M., Robinson, C. V., Baumeister, W., and Jap, B. K. (2006). Proteasome assembly triggers a switch required for active-site maturation. *Structure*, 14(7):1179–88.

[Wong et al., 2004] Wong, H. C., Chen, J., Mouche, F., Rouiller, I., and Bern, M. (2004). Model-based particle picking for cryo-electron microscopy. *Journal of structural biology*, 145(1-2):157–67.

[Wu, 1983] Wu, C. F. J. (1983). On the Convergence Properties of the EM Algorithm. *The Annals of Statistics*, 11(1):95–103.

[Xu et al., 2012] Xu, M., Beck, M., and Alber, F. (2012). High-throughput subtomogram alignment and classification by Fourier space constrained fast volumetric matching. *Journal of structural biology*, 178(2):152–164.

[Yu and Fessler, 2002] Yu, D. F. and Fessler, J. A. (2002). Edge-preserving tomographic reconstruction with nonlocal regularization. *IEEE transactions on medical imaging*, 21(2):159–73.

[Yu et al., 2010] Yu, L., Snapp, R. R., Ruiz, T., and Radermacher, M. (2010). Probabilistic principal component analysis with expectation maximization (PPCA-EM) facilitates volume classification and estimates the missing data. *Journal of structural biology*, 171(1):18–30.

[Zanetti et al., 2009] Zanetti, G., Riches, J. D., Fuller, S. D., and Briggs, J. a. G. (2009). Contrast transfer function correction applied to cryo-electron tomography and sub-tomogram averaging. *Journal of structural biology*, 168(2):305–12.

[Zhang et al., 2008] Zhang, W., Kimmel, M., Spahn, C. M. T., and Penczek, P. A. (2008). Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis. *Structure*, 16(12):1770–6.

[Zheng et al., 2007] Zheng, S. Q., Keszthelyi, B., Branlund, E., Lyle, J. M., Braunfeld, M. B., Sedat, J. W., and Agard, D. A. (2007). UCSF tomography: an integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. *Journal of structural biology*, 157(1):138–47.

[Zheng et al., 2010] Zheng, S. Q., Sedat, J. W., and Agard, D. A. (2010). Automated data collection for electron microscopic tomography. *Methods in enzymology*, 481:283–315.

[Zhu et al., 2004] Zhu, Y., Carragher, B., and Glaeser, R. (2004). Automatic particle selection: results of a comparative study. *Journal of structural*, 145(1-2):3–14.

[Zhu et al., 2003] Zhu, Y., Carragher, B., Mouche, F., and Potter, C. S. (2003). Automatic particle detection through efficient Hough transforms. *IEEE transactions on medical imaging*, 22(9):1053–62.