# ALIGNMENT OF CRYO-ELECTRON TOMOGRAPHY IMAGES USING MARKOV RANDOM FIELDS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Fernando Amat Gil

May 2010

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Mark A. Horowitz) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Daphne Koller)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Kenneth H. Downing)

Approved for the University Committee on Graduate Studies.

# Abstract

Cryo-Electron tomography (CET) is the only imaging technology capable of visualizing the 3D organization of intact bacterial whole cells at nanometer resolution *in situ*. However, quantitative image analysis of CET datasets is extremely challenging due to very low signal to noise ratio (well below 0dB), missing data and heterogeneity of biological structures.

In this thesis, we present a probabilistic framework to align CET images in order to improve resolution and create structural models of different biological structures. The alignment problem of 2D and 3D CET images is cast as a Markov Random Field (MRF), where each node in the graph represents a landmark in the image. We connect pairs of nodes based on local spatial correlations and we find the "best" correspondence between the two graphs. In this correspondence problem, the "best" solution maximizes the probability score in the MRF. This probability is the product of singleton potentials that measure image similarity between nodes and the pairwise potentials that measure deformations between edges. Well-known approximate inference algorithms such as Loopy Belief Propagation (LBP) are used to obtain the "best" solution.

We present results in two specific applications: automatic alignment of tilt series using fiducial markers and subtomogram alignment. In the first case we present RAPTOR, which is being used in several labs to enable real high-throughput tomography. In the second case our approach is able to reach the contrast transfer function limit in low $SNR$ samples from whole cells as well as revealing atomic resolution details invisible to the naked eye through nanogold labeling.

# Acknowledgments

Writing this page is the last step of a journey that started six years ago. It is a great feeling to reach a conclusion, but it is even better knowing that this is just the beginning of my career. The experiences I have had during my PhD at Stanford will follow me wherever I go, and all of those experiences have names and faces associated with them that I would like to thank here.

First and foremost, I would like to thank my principle advisor Mark Horowitz. He has been a real mentor in all aspects of a PhD, from showing which questions are important to answer to how to exchange knowledge better. Through his example, he has shown me that it is possible to seek knowledge forever. No matter how many things you know, there are still many things to be discovered. Finally, I am grateful to him for giving me the opportunity to apply my engineering and mathematical background in a new field for me and for opening the door to a great group of collaborators from many different disciplines, which dramatically enriched my experience at Stanford.

One of those collaborators is my co-advisor Daphne Koller, who I would like to thank for guiding me through the rich subject of graphical models and the complicated world of research. Daphne's advice is always sharp and stimulating, pushing me to elevate my own standards. She set an example for me with her brilliance, high standards and drive for excellence.

Kenneth H. Downing is another of those collaborators and he is responsible for everything I know about electron microscopy. I would like to thank him for opening the doors of Donner Lab at LBL for me and for showing me the nuts and bolts of the electron

microscope over informal discussion sessions. I am also grateful that he kindly accepted to be in my reading committee for this thesis.

I would also like to acknowledge my office mate and great friend Farshid Moussavi. Most of the work presented here is shared with him through many hours at the office and many hours of joy and camaraderie outside the office. I really admire his determination and his character. I would also like to thank Luis R. Comolli for never ending discussions about microscopy, Fourier and any other subject not strictly related to research. I consider Luis a true artist of the microscope and without his talent and all his images this thesis would not be possible. I would also like to thank Lucy Shapiro and Harley McAdams for all their support and kindness during all these years. The list of collaborators could go on forever, but I definitely do not want to forget to thank Albert Lawerence at UCSD, Grant Jensen at Caltech, Puey Ounjay at LBL and John Smit at British Columbia for working together in different projects presented in this thesis. Finally, thanks to all the VLSI research group for support and fruitful discussions, and to Teresa and Rafael for making visits to administration offices fun.

In my experience, the PhD journey goes beyond research and it affects all aspects of life. Without the personal friendship of many people this would have not been possible. I am thinking of friends and family such as Mario, Argyris, Luciana, Sewoong, Yorgos and many others that made these years in San Francisco unforgettable. However, I would like to dedicate this thesis to Margie Marx. For me, she represents the beginning and the end of my years at Stanford, and she taught me many valuable lessons that can not be learned in any university.

Last but not least, I would like to dedicate this thesis to the three most important people in my life: my fiancee Hillary, my dad Nicolas and my mom Josefina. Hillary has given me love and support throughout these years and she is always able to light up a gray day with a single smile. I am looking forward to sharing the rest of my life with her. My parents taught me since I was little that education will open doors to fanstatic places... they were so right.

# Contents

# List of Tables

# List of Figures

## List of Symbols

| | |
|---|---|
| $Cc.$ | Caulobacter crescentus |
| $CET$ | Cryo-Electron Tomography |
| $CCD$ | Charge Couple Device |
| $MRF$ | Markov Random Field |
| $ET$ | Electron Tomography |
| $EM$ | Electron Microscopy |
| $SNR$ | Signal-to-Noise Ratio |
| $2D$ | Two-dimensional |
| $3D$ | Three-dimensional |
| $NCC$ | Normalized Cross-Correlation |
| $TCCC$ | Thresholded Constrained Cross-Correlation |
| $ML$ | Maximum Likelihod |
| $FSC$ | Fourier Shell Correlation |
| $OM$ | Outer Membrane |
| $IM$ | Inner Membrane |
| $CTF$ | Contrast Transfer Function |
| $TEM$ | Transmission Electron Microscope |
| $r_j$ | 3D coordinates of j-th marker |
| $p_{ij}$ | 2D coordinates of j-th marker in i-th image |
| $p'_{ij}$ | 2D coordinates of markers estimated by projection model |
| $G_i$ | Projection matrix for i-th image |
| $n_P$ | Number of projections in tilt series |
| $n_T$ | Number of trajectories |
| $\psi_i$ | Singleton potentials |
| $\psi_{ij}$ | Pairwise potentials |
| $f^k$ | k-th subtomogram |
| $F^k$ | k-th subtomogram in Fourier space |

| | |
|---|---|
| $\langle , \rangle$ | Inner product |
| $B$ | Number of subtomograms |
| $M^k$ | Missing wedge mask for k-th subtomogram |

# Chapter 1

# Introduction

It is a cliche that a picture is worth a thousand words but in many fields data visualization is critical. In fact, imaging in life sciences is one of the main tools for new discoveries, irrespective of wether one is studying human beings or prokaryotic cells. Through improved technology methods, the capabilities of different imaging modalities and the amount of acquired data is increasing very rapidly. Our current challenge is to develop new algorithms to efficiently transform all this data into knowledge. In particular, the accumulation of data observing similar structures opens the possibility to reveal patterns through statistical methods that otherwise would be impossible to extract manually.

In the past 15 years, cryo-electron tomography ($CET$) has proved itself as a critical $3D$ imaging technology for structural biologists due to its capability of imaging whole cells without staining artifacts at resolutions of few nanometers [Bau02]. This resolution provides crucial information by observing the spatial interaction of different macromolecules within cells. However, typical $CET$ datasets have very low $SNR$ and are very large, which demands the extended use of image processing techniques to extract useful information.

Graphical models, and in particular Markov Random Fields ($MRF$), provide a natural framework to successfully tackle problems with large amounts of uncertainty and complexity. The key point is to decompose the problem into simpler subproblems and "exchange"

possible outcomes between the different modules to reach a coherent solution. In this thesis, we present several applications of inference in $MRF$ applied to $CET$ datasets in order to align $2D$ and $3D$ images with very low $SNR$ and missing information. In particular, we use the probabilistic framework to model local spatial constraints between different biological structures as a guidance to increase the accuracy of the alignment.

## 1.1 Imaging in life science

Fig. 1.1 shows a schematic with the attainable resolution for different imaging sources. At one end of the spectrum, we have light microscopy, which is crucial to study bacterial cells *in vivo*. It offers a few hundreds nanometer resolution and the possibility of using fluorescence labeling to localize and dynamically track proteins of interest through the cell cycle. Using these techniques, it has been shown [SLR97] that bacterial cells are not just "bags of enzymes", but they are highly organized at the protein level. For example, many proteins are found at specific points, such as poles or division planes, and changing this location can alter different cell functions. Even with the development of new super-resolution light microscopy techniques like photo-activated localization microscopy (PALM) [BPS$^+$06] or stochastic optical reconstruction microscopy (STORM) [RBZ06], images can not resolve structures smaller than few tens of nanometers.

At the other end of the bioimaging spectrum resolution we encounter high-resolution structural techniques, such as X-ray crystallography, nuclear magnetic resonance spectroscopy or single particle reconstructions. These methods can achieve atomic resolution in $3D$ and solve primary and secondary structures of small purified molecules, which can be used to fit a $3D$ atomic-model of the amino acid sequence to gain biochemical insight. Moreover, research efforts have allowed relatively simple high-throughput data collection and reconstruction. However, in order to achieve such high resolution information molecular complexes need to be isolated *in vitro*, be relatively small and have a homogeneous conformation.

Looking again at Fig. 1.1, we observe that there is an "imaging gap" between the

resolution of the two sources described above. Electron tomography covers this gap, which makes it the highest resolution imaging technology available capable of obtaining unique structures of individual cells, organelles, flexible macromolecules and certain viruses. In particular, $CET$ is capable of obtaining $3D$ reconstructions of whole bacterial cells at "molecular resolution" (approx. 4-8nm) *in situ* [MNM05, LFB05]. As mentioned above, these capabilities are critical to understand the function and complex spatial relationships of organelles, microtubules, vesicles, ribosomes, and other large structures within cells [LBB$^+$08, MVMH04, SD06, MLJ06]. However, $CET$ images present several challenges such as low $SNR$ (well below 0dB), missing information, heterogeneity and large data sizes, which make acquisition, reconstruction and data analysis far from straight forward (Fig. 1.2). New statistical image processing methods are needed to efficiently process and extract higher resolution structural information from hundreds of tomograms.

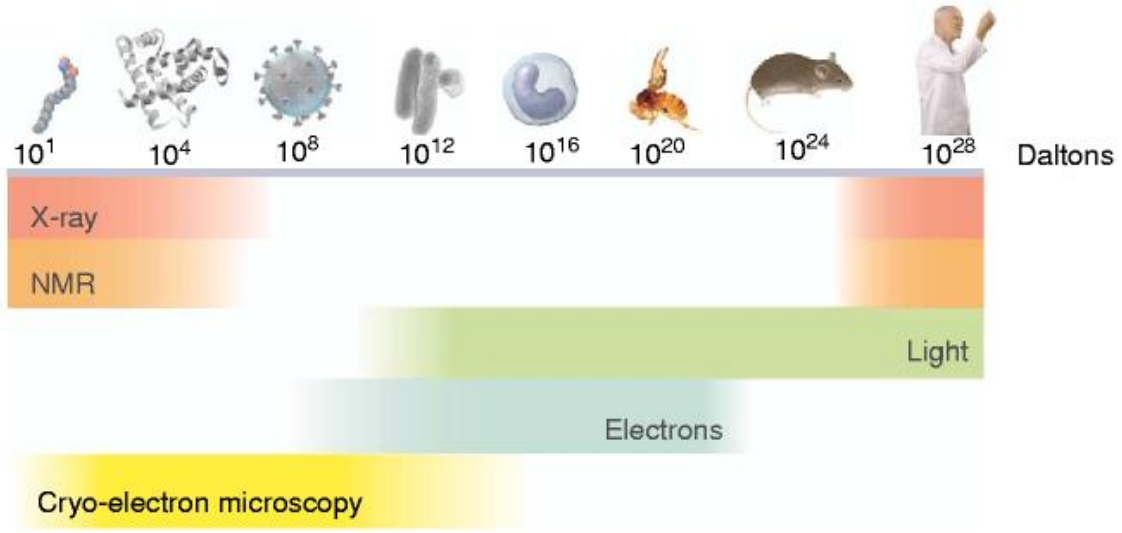


**Figure 1.1:** Bioimaging spectrum figure from [Sub06] representing the different resolutions attainable with different imaging techniques. Cryo-electron microscopy bridges the gap between *in vivo* low resolution imaging and *in vitro* high resolution.

## 1.2 Statistical image processing in CET

The goal of image processing methods for $CET$ is two fold: first, to reach a true high-throughput pipeline where hundreds of tomograms can be acquired per month, expanding the kind of experiments and questions that can be answered by $CET$. And second, to efficiently combine these large amounts of data in order to resolve biological structures at higher resolutions than the ones provided by a single tomogram. For example, a high-throughput pipeline would enable the reconstruction of mammalian cells at nanometer resolution by stitching hundreds of tomograms together.

Fig. 1.2 shows a typical $CET$ dataset of whole cells that will be used in this thesis. The low SNR and the fact that the tomogram contains missing data makes the $CET$ image statistics very different from natural images or even medical images from MRI, PET and ultrasound systems. It is easy to see that standard image processing techniques to extract information from multiple $CET$ images, such as edge detection and feature descriptors, can not be applied directly using off-the-shelf procedures. Moreover, standard techniques to reduce noise such as denoising and downsampling can only be applied in a first coarse step, since the goal in structural biology is always to achieve as high structural resolution as possible, requiring high precision in any image analysis task. Therefore, $CET$ requires new statistical image processing methods that can handle these constraints to help extracting biological knowledge from $CET$ images.

## 1.3 Thesis outline and contributions

In this thesis, we present a probabilistic framework based on $MRF$ to improve and automate data analysis of $CET$ images. We apply this framework to two critical problems in $CET$: first, to align a set of $2D$ images to create a $3D$ reconstruction of the area of interest. We will see how solving this problem removes one of the bottlenecks in the tomographic acquisition pipeline, opening the possibility to acquire hundreds of tomograms per month. Second, we use the same framework to align many of these $3D$ volumes to create

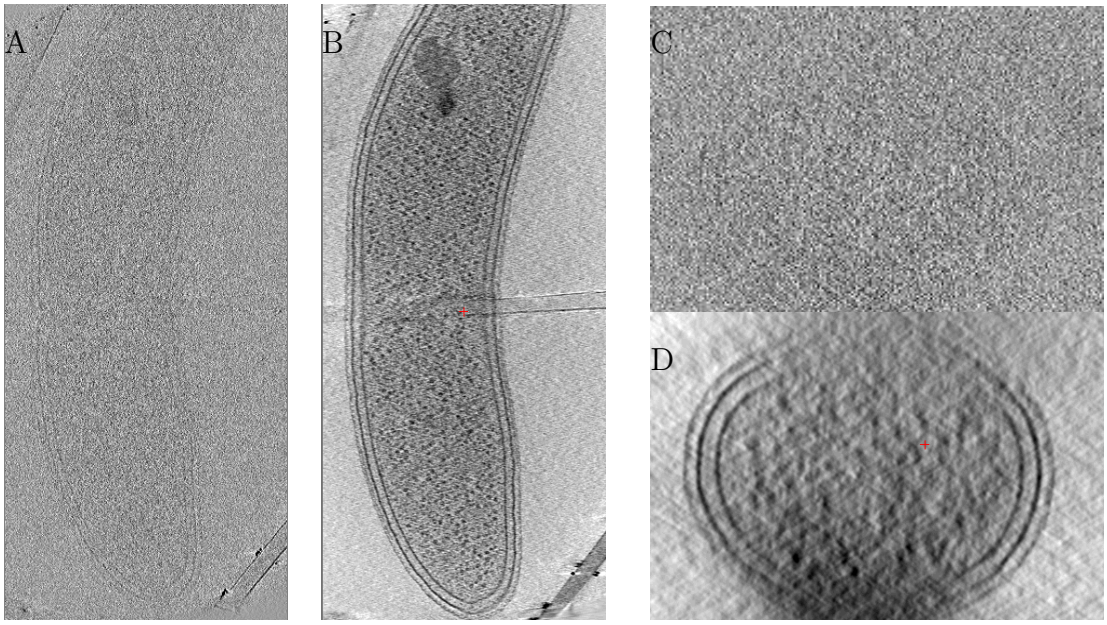


**Figure 1.2:** Slices of $3D$ volume exemplifying the main two challenges of $CET$ of whole cells: noise and missing data. (A) XY slice of *Caulobacter crescentus* tomograms without filtering to preserve high frequency information. (B) Same image as in (A) with low-pass filter. (C) XZ slice of the same tomogram without filtering. (D) Same image as in (C) with low-pass filter. Features of cell wall components on top and bottom are blurred due to the missing wedge effect.

higher resolution averaged images. In order to accomplish these tasks the probabilistic framework contains two elements that are problem specific: a metric used to compare individual image regions to determine good matches and the cost of deforming the space between these features that is allowed in the model. We not only show how to use probabilistic methods to perform these alignments for $CET$ data, but we also present a new dissimilarity score that works better for noisy images.

Chapter 3 describes in more detail the $2D$ alignment problem, and presents the fundamentals for automatic alignment of $2D$ projections of $CET$ tilt series using gold fiducial markers. $2D$ image alignment is easier than $3D$ because of fewer degrees of freedom and the existence of high-contrast point-like features as markers. Therefore, it is a perfect example to show the basic principles of the $MRF$ alignment framework for low $SNR$ images. Chapter 4 describes the second application and presents the extension of the alignment

framework to $CET$ $3D$ density volumes in order to recover higher resolution structures by averaging the aligned volumes. The alignment and averaging of thousands of volumes allows one to "see" details otherwise hidden in single tomograms. Finally, Chapter 5 uses these techniques to present a case study of the surface layer structure of *Caulobacter crescentus* using $CET$ of whole cells. Combining all the methodology explained in previous chapters we bring new biological insights and we show the strengths and limitations of the framework presented in this thesis in a real case scenario.

Before jumping to the applications, the next chapter explains the fundamentals of image formation in the electron microscope and the main steps involved in acquiring, processing and analyzing $CET$ datasets to provide the needed background. This background is important for two reasons: it provides an understanding of the data at hand (noise characteristics, missing data distribution, etc.), which is necessary to develop better algorithms, and it brings a high level perspective to frame the work described by the rest of the thesis.

# Chapter 2

# Background

Cryo-electron tomography ($CET$) is the highest resolution imaging technology available capable of obtaining unique structures of individual cells, organelles, flexible macromolecules and certain viruses. Understanding the image formation process in the electron microscope allows one to obtain better images, design new statistical image processing algorithms and identify bottlenecks in the $CET$ acquisition pipeline. Sections 2.1, 2.2 and 2.3 provide this background to familiarize the reader with the kind of images presented in this thesis.

We will see in this chapter that $CET$ creates very low $SNR$ images. This phenomenon makes some steps that are easy in other contexts, such as image alignment, non-trivial. The second part of this chapter reviews existing methodologies from different fields that are related to the ones introduced in this thesis and explains why we need to adapt them to the $CET$ image characteristics in order to improve their performance. In particular, Section 2.4 describes the formalism of Markov Random Fields (MRF), which would be used as a framework to solve the alignment problem in CET images.

## 2.1 Image formation in transmission electron microscope

Fig. 2.1 shows a schematic of the main components of a standard transmission electron microscope ($TEM$). The basic setup is very similar to a common light microscope, where we have an illumination source, a condenser lens to focus the source into the specimen and an objective and projector lens to magnify and create an image that can be recorded with a camera or viewed in a phosphor screen. However, in $TEM$ the source is a coherent beam of electrons instead of light, which accounts for many differences in the components and the image formation that will be described in this section. The reader is referred to Chapter 3 by Frank [Fra06] and Chapters 3,4 and 5 by Glaeser et al. [RG07] and Chapters 3 and 6 by Reimer and Kohl [RK08] for detailed explanations of these topics.

The illumination source consists of an *electron gun* that emits an intense beam of electrons that can be considered coherent. Intensity is required to minimize exposure times, which can reduce resolution by spatial incoherent due to movement in the sampler holder during acquisition. Temporal coherence around a nominal wavelength $\lambda$ is required to achieve high resolution through the mechanism of phase contrast, which will be explained below. In many high performance microscopes, the electron gun is a tungsten field emission gun (FEG) tip subject to strong electric fields to allow electrons to "escape" from the tip and be accelerated towards the anode, which is kept at ground potential. Another option for the electron gun is a thermal filament of lanthanum hexaboride ($LaB_6$). The filament is heated by passing a current through it and when it reaches a certain temperature the electrons escape out of the filament into free space. FEG electron guns are preferred for samples with phase contrast (explained below), since it generates a more coherent source, while $Lab_6$ filaments are preferred for samples dominated by amplitude contrast since they radiate higher intensity. In both cases, typical energies for biological applications are 100 to 300KeV, which gives a range of wavelength from 0.037 to 0.020Å.

The electron source brings two main differences with respect to light microscopy in the design of the $TEM$ optics. First, electric current lenses that generate magnetic fields (Fig. 2.2) instead of glass lenses, are used to focus the electron beam. The magnetic fields

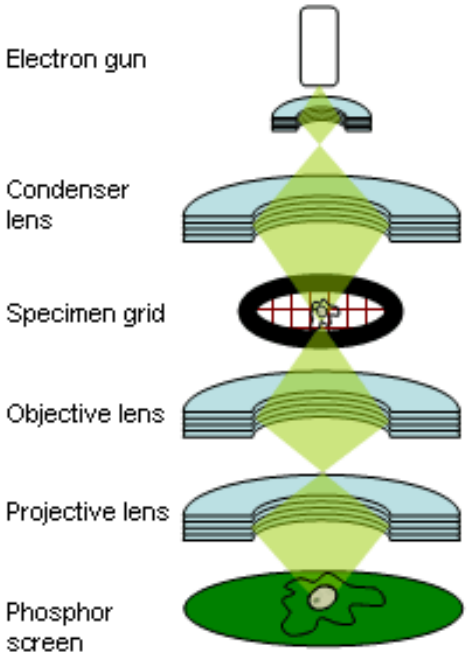


**Figure 2.1:** Schematic showing the different components of a transmission electron microscope. The basic setup is very similar to a light microscope, with the main difference that the illumination source is a coherent electron beam instead of light. This difference forces the use of electron lenses which generate magnetic fields to focus the electron beam. Figure from http://astronomy.nmsu.edu/jlevans/seminarSpr07/.

exert an orthogonal Lorentzian force[1] on the electrons altering their direction but not their energy. By changing the current in the lens, we adjust the intensity of the magnetic field, which changes the focal length. Second, high vacuum within the column is required in order to have scattering only between the specimen and the electrons.

The energy and size of electrons allows to study them as particles or as waves. Considering them as particles and the specimen under observation as a set of atoms, we can describe the possible interactions between each electron emitted by the electron gun and each atom in the sample (Fig. 2.3). First, many electrons will not pass close enough to any

[1]The force that a magnetic field $\boldsymbol{B}$ exerts over an electron with velocity $\boldsymbol{v}$ is $\boldsymbol{F} = q(\boldsymbol{v} \times \boldsymbol{B})$, where $\times$ is the cross-product operator and $q$ is the charge of the electron.

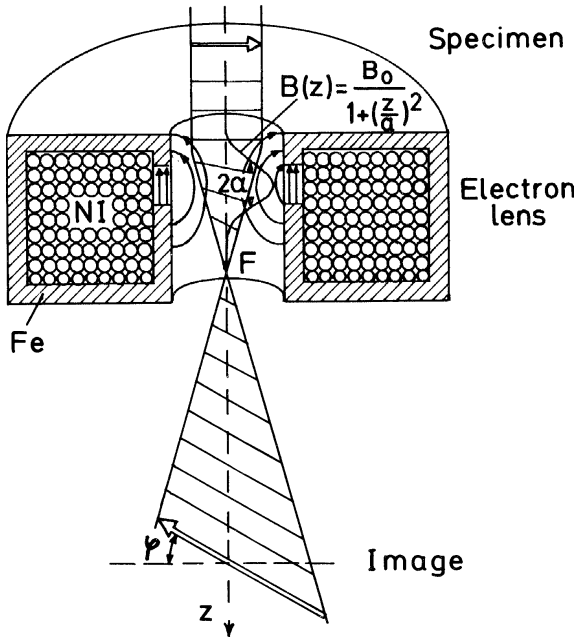


**Figure 2.2:** The Lorentz force generated by the magnetic field is normal to both the direction of the electrons and the magnetic field $\boldsymbol{B}$. An electron entering the lense with velocity $\boldsymbol{v}$ undergoes an acceleration that is everywhere normal to the local velocity vector. Therefore, $\boldsymbol{v}$ does not change in magnitude but it alters its direction, concentrating the electron beam at a distance F of the lense (focal length). In order to increase focusing power, we want a lense as thin as possible. Figure from [RK08].

atom and their trajectory will not be affected. We define these electrons as *unscattered*. Second, electrons that pass close enough to a specimen atom to be inside the electron cloud will be deflected laterally due to the attractice Coulomb force. If the energy transferred during the interaction is neglegible we define these electrons as *elastically scattered*, and *inelastically scattered* otherwise. Therefore, at a qualitative level, the output electron beam contains information related to the Coulomb potential distribution in the sample, which is an indirect measure of the $3D$ density of the sample.

When using the $TEM$ for biological specimens several extra remarks need to be considered before analyzing the image formation process. First, inelastic scattered electrons add low resolution background noise because they introduce incoherence by generating new wavelengths in the output electron wave, which reduces the contrast in the image

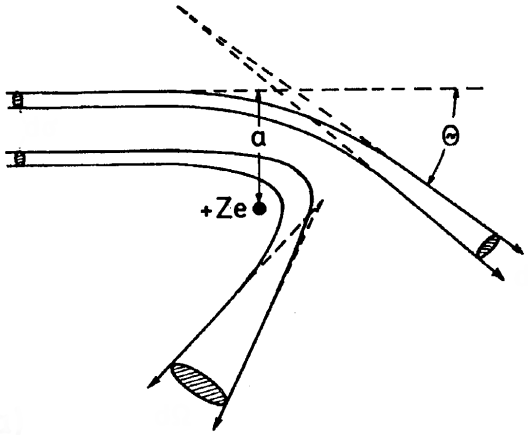


**Figure 2.3:** Elastic electron scattering produced by the interaction between the electron beam and the nucleus of an atom $(+Ze)$ present in the specimen due to the attractive Coulomb force. The electron is scattered by an angle $\theta$. Figure from [RK08].

(Fig. 2.4). Thus, current $TEM$ are equipped with an energy filter to improve the image quality. In particular, the most common modality in $CET$ is known as zero-loss imaging, since we filter out electrons which have lost energy. Second, multiple scattering events for a single electron should be minimized as much as possible since they complicate the interpretability of the images and they increase the chance of having an inelastic scattering event. Specimen thickness is directly related to the probability of having multiple scattering events, so the thinner the specimen the better the achievable resolution.

Even if a particle based explanation is appropriate to gain intuition, we need to study the electron as a wave to understand the phase contrast in image the image formation process. Fig. 2.5 shows a diagram to obtain a more quantitative picture of the image formation process in the microscope. We define

$$\psi_{in}(x, y) = \psi_0 \exp\{2\pi i k z\} \tag{2.1}$$

as the uniform planar wave incident on the specimen, where $(x, y)$ represent the coordinates on the specimen plane, $z$ represents the direction of the microscope axis and $k = 1/\lambda$ is the wave number of the electron. At the other end of the specimen we obtain an output

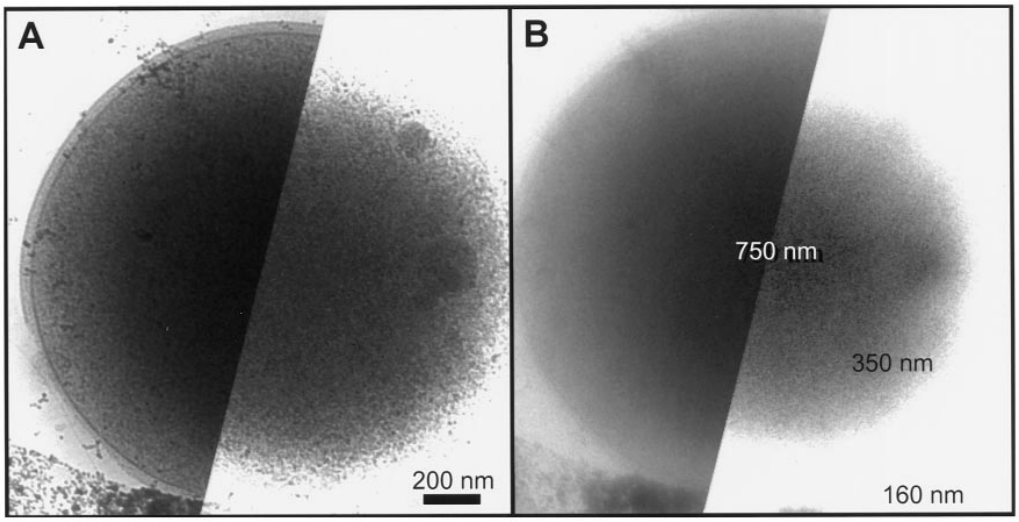


**Figure 2.4:** Images of a whole Sulfolobus HI15 PING cell in ice. (A) Zero-loss energy filtered image. (B) Unfiltered image. The better resolution of different structural details in A versus B is due to the removal of inelastic scattered electrons, which affect high resolution details generated by phase contrast mechanisms due to chromatic aberrations. The images are divided into two halves with different thickness and scaling, to visualize the full dynamic range of the recorded data. Microscope magnification was 14,500X. The numbers in B indicate the local thickness of the cell, which is about equal to the thickness of the embedding ice layer. Figure from [GSR+98].

wave from combining the unscattered electrons and the elastically scattered ones[2]:

$$\psi_{out}(x,y) = \psi_{in}(x,y) + \psi_{sc}(x,y) \tag{2.2}$$

where $\psi_{sc}(x,y)$ is the elastically scattered wave. For the most part, we can understand the image formation in the electron microscope using Fourier light optics as commonly used to describe laser optics systems to describe the coherent beam. As a result of the source coherence, electrons scattered with angle $\theta$ in the specimen will converge at the same point in the back focal plane of the objective lens (Fig. 2.5), generating the following wave:

$$\psi_F(u,v) = \mathscr{F}\{\psi_{out}\}A(u,v) \tag{2.3}$$

[2]From now on we consider a zero-loss imaging mode using the energy filter.

where $(u, v)$ are coordinates in the focal plane, $\mathscr{F}$ is the Fourier transform operator and $A(u, v)$ represents the aperture transfer function equal to 1 in the objective aperture and 0 otherwise. Finally, the wave in the image plane is represented by:

$$\psi_I(x, y) = \mathscr{F}^{-1}\{\psi_F(u, v)\} \tag{2.4}$$

$$= \psi_{out}(x, y) * \mathscr{F}^{-1}\{A(u, v)\} \tag{2.5}$$

$$= \tilde{\psi}_{out} \tag{2.6}$$

where the operator $*$ represents convolution and $\mathscr{F}^{-1}\{A(u, v)\}$ can be considered the point spread function of the microscope seen as a linear system (we will call it contrast transfer function (CTF) if we use it in Fourier domain). We use the $\tilde{\psi}$ notation to denote the function obtained after the convolution with the point spread function. The observable quantity recorded by the camera is proportional to the magnitude of the incident wave in the image plane:

$$I(x, y) \propto \psi_I(x, y)\psi_I^*(x, y) = |\psi_I(x, y)|^2 \tag{2.7}$$

Looking at the derivation above, the only unspecified function is $\psi_{sc}$, which contains the information related to the interaction between electrons and specimen. In order to quantify $\psi_{sc}$ we need to understand two possibilities to obtain information from scattering: amplitude and phase contrast. The first one generally generates medium-resolution information (above 2-3nm) and it is caused by variations on the amplitude due to absorption. Briefly, specimens containing heavy atoms such as stained membranes deflect electrons with an angle $\theta$ larger than the angle of acceptance of the objective aperture, which effectively blocks these electrons. However, the high resolution information in $TEM$ arises from the interference in the image plane between the unscattered electron wave and the scattered one in what we define as phase contrast. Briefly, if we can achieve constructive interference in the image plane between the unscattered electron wave and the elastically scattered one, the magnitude $I(x, y)$ will fluctuate proportionally to the projection of the

specimen. We will show how we can achieve this constructive interference by changing the defocus of the TEM using the wave-optical theory of image formation to mathematically formulate the concepts of amplitude and phase contrast, and precisely define $I(x, y)$.

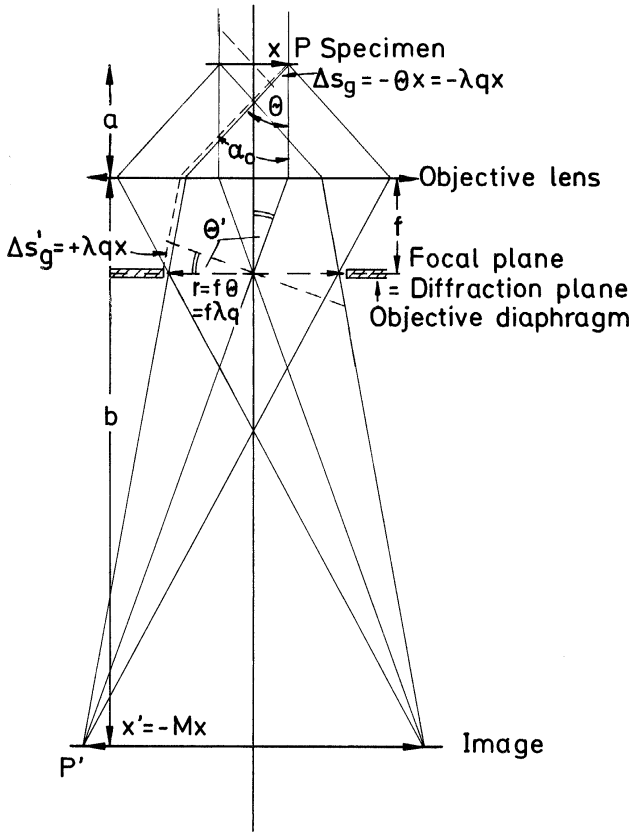


**Figure 2.5:** Quantitative description of the image formation process in the $TEM$. $\theta$ represents the scattering angle for each electron and all the electrons scattered with the same $\theta$ converge at the same point in the back focal plane of the objective lens. In particular, the unscattered electrons ($\theta = 0$) converge at the origin. The rest suffer a phase delay $\Delta s_g = -\theta x = -\lambda q x$, where $q = \sqrt{u^2 + v^2}$ represents the radial component in the focal plane coordinates. Therefore, the wave amplitude in the focal plane represents the Fourier transform of the wave amplitude exiting the specimen. Finally, the aperture (objective diaphragm) blocks all the electrons scattered with angle $\theta > \alpha_c$, effectively generating amplitude contrast in the image. $M$ represents the magnification generated by the objective and the projective lenses. Figure from [RK08].

Since $TEM$ uses high energy electrons we can use the phase-grating approximation to

decompose the scattered wave as

$$\psi_{sc}(x,y) = \psi_{in}(x,y)(1 - s(x,y)) \exp\{i\eta(x,y)\} \tag{2.8}$$

Briefly, the phase-grating approximation considers the specimen as a superposition of infinitesimally thin layers, each of them creating a possible scattering event. Using this approximation the term $s(x,y) \propto \exp\{-\int V(x,y,z)dz\}$ is a real-valued function containing the absorption and $\eta(x,y) \propto -\int V(x,y,z)dz$ is a real valued function containing the phase differences in the wave due to interactions with the specimen. In both cases the terms are related to the integral projection of the Coulomb potential in the sample along the optical axis, which will allow us to recover $3D$ density information (Section 2.2). Also in both cases we assume a weak-phase object ($\eta(x,y) \ll 1$) and a weak-scattering object ($s(x,y) \ll 1$) to perform calculations using Taylor series expansions. Using only linear terms in Eq. (2.2) and (2.8), we can approximate the amplitude of $\psi_{out}$ as follows:

$$\begin{aligned} s(x,y) &= 1 + \epsilon(x,y) + \dots \\ \exp\{i\eta(x,y)\} &= 1 + i\varphi(x,y) + \dots \\ \psi_{out}(x,y) &\simeq \psi_0(1 + \epsilon(x,y) + i\varphi(x,y)) \end{aligned} \tag{2.9}$$

First, we assume $\varphi(x,y) \equiv 0$ to study only amplitude contrast. Combining Eq. (2.5), (2.7) and (2.9) we obtain:

$$\psi_I(x,y) = \psi_0(1 + \epsilon(x,y)) * \mathscr{F}^{-1}\{A(u,v)\} \tag{2.10}$$

$$= \psi_0(1 + \tilde{\epsilon}(x,y)) \tag{2.11}$$

$$\begin{aligned} I(x,y) &\propto |\psi_I(x,y)|^2 \\ &= (1 + \tilde{\epsilon}(x,y))^2 \\ &\simeq 1 + 2\tilde{\epsilon}(x,y) \end{aligned} \tag{2.12}$$

$$\tilde{\epsilon}(x,y) = \epsilon(x,y) * \mathscr{F}^{-1}\{A(u,v)\}$$

where we have disregarded $\psi_0$ in Eq. (2.12) since it is a uniform illumination[3] and we have disregarded the quadratic terms because of the weak-scattering object approximation. Eq. (2.12) shows how absorption variations in the specimen introduce changes in the recorded amplitude by the detector.

However, if we to use the same analysis to derive the phase-contrast contribution by setting $\epsilon(x, y) \equiv 0$, we obtain $I(x, y) \propto 1$ because $\varphi(x, y)$ is in quadrature with the unscattered wave. This problem with pure phase objects was also encountered in light microscopy and led to the invention of the phase-contrast microscope by Zernike in the 1930s, which allowed one to study cells *in vivo* without staining them. In light microscopy, a phase plate is located after the specimen in the primary aperture plane with an area complementary to the illumination that does not affect the unscattered light, while the rest of the plate contains a material delaying the scattered waves by $\lambda/4$. Therefore, when both elements reach the image plane now they are in phase and generate variations in the image intensity (Fig. 2.6).

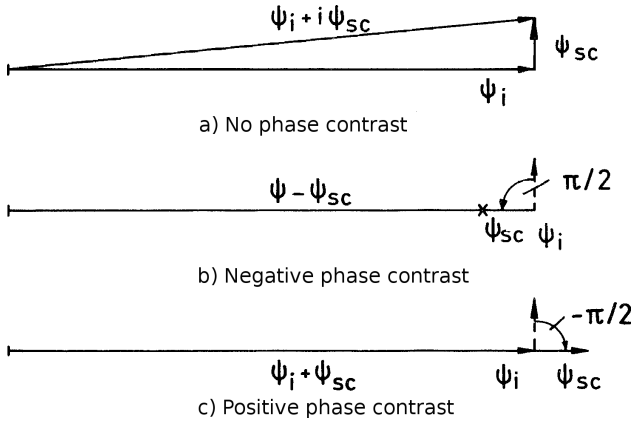


**Figure 2.6:** (a) Vector addition of the image amplitude $\psi_i$ and the scattered amplitude $\psi_{sc}$ phase shifted by $\pi/2$ radians. Variations in the amplitude of $\psi_{sc}$ do not affect the image amplitude because $\psi_{sc} \ll \psi_i$. (b) Negative phase contrast produced by an additional phase shift of $-\pi/2$. Now variations in the amplitude of $\psi_{sc}$ do affect the image amplitude. (c) Positive phase contrast produced by an additional phase shift of $+\pi/2$. Figure from [RK08]

Unfortunately, a device equivalent to Zernike phase plates is not easy to implement

[3]This normalization will be performed in all the following derivations.

in electron microscopy. Several attempts have been made [DN01, MBS+07, CDT+07] to generate $\lambda/4$ delays in the electron path. However, practical difficulties arise due to charging of the phase plates after operating them several times, which produces unpredictable phase shifts and avoids its practical use due to the requirement of replacing the plates after every image acquisition. Research and development efforts to design commercially available phase plates for TEM are work in progress, since it will improve the ability to see phase objects under the microscope.

So far we have assumed perfect lenses with no aberrations. In the following paragraphs we will explain how understanding aberrations and producing changes in the focal length it is possible to generate phase delays in $\psi_{sc}$ that allow image formation by phase contrast at certain frequencies. First, changing the focal length of the objective lense by $\Delta z$ allows us to defocus the electron beam with respect to the aperture plane, which introduces a frequency dependent phase delay in the aperture plane that can be added to Eq. (2.3) as follows:[4]

$$\psi_F(u,v) = \mathscr{F}\{\psi_{out}\}A(u,v)\exp\{-iW(u,v)\} \tag{2.13}$$

$$W(u,v) = \frac{\pi}{2}(-2\Delta z\lambda q^2) \tag{2.14}$$

$$q = u^2 + v^2 = \frac{\theta}{\lambda} \tag{2.15}$$

where $q$ represents the radial component of the Fourier coefficients. Second, even if it seems counter intuitive, in $TEM$ some lens aberrations are desired to enable contrast. Spherical aberration reduces the focal length of electrons that have been scattered by large angle $\theta$. The further from the center the electrons interact with the objective lens, the more curved the rays generated are, which creates a difference in optical paths between different scattered electrons. This difference in paths is equivalent to a frequency dependent phase delay that adds one more term in the phase delay $W(u,v)$. Again, following derivations

[4]The detailed derivations of this phase shift can be found in Section 3.3 of Reimer and Kohl [RK08].

from Section 3.3 in [RK08], the new expression is:

$$W(u,v) = \frac{\pi}{2}(C_s\lambda^3 q^4 - 2\Delta z\lambda q^2) \tag{2.16}$$

where $C_s$ is the spherical aberration coefficient of the lens provided by the manufacturer. However, for typical values of $\lambda$ and $C_s$ found in structural biology, this phase delay only plays important roles in very high frequency and the defocus value $\Delta z$ is the main element generating phase contrast.

Adding this phase delay to the derivation of Eq. (2.12) adapted to phase contrast only ($s(x,y) \equiv 0$) we obtain the following wave amplitude at the image plane:

$$\begin{aligned}\mathscr{F}\{\psi_I(x,y)\} &= \psi_0(\delta(u,v) + i\mathscr{F}\{\varphi(x,y)\})A(u,v)\exp\{-iW(u,v)\}) \\ &= \psi_0(\delta(u,v) + i\mathscr{F}\{\varphi(x,y)\})A(u,v)(\cos(W(u,v)) - i\sin(W(u,v))) \\ &= \psi_0(\delta(u,v) - \mathscr{F}\{\varphi(x,y)\}A(u,v)\sin(W(u,v)) + i\mathscr{F}\{\varphi(x,y)\}A(u,v)\cos(W(u,v)))\end{aligned} \tag{2.17}$$

where we use the Euler formula $\exp i\phi = \cos(\phi) + i\sin(\phi)$. Following the same first-order approximations as in Eq. (2.12), only the real term $\mathscr{F}\{\varphi(x,y)\}A(u,v)\sin(W(u,v))$ contributes to variations in the recorded image intensity. In this case, the CTF is $A(u,v)\sin(W(u,v))$, which brings a trade-off between generating phase contrast and limiting the image resolution due to zeros in the sinusoid. Fig. 2.7 shows different plots for the function $\sin(W(u,v))$ using different defocus to generate phase contrast at different frequencies. Fig. 2.8 shows a series of images of the specimen using different defocus to demonstrate how the CTF affects the image formation by phase contrast of $TEM$ images. Since the number of electrons that pass through the objective aperture is constant for a pure phase object, if the intensity at some points of the specimen is increased by adding the amplitudes with favorable phase shifts, the intensity at neighboring points will be decreased. Thus, if the image of a particle is darker in the center as a result of negative phase contrast, it will be surrounded by a bright rim and vice versa. Moreover, spots that are dark when $\Delta z > 0$,

will be bright for $\Delta z < 0$ and vice versa.

Finally, we can combine amplitude and phase with the above derivations using superposition to obtain the following equation:

$$
\begin{aligned}
I(x,y) &\propto 1 + 2\tilde{\epsilon}(x,y) + 2\tilde{\varphi}(x,y) \\
\tilde{\epsilon}(x,y) &= \epsilon(x,y) * \mathscr{F}^{-1}\{A(u,v)\cos(W(u,v))\} \\
\tilde{\varphi}(x,y) &= \varphi(x,y) * \mathscr{F}^{-1}\{A(u,v)\sin(W(u,v))\}
\end{aligned}
\tag{2.18}
$$

Eq. (2.18) and Fig. 2.7 illustrate the trade-off that the mechanism of achieving phase contrast through variations on the defocus brings to the $CTF$ of the $TEM$. On the one hand, we need to increase the defocus to generate phase contrast, otherwise phase objects would be invisible under the $TEM$. On the other hand, increasing the defocus generates nulls at different frequencies in the transfer function of the microscope that limit the achievable resolution of $TEM$ images. Moreover, the optimal defocus depends on the sample under the microscope and the resolution objectives of the imaging project at hand. Thus, setting the right defocus for each image is one of the key decisions that an electron microscopist faces when acquiring data.

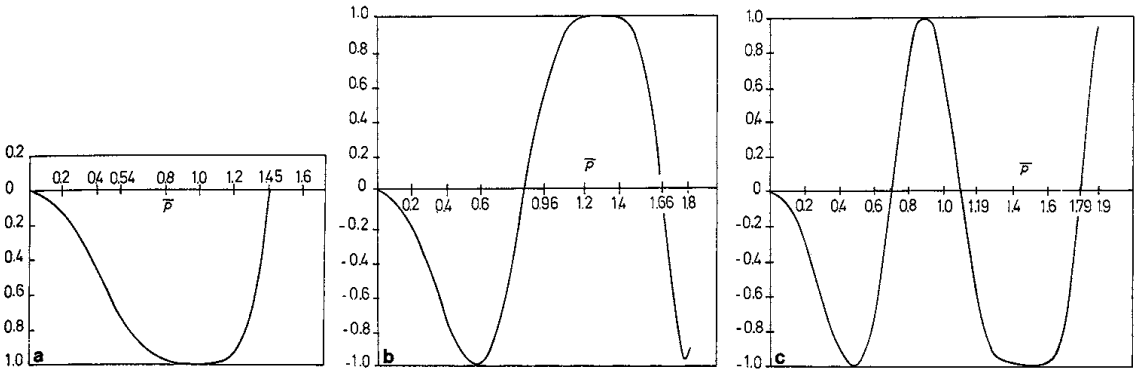


**Figure 2.7:** Plots for the function $\sin(W(u,v))$ at different defocus to generate phase contrast at different frequencies. Maximum phase contrast is generated in the frequencies with values $\pm 1$. Zero crossings indicate a null in the $CTF$ of the microscope. (a) Defocus is $\Delta z = \sqrt{C_s\lambda}$. (b) Defocus is $\Delta z = \sqrt{3C_s\lambda}$. (c) Defocus is $\Delta z = \sqrt{5C_s\lambda}$. The $x$-axis represents the reduced spatial frequency $\bar{p} = \sqrt{2}(C_s\lambda^2)^{\frac{1}{4}}$. Figure from [Fra06]

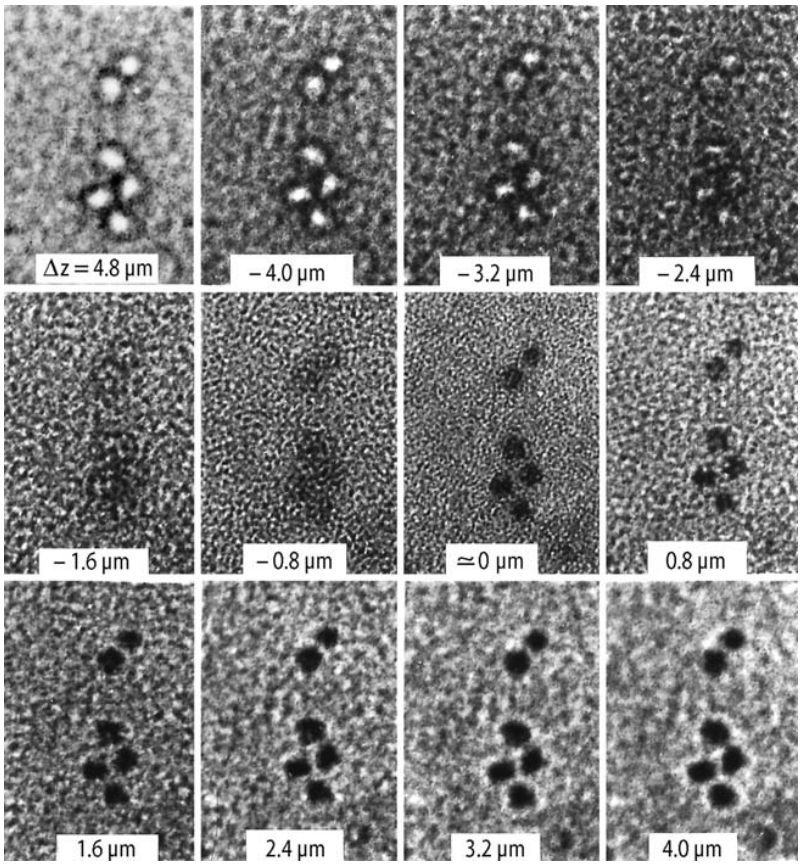


**Figure 2.8:** Defocus series of ferritin molecules on a 5 nm carbon supporting film and changes in the granulation of the carbon film (E = 100 keV). If the image of a particle is darker in the center as a result of negative phase contrast, it will be surrounded by a bright rim and vice versa. Also, spots that were dark when $\Delta z > 0$ are bright for $\Delta z < 0$ and vice versa. Figure from [RK08].

All the derivations above that show that the $2D$ image represents an integral projection of the Coulomb potential convolved with a point spread function were done based on the hypothesis of weak phase and amplitude object. Defining a clear line where this approximation does not hold is not easy, since it depends on many factors such as the heaviest atom present in the specimen, the thickness of the specimen or the energy of the electron beam. However, there are some rules of thumb based on experience with the most common acquisition practices. Samples stained with heavy atoms are definitely not weak-amplitude or weak-phase objects and most of the information recorded is through

absorption of electrons by the aperture. In this thesis we use unstained cryogenic samples, where high resolution can be obtained with a standard 300KeV microscope and zero-loss energy filter. In par actice, the weak-phase and weak-scattering approximation usually holds for specimens up to 600nm in thickness. However, the thinner the specimen the higher the attainable resolution (Fig. 2.9), because it decreases the probability of multiple scattering events.

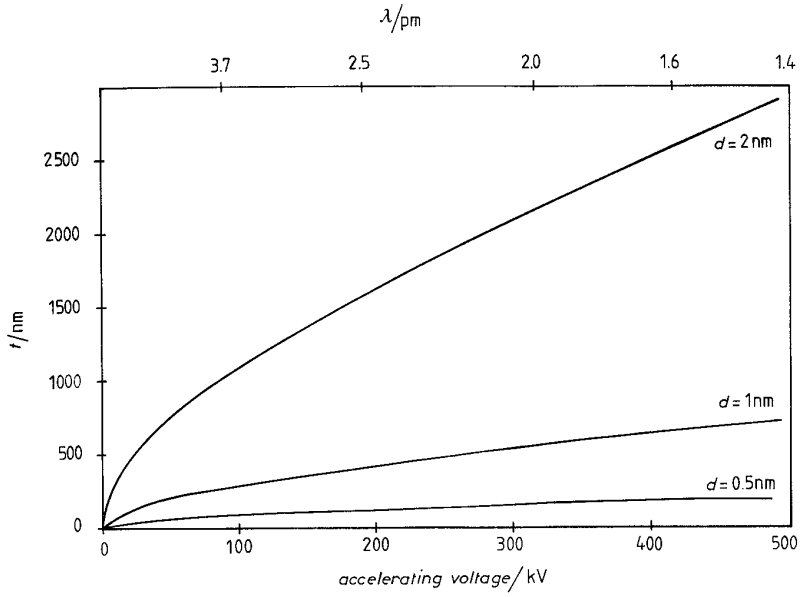


**Figure 2.9:** Maximum thickness $t$ as a function of accelerating voltage or wavelength ($\lambda$) for a resolution of 0.5, 1 and 2nm. Curves are calculating using theoretical formulations to estimate when the phase-grating approximation stops being valid for a certain resolution and a certain thickness given the energy of the electrons. Figure from [Fra06].

On top of the theoretical approximations to understand image formation in $TEM$, real microscopes have practical limitations. Two of these constraints are the shot noise and the spatial and temporal incoherence in the electron beam. First, the random process inherent to any image of electron arrival introduces Poisson noise proportional to the number of electron counts received in each pixel. In images with low counts, such as the ones in $CET$, the noise added to the low phase contrast will prove as one of the main challenges. Second, throughout the theoretical derivation we have assumed an electron

beam which has perfect temporal and spatial coherence. In practice, since the electron gun emits a certain spread of wavelengths around the nominal value $\lambda$ and the incident wave vector $\overrightarrow{k}$ is not limited to the direction of the $z$-axis, the electron beam presents limited spatial coherence. Moreover, the jittering of the high voltage around a nominal value produces incoherence in the illumination over time. All these effects can be encoded in an empirical envelope function usually modeled as an exponential decay that multiplies the $CTF$ shown in Eq.( 2.18).

All the elements summarized in this section are necessary to understand what resolution can be expected for different specimens and how different imaging parameters can be optimized depending on the end goal of the project at hand. However, each single image contains only information of the projection of the specimen from a given angle, which is not enough to reconstruct a $3D$ density map. In the next section we will show how we can reconstruct this map by combining multiple $2D$ images obtained from different angles.

## 2.2 3D reconstruction from multiple 2D TEM images

The fundamental result that allows us to obtain $3D$ density maps from $2D$ projections is the Projection-Slice Theorem [Bra86]. The theorem states that given an N-dimensional function $f(x)$ the following two operations return the same result (Fig. 2.10):

1. Compute the integral projection of $f(x)$ along a given direction to obtain an N-1 dimensional function and then take the Fourier transform of that projection.

2. Compute the Fourier transform of $f(x)$ and then return the coefficients along a "slice" with the same orientation as the projection

In other words, the Fourier transform of each projection recorded in a $TEM$ image returns a hyperplane of Fourier coefficients of the original $3D$ density. Thus, if we obtain $TEM$ images from multiple angles to fill the Fourier space in $3D$, we can recover the original $3D$ density by a simple inverse Fourier transform. Qualitatively, this is the idea behind any tomographic application, although there are many implementation details

that are beyond the scope of this introduction. In this section, we will present different methodologies in electron microscopy used to obtain the necessary projections to recover the $3D$ density of different biological structures.

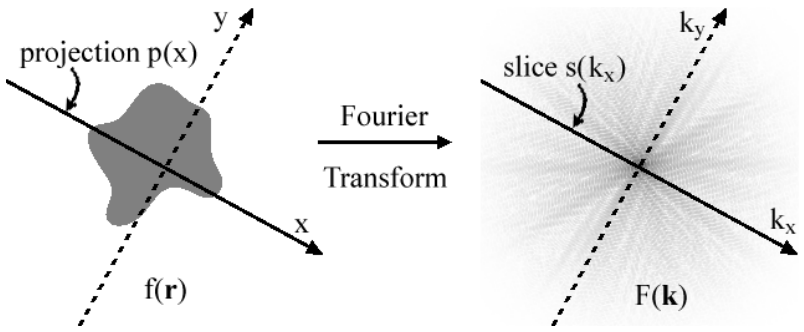


**Figure 2.10:** Schematic representing the projection-slice theorem in $2D$. Projecting the density into a 1D line and then taking the 1D Fourier transform of that line is equivalent to computing the $2D$ Fourier transform of the density and taking a "slice" of coefficients with the same angle as the projection. In other words, $\mathscr{F}\{p(x)\} = s(k_x)$ in the figure from http://en.wikipedia.org/wiki/Projection-slice_theorem.

Before we do that, we will present the proof of the Projection-Slice theorem for completeness. For simplicity, take $f(x, y)$ as a $2D$ function. We can define the following operators:

$$p(x) = \int_{-\infty}^{\infty} f(x, y)dy \tag{2.19}$$

$$F(u, v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\exp\{-2\pi i(xu + yv)\}dydx \tag{2.20}$$

where $p(x)$ represents the projection along the $y$-axis and $F(u, v)$ is the $2D$ Fourier transform. Simply setting $v = 0$ in Eq. (2.19) proves the equivalence:

$$F(u,0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\exp\{-2\pi ixu\}dydx \tag{2.21}$$

$$= \int_{-\infty}^{\infty}\exp\{-2\pi ixu\}\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx \tag{2.22}$$

$$= \int_{-\infty}^{\infty}\exp\{-2\pi ixu\}p(x)dx \tag{2.23}$$

$$= \mathscr{F}\{p(x)\} \tag{2.24}$$

This proof can be extended using the same methodology to a function of N-dimensions and a projection along any direction. These concepts are intrinsically related to the mathematical operator defined as the Radon transform [Her09], which calculates the projection of the function along any "slice". In general, reconstructing the $3D$ density map is equivalent to obtaining the inverse Radon transform, a problem extensively studied in different fields.

The methodology to acquire all the necessary projections to reconstruct the $3D$ density of the specimen under the $TEM$ depends on the type of specimen. Ideally, we would like to be able to operate the same way as a computerized tomography (CT) scan operates in human patients. Briefly, the patient lies inside a cylindrical machine that takes projections from angles within 180 degrees range5, effectively oversampling the $3D$ Fourier space to obtain a perfect reconstruction of each individual patient. Unfortunately, the situation with small biological structures under the $TEM$ is not as ideal. Due to limitations in the dose that biological material can tolerate and the sample fixture, we can not obtain projections from every possible angle. We will see this in more detail in the next section, but the main problem is that when the sample is tilted too much, it becomes too thick and most of the electrons suffer multiple scattering events, eliminating any information in the recorded image. Therefore, we need to try to overcome this problem making different approximations depending on the sample under the microscope.

First, if the specimen has a crystalline structure we can assume that different samples will have exactly the same conformation because of the crystalline regularity. Therefore, we can obtain a single high $SNR$ image per sample at different orientations by tilting the specimen, recording multiple noisy images, and combining the information from images of different samples. If we are able to find the common origin of the crystal for all the images and to determine the tilt angle for each projection, we can combine all the images to recover the $3D$ crystalline structure using a least-square fitting of all the diffraction measurements. These two tasks are greatly simplified by the fact that a crystalline structure has a discrete representation in Fourier domain. Briefly, the common phase origin is found by taking the zero tilt image as a reference and aligning the rest to it. This alignment is possible by using pairwise *common lines*: according to the Projection-Slice theorem presented above, two projections should have a common line of Fourier coefficients along the origin. We can use that common line to determine the phase shift between projections. The tilt of each specimen can be derived from the distortion that is observed in the diffraction pattern between the untilted image and the tilted one. The reader is referred to Glaeser at al. [RG07] for more details.

The second category of biological structures is known as *single particles*. These specimens are macromolecules that can be purified in the sample prepration process into a homogeneous population. Homogeneity is the key element in this kind of analysis: we do not require for the specimen to be crystalline anymore, but to be identical from one sample to the next so we can still combine information from multiple images. Using this assumption, we can follow the same procedure as for crystalline structures obtaining a single high-dose image per sample, although in this case we do not need to tilt the sample, since each particle would be randomly oriented in the ice. As before, we need to determine a common origin and the orientation of each particle in order to combine all the images in a unique $3D$ density map. However, now the biological structures have a continuous spectrum, which complicates the task. There is a vast literature proposing efficient methods to solve the alignment problem in single particle work and arguably the most common

one is known as iterative projection matching. Briefly, we obtain a coarse $3D$ model of the specimen. Then we project that model in all possible directions to identify the orientations of all the particles by maximizing the cross-correlation between the reprojections and the raw data. Once we have classified each particle, we can obtain a new $3D$ density map by solving the inverse Radon transform with standard backprojection algorithms. This cycle of $3D$ reconstruction and alignment is iterated until convergence. In general, a classification step is also necessary to detect particles that do not satisfy the homogeneity assumption. Typical structures analyzed by single particle methods are viruses, ribosomes or microtubules and $3D$ density maps with near-atomic resolution are generally achieved by merging thousands of particles. The reader is referred to Frank [Fra96] for more details.

Larger macromolecules or structures such as the chemotaxis apparatus or the bacterial cell wall can not be purified into a homogeneous population and need to be studied *in situ* to understand the complex spatial relationships within the cell. In these cases, each structure is unique and we can not merge data from different projections in a simple manner, so we need to obtain multiple projections from a single specimen (Fig. 2.11). This methodology is known as electron tomography ($ET$) and has two main limitations: first, the maximum electron beam dose permitted before damaging the biological structure needs to be spread across multiple projections, which makes each individual image noisier. Second, we can not obtain projections from all possible angles, which will cause missing information in the sampling of the Fourier coefficients of the $3D$ density map, effectively distorting the resulting reconstruction in certain directions. In the next section we present a detailed explanation of all the steps and challenges associated with cryo-electron tomography ($CET$), since it is the methodology used to acquire the images analyzed in this thesis.

## 2.3 Cryo-electron tomography pipeline and challenges

Acquiring and processing an electron microscope tomogram requires different steps outlined in Fig. 2.12. Thanks to multiple research efforts many of the steps have been already

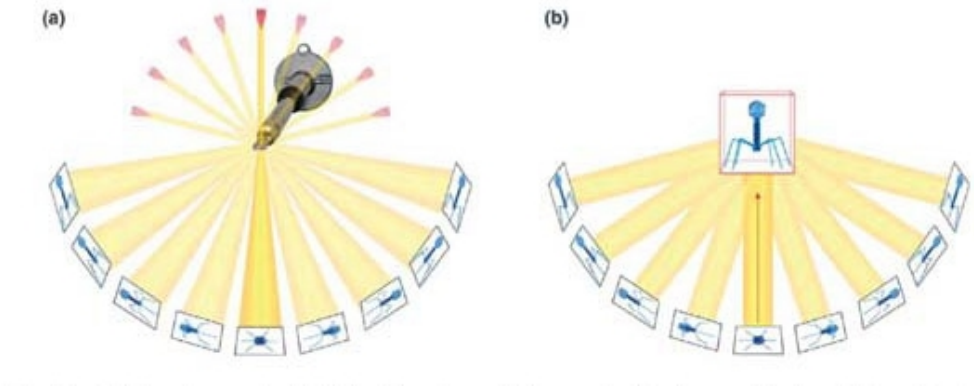


**Figure 2.11:** Figure from [MNM05] showing the main principle of electron tomography. We obtain projections of the sample from different tilt angles so we can reconstruct a $3D$ density volume of the specimen under observation through standard computerized tomography techniques.

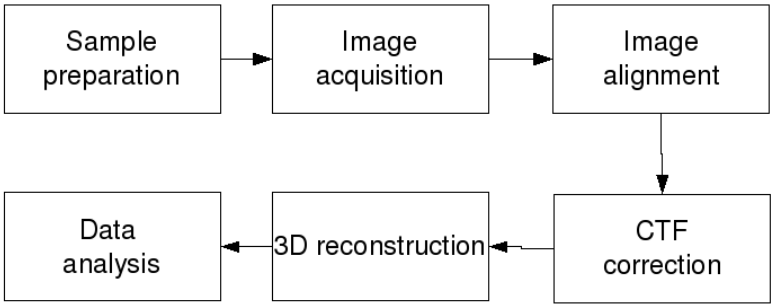


**Figure 2.12:** $CET$ pipeline to acquire and process a single tomogram. We need to automate as much as possible every block to achieve real high-throughput tomography.

automated. However, there are still many open problems to achieve real high-throughput $CET$. One of the goals of the work presented in this thesis is to solve some of these bottlenecks. Being able to obtain hundreds of tomograms per months will extend the horizons of $CET$. In this section we will describe each step of the pipeline and the associated challenges to identify bottlenecks and their associated constraints.

### 2.3.1 Sample preparation

Tomography samples can be prepared through a variety of methods. For cryo-electron tomography, the samples start in a liquid medium placed into a thin carbon grid. These grids are then flash frozen using liquid ethane. The temperature drops at rates of 10,000 K/s, removing kinetic energy too quickly for water molecules to rearrange into ice crystals. Instead, amorphous ice is formed, which preserves samples in their native state with minimal artifacts. Once such specimens are frozen, they can be examined artifact-free under the microscope. However, the high energy electrons used in $CET$ cannot penetrate more than about 600nm into biological material without being inelastically scattered. Therefore, thicker samples must first be sectioned before imaging, which can introduce artifacts. In order to preserve the native structure of cells and tissues, thick samples can be quickly frozen under high pressure, which prevents internal water from expanding to form ice crystals. Such frozen blocks of material can be sectioned directly ("cryosectioning") and imaged using $CET$, or they can first be chemically fixed, plastic-embedded, and stained, and then sectioned at room temperature ("plastic sections"). Alternatively, thick samples can simply be chemically fixed, plastic embedded, stained, and sectioned directly (without high pressure freezing). In this thesis, we present algorithms developed for cryo-samples of whole cells without sectioning, although most of them can be applied to other TEM samples.

### 2.3.2 Image acquisition

Because no goniometer is mechanically perfect, and because the specimen is never at its exact eucentric height as the sample is rotated in the microscope, it moves laterally and vertically within the column as it is tilted. Thus, to keep the target centered under the beam and at a constant focus, before each image is taken the beam must be electronically shifted to center the image and the strength of the objective lens has to be adjusted to maintain proper focus. In order to find the magnitude and direction of the shifts and focus changes that should be applied before each image is taken without exposing

the target to any unnecessary dose, the movements can either be predicted by modeling trends from previous images in the tilt-series [ZMB+09, SSC+09] or measured indirectly by deflecting the beam to an off-target, tracking position, and recording reference images [NFL+05, Mas05]. Existing data acquisition packages are a mature technology and perform all the above tasks automatically.

### 2.3.3 Image alignment

If the image acquisition were perfect, we could obtain a $3D$ density map from the raw data. However, we need to align the series of tilted images in order to correct for the mechanical imperfections and tracking errors during acquisition. In order to do that, we need to identify and correspond common features between different projections. Most available methods [KMM96, HCWS08] perform this step in a semi-automatic fashion. Briefly, microscopists put gold fiducial markers on the sample to create high contrast point-like features. Then, the user seeds a fiducial model in the zero tilt image and the program tries to find the markers corresponding to those seeds in the rest of the images. Then, the tracking results are presented to the user with a GUI to correct possible mistakes. The process is iterated until the user is satisfied with the alignment. Since the process involves a substantial amount of user interaction, a fully automatic alignment algorithm is needed to obtain real high-throughput in the tomography pipeline and will be presented in Chapter 3.

Alignment of $CET$ images with fiducial markers would be trivial if it were not for the very low $SNR$ generated by several factors: first, any biological sample suffers structural damage from electron beam radiation. Thus, there is a specific electron dose limit before the sample is irreversibly damaged and structural information is lost (Fig. 2.13). This dose limitations effectively reduces the attainable $SNR$ in $CET$ volumes because we have to divide the dose among all the projections. The $SNR$ difference between different exposures is apparent in Fig. 2.14.

Second, sample thickness dramatically influences the final quality of the image as shown

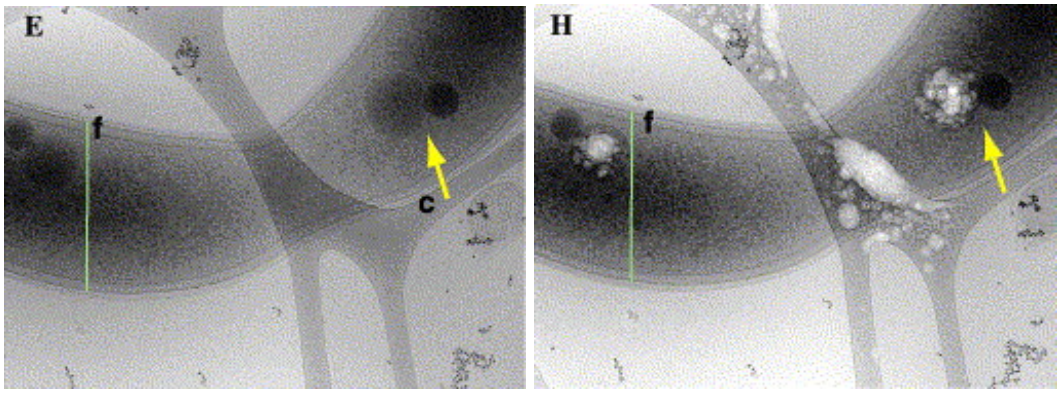


**Figure 2.13:** Energy-filtered projection images of *Caulobacter crescentus* cells in amorphous ice to exemplify radiation damage in the sample caused by electron beam overexposure. This dose restriction effectively limits the amount of electrons that can be used in the imaging process and the attainable $SNR$ in $CET$ volumes. Left: image with accumulated exposures of 1560 e/nm$^2$. No structural damage by beam radiation is noticeable. Right: image with accumulated exposures of 21800 e/nm$^2$. Arrow marks area that shows radiation damage of the biological material. Figure from [CD05].

in Section 2.1. Typical whole cell preparations of *Caulobacter crescentus* range between 400nm to 600nm in thickness, which is considered the limit for $CET$ samples. Finally, because $CET$ samples are not stained with heavy metals, they produce in low contrast images. In other words, the electron interaction with amorphous ice or with biological specimens is very similar. The three factors combined together result in images with $SNR$ well below 0dB and very faint features, which is the reason why in general off-the-shelf image processing procedures such as edge detection or feature descriptor matching do not to work in $CET$ volumes and we need to develop specific robust statistical procedures to handle the images.

### 2.3.4 CTF correction

We have seen in Section 2.1 how setting the right defocus to balance the trade-off between phase contrast and resolution is a key decision when obtaining TEM images, since the $CTF$ modulates each frequency depending on the defocus value. When combining multiple $2D$ images to generate a $3D$ volume, a third problem arises: the same frequency band from

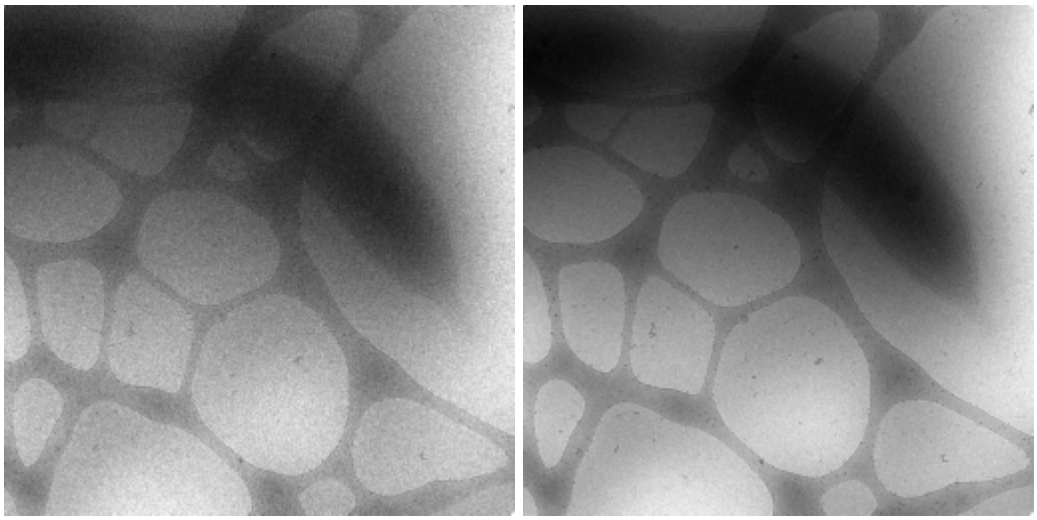

**Figure 2.14:** Example of how dose limitations affects image $SNR$. Left: image with 60ms exposure. Right: same image with 240ms exposure. The number of accumulated counts per pixel in the left image is lower than in the right image, which decreases the $SNR$ due to the Poisson statistics of the $CCD$. For example, the ice areas inside the grid show much less salt and pepper noise on the right than on the left image.

different image areas contains a different transfer function because the defocus changes across tilted specimens. Since $CTF$ changes the phase of the information after each zero-crossing (Fig. 2.7), the reconstruction combines intensities with opposite phase, effectively reducing the final resolution. Thus, we need to estimate and deconvolve the CTF in each pixel if we want to obtain higher resolution tomograms. Section 4.4.2 presents numerical simulations demonstrating these effects.

Correcting CTF in tilt series images has proved elusive and it is still an open problem. Most microscopists acquiring $CET$ tilt series adjust the defocus to optimize the trade-off between contrast and resolution attainable to the first zero of the CTF, so no correction is needed afterwards. In single particle, where hundreds of images are taken at zero tilt, estimating the CTF is possible through averaging. The most common approach is to parametrize the CTF following the weak-phase and weak-scattering approximation introduced in Section 2.1 [ZPSF97, LBC99] and present the user with a GUI that allows one to tweak the parameters until a fit to an averaged power spectrum of the data is

achieved. Then, phase-flipping operation [PEA07] or Wiener filter [FLC06] is used to correct each projection. In the past few years, there have been attempts to extend some of these ideas to $TEM$ tilt-series [FLC06, XMS$^+$09], although the results depend on the $SNR$ of the images.

### 2.3.5 Three-dimensional reconstruction

Once the images have been aligned, obtaining a $3D$ density map can be formulated similar to other imaging modalities such as X-rays or CT scans. Standard techniques such as weighted backprojection (WBP), algebraic reconstruction technique (ART) and simultaneous iterative reconstruction technique (SIRT) have been implemented in different software packages to generate high resolution $3D$ density maps [Mas97, NFL$^+$05, SMVM$^+$04]. For cases where the first-order approximation of the microscope as a linear structure projector presented in Section 2.1 does not hold, such as large field tomography of plastic sections, advanced mathematical formulations based on algebraic geometry have been introduced [LBPE06]. These approaches correct for non-linear effects present in the images by prolonged exposure to electron beam radiation using inverse Radon transform of curvilinear projections. All the methods for reconstruction depend on a high accuracy sub-pixel alignment to obtain a high resolution density map.

### 2.3.6 Data analysis

The term data analysis is generally applied to any post-processing operation that uses the datasets to provide information to the users from a large set of voxels, and includes fields such as visualization, segmentation, denoising or template matching. Usually, users are interested in very specific information from each tomogram that can be encoded efficiently. However, several reasons make the extraction of information challenging. The first obvious one is the fact that $CET$ density maps have low $SNR$, since a reconstruction from a 100+ noisy projections is still noisy. Second, from the Projection-Slice theorem introduced in Section 2.2, we know that each projection represents a slice of the $3D$ density volume in

Fourier space. Since for slab-like grid, such as the ones used in this work, the thickness of the sample increases as $1/\cos(\beta)$, where $\beta$ is the tilt angle of the specimen, at $\beta = 60^\circ$ the effective thickness of the sample has doubled. At this point, depending on the original thickness of the sample, almost no electron contributes to image formation. This limitation implies that in general we acquire data in a range well below $\pm 70^\circ$ (Fig. 2.15). In Fourier space, that is equivalent to having a "missing wedge" of roughly 25% of the data. This effect blurs the $3D$ reconstruction in the direction of the electron beam, affecting different features depending on the orientation of the objects inside the $3D$ volume. Thus, the same structure can look very different depending on its relative orientation with respect to the missing wedge (Fig. 2.15). Different acquisition schemes, such as dual-axis tilt [Mas97] or conical tomography [LCB$^+$05], have been developed to fill the Fourier space better given the tilting limitations. However, due to the dose limitation factor and the large thickness of whole bacterial cells, these approaches are not commonly used in the $CET$ samples presented in this thesis.

Third, any biological structure is flexible and might exist in different conformational states (Fig. 2.16). Therefore, there is an intrinsic variability in the data that contains relevant biological information and needs to be differentiated from the noise explained above. Proposed approaches [LWR$^+$06, IMF$^+$05, PYFS06] to distinguish noise from real significant variation are based on collecting enough samples of similar structures, aligning them with respect to each other and perform classification to obtain different conformational states. Until recent years, the amount of samples needed to perform these studies was out of practical reach due to the high noise in the images. Real high-throughput tomography allows us to collect more samples for each experiment, facilitating the incorporation of heterogeneity in structural models.

Fourth, a standard whole cell $CET$ volume is recorded using a 2Kx2K or 4Kx4K $CCD$ device for 100+ projections. Considering a common sampling of 1nm/pixel and an average thickness of 500nm, we obtain a volume with 4096x4096x500 voxels in float precision occupying 32GB of memory. Thus, any $3D$ exploration or computation in such

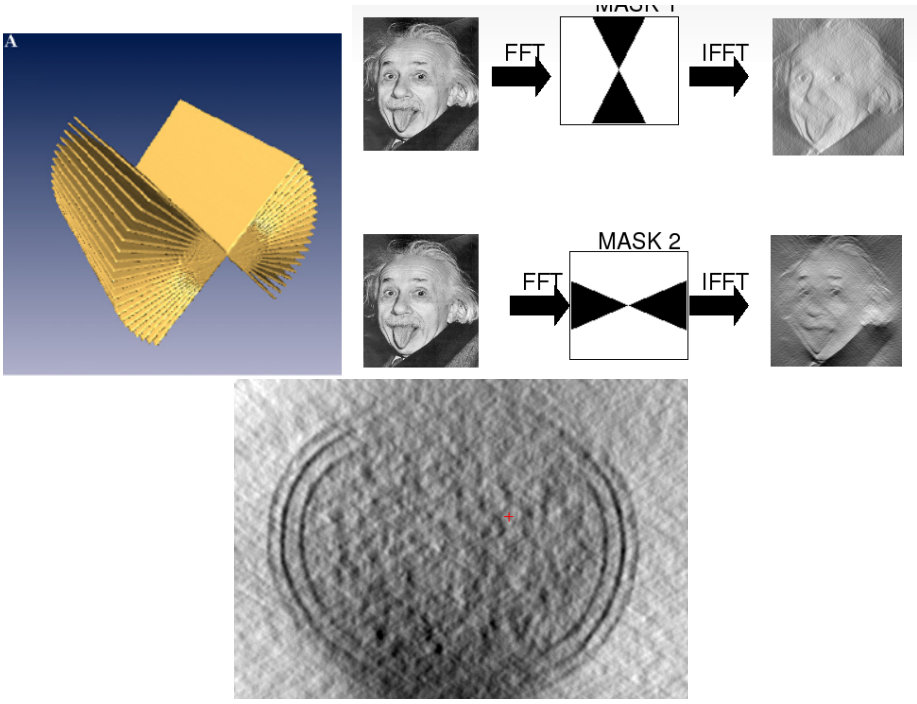


**Figure 2.15:** Top left: figure from [LCB+05] to represent how Fourier space is filled with each projection. In single axis tomography we have a "wedge" of missing data, which blurs features in the direction parallel to the optical axis. Top right: $2D$ synthetic example of missing wedge effect. Different features are blurred for different orientations of the missing wedge: top image shows no eyebrows while bottom image shows no hair threads. Bottom: cross-Section of a *Caulobacter crescentus* tomogram. Features of cell wall elements disappear on top and bottom of the image due to the blurring caused by the missing wedge.

a large volume needs to be efficient in order to be implemented in practice. In practice, there are several strategies that facilitate handling large datasets. For example, most $CET$ labs have computer clusters to allow parallel computations over multiple datasets. Also, stacks of the same tomogram at different resolution levels (pyramids) help reducing the analysis time using a coarse-to-fine approach. It is also important to incorporate as much prior information as possible from other sources such as biochemistry, X-ray or light microscopy to reduce the search space in each dataset.

Finally, one of the main advances in light microscopy has been the possibility of labeling targeted structures using different fluorescent markers [Phi01]. Labeling simplifies

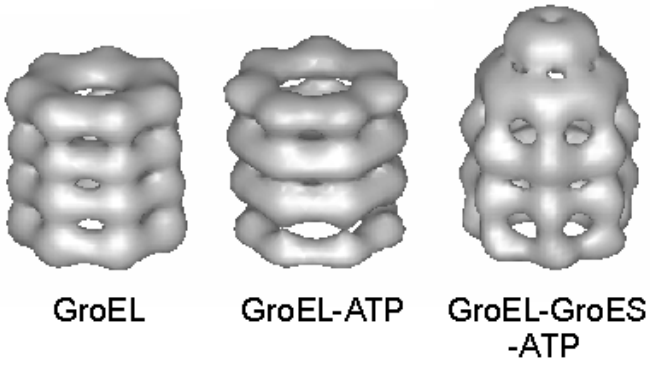


**Figure 2.16:** Surface-rendered views of $3D$ reconstructions of GroEL, GroEL-ATP and GroEL-GroES-ATP from cryo-$EM$ to exemplify heterogeneity in biological structures. Image from [CRH+94].

image analysis tasks such as localization and expression level measurements. Several attempts [HJB09, SAK+07, LAS+08, EHSH09] have been made to introduce heavy atom labels such as nanogold particles into bacterial specimens to facilitate localization of targeted structures. However, most of them have been at best moderately successful. That generates the necessity to develop image based localization and alignment techniques using low contrast biological structures as features, which produces large uncertainties. Any new method needs to take this uncertainty into consideration before making final decisions.

Given that until recently it was hard to obtain large numbers of images for the same experiment, it was feasible for biologists to try to analyze each dataset by hand. Real high-throughput tomography is changing the $CET$ panorama because it is increasing the need to automate the data analysis phase, since the amount of data that needs to be processed is increasing exponentially. Many experiments are now limited by data analysis and not by data collection. On the other hand, with the increasing number of datasets we can start to use a number of powerful statistical processing methods to aid image analysis, which also makes automated data analysis more feasible.

As stated above, image alignment is one of the first bottlenecks in the $CET$ pipeline to achieve this automation. In this thesis we will see how we can leverage existing probabilistic frameworks from fields such as computer vision and computer graphics to remove

this bottleneck. Moreover, we will show how similar techniques can be applied for data analysis to combine thousands of tomograms to obtain higher resolution volumes in a similar way as is done in single particle. The next section presents related work in image alignment to show the strengths and weaknesses of different approaches with respect to the characteristics and challenges of $CET$ images explained above.

## 2.4 Related work in image alignment

High accuracy image alignment between $CET$ images is critical to obtain high-resolution structural information. For example, alignment of $2D$ projections directly affects tomogram resolution; and alignment of $3D$ volumes allows to obtain better resolution structures through averaging. However, we have seen that the alignment task is challenged by noise, missing wedge and heterogeneity. Moreover, we do not have ground truth of the structures we are looking for, which poses a problem during validation. Most of the alignment or template matching in $CET$ uses some sort of cross-correlation-like measure to find optimal matches [BSL$^+$08, OFK$^+$06, WLR$^+$09] in an exhaustive search among all possible rigid transformations (translations and rotations). Using standard techniques such as coarse-to-fine search or exploiting symmetry reduce the computational time. The main issue remaining is the fact that high noise and missing data provide many false positives when comparing two structures. In this thesis, we exploit local spatial correlations and sparsity to increase accuracy and robustness of alignment procedures in $CET$ images. Spatial correlations help disambiguating all the possible matches coming from template matching to impose coherence between neighboring features. Sparsity helps finding a representation of the signal with fewer coefficients, which facilitates detection in low $SNR$ environments. This section presents a brief summary of work in image alignment related to the techniques presented in this thesis.

### 2.4.1 Alignment using the whole image

We can try to calculate the transformation that registers image A into image B by minimizing a dissimilarity score or metric between all the gray values in an image. This approach can be expressed mathematically as follows:

$$\underset{f}{\operatorname{argmin}} \sum_{i,j} \rho(I_A(i,j), I_B(f(i,j))) \text{ for 2D images} \tag{2.25}$$

$$\underset{f}{\operatorname{argmin}} \sum_{i,j,k} \rho(I_A(i,j,k), I_B(f(i,j,k))) \text{ for 3D volumes} \tag{2.26}$$

where $\rho$ is the dissimilarity measure between gray values. Popular dissimilarity measures include mutual information [VV95] and cross-correlation [Guc82]. $f$ represents the transformation between two datasets, and its parametrization is crucial to define which distortions are allowed between images [CHH04]. For example, in the $2D$ case of $CET$ tilt series, an affine transformation should be sufficient because the size of the sample is very small in comparison with the focal distance. Therefore, we do not need to consider perspective effects and we can model $f$ as follows:

$$f(i,j) = (a \cdot i + b \cdot j + u, c \cdot i + d \cdot j + v) \ ; \ a,b,c,d,u,v \in \mathbb{R} \tag{2.27}$$

This approach is chosen to perform the coarse alignment of $CET$ images. However, there are three main issues: first, $CET$ images are low contrast and alignment using the whole image does not deliver subpixel accuracy to obtain optimal resolution reconstructions. Second, the error propagates along the tilt series because we just align each pair of images sequentially. Third, in thick samples, transformations estimated from $2D$ projections are not enough to correct for all possible distortions introduced by the microscope lenses, the electron trajectories through the specimen and the specimen shrinkage itself.

In the case of registering $3D\ CET$ volumes, we usually restrict $f$ to rigid transformations to reduce the large search space. Moreover, low $SNR$ and missing wedge present challenges for non-rigid registration approaches. In Chapter 4 we will show how to define $\rho$ to make it more robust against $CET$ challenges.

### 2.4.2 Alignment using points of interest and local descriptors

One way to overcome some of the limitations exposed above is to align images based on common landmarks: instead of comparing every pixel, we try to find prominent features in image A and correspond them to prominent features in image B. Once we have the correspondence, we can find the parameters of the transformation $f$ using optimization techniques. Therefore, the key point is to find the features and to be able to compare them.

The list of possible interest point detectors and descriptors is too large to summarize here. However, edges and corners [HS88, MFM04] are among the most popular ones for natural images. Descriptors based on local histograms of intensities around each feature [Low04, MS05, TM08] have proved to be computationally efficient and highly discriminative, which prunes the number of possible corresponding points between images. These descriptors are also robust to change in scale, orientation and illumination between images.

Once we have possible correspondences, the most common approach uses RANdom SAmple Consensus (RANSAC) [FB81] to robustly estimate $f$ even in the presence of outliers. Briefly, we sample $P$ features from image A and match them to similar landmarks in image B, where $P$ is the minimum number of points to estimate $f$ (3 in the case of affine transformations). Then we compute the residual error of the estimated $f$ with respect to the given set of points. If we repeat the process multiple times using random sampling with replacement on the selected $P$ points, we are guaranteed to find an optimal solution with a certain probability. From all the trials we choose the one with the lowest residual error.

In tilt $CET$ series we can detect points of interest using a gold bead template. However,

we can not discriminate between different landmarks because all the gold beads look the same. The only possible distinction is isolated markers versus group of markers (Fig. 3.6), although this distinction does not prune enough the number of possible correspondences to make RANSAC a feasible approach. In $CET$ volumes similar salient features might look different due to low $SNR$ and missing wedge effects, which ruins the performance of standard local descriptors.

### 2.4.3 Alignment using Markov Random Fields

Anguelov et al. [AScPK04] and Coughlan and Ferreira [CF02] present a probabilistic formulation of the alignment problem for deformable shapes such as range laser scans of human bodies and hand-written letters respectively. Both references effectively define a probability distribution over the set of possible correspondences and search for high probability assignments as good solutions. Using probabilities as a mathematical tool to quantify uncertainty has proved very useful in many image processing and pattern recognition tasks because it does not disregard information by making decisions in intermediate steps. However, the improvement comes at the price of an exponential increase in complexity. Both Anguelov et al. and Coughlan and Ferreira use graphical models, and in particular Markov Random Fields ($MRF$), to model the probability distributions and reduce the complexity without sacrificing performance. The approach presented in this thesis follows a very similar probabilistic framework for the alignment of $CET$ images. In this section, we will give a brief introduction to $MRF$ and how they have been successfully applied to alignment problems. The reader is referred to books by Jordan [Jor98] and Koller and Friedman [KF09] for an in depth treatment of the subject. Chapter 8 in [Bis07] provides a good introduction to the topic for a first time reader.

### Markov Random Fields

The introduction from [Jor98] provides a qualitative picture of the key ideas behind graphical models: "Graphical models, a marriage between probability theory and graph theory, provide a natural tool for dealing with two problems that occur throughout applied mathematics and engineering–uncertainty and complexity. In particular, they play an increasingly important role in the design and analysis of machine learning algorithms. Fundamental to the idea of a graphical model is the notion of modularity: a complex system is built by combining simpler parts. Probability theory serves as the glue whereby the parts are combined, ensuring that the system as a whole is consistent and providing ways to interface models to data. Graph theory provides both an intuitively appealing interface by which humans can model highly interacting sets of variables and a data structure that lends itself naturally to the design of efficient general-purpose algorithms."

The goal is to decompose probability joint distributions into a product of local distributions (*factorization*) to perform probabilistic queries more efficiently. Imagine we have a set of random variables $\boldsymbol{X} = (X_1, \ldots, X_N)$ and we want to characterize the joint probability distribution $P(X_1, \ldots, X_N)$. In general, without any other information, this task is exponentially complex. However, if we know that only certain random variables directly interact with each other, we can try to decompose the joint probability as follows:

$$P(X_1, \ldots, X_N) = \frac{1}{Z} \prod_C \psi_C(\boldsymbol{X}_C) \text{ with } \psi_C(\boldsymbol{X}_C) \geq 0 \tag{2.28}$$
$$Z = \sum_{\boldsymbol{X}} \prod_C \psi_C(\boldsymbol{X}_C)$$

where $Z$ is the partition function to normalize all the products and $\psi_C$ represents a factor over a set $C$ of random variables which are directly correlated to each other. Note that in general, each factor can not be interpreted strictly as a probability distribution, but as an encoding of preferred assignments of different sets of random variables.

A Markov chain is a simple example to illustrate the idea of factorization. Imagine

four random variables $\boldsymbol{X} = (X_1, X_2, X_3, X_4)$ representing a Markov chain. Using the chain rule for probability distributions we can decompose the joint distribution as follows:

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)P(X_4|X_1, X_2, X_3) \tag{2.29}$$

Thanks to the Markovian property, we can simplify the structure further, since we know that each random variable $X_i$ only interacts directly with its neighbors. Therefore, we have:

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3) \tag{2.30}$$

In this simple case, we have four factors, one with cardinality of $|C|$ equal to 1, and three with cardinality of 2. Without the Markovian property the cardinalities would have been much larger, creating a more complex structure. This example of a Markov chain can also be captured graphically in a very intuitive manner (Fig. 2.17). Each node in the graph represents a random variable and we insert an edge only between random variables that appear together in some factor. In this case, the edges are only between adjacent random variables.

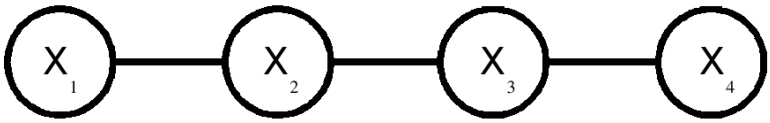


**Figure 2.17:** Markov Random Field encoding a Markov chain over four random variables.

In fact, following this procedure, we can associate a graph $G = (V, E)$ to any joint probability distribution $P(X_1, \ldots, X_N)$ that factorizes as shown in Eq. (2.28). Each node $V_i$ represents the random variable $X_i$ and we insert an edge $E_{ij}$ between $V_i$ and $V_j$ if and only if a factor $\psi_C$ contains $X_i$ and $X_j$. If there is a cause-effect relationship between random variables, the edges are directed to encode this information. In this case, we

call the graphical model a Bayesian Network. If the factors only encode preferences or constraints between random variables, the edges are undirected, and we call the graphical model Markov Random Field ($MRF$). In this chapter, and in the rest of this thesis, we explain and use the formalism of $MRF$ to encode preferences coming from image features and local spatial correlations.

At the core of the factorization of probability distributions resides the notion of *conditional independence*: we say that $X_i$ and $X_j$ are independent given $X_k$ if and only if $P(X_i, X_j|X_k) = P(X_j|X_k)P(X_i|X_k)$. This concept formalizes the idea of whether two random variables are directly interacting with each other. Fig. 2.18 shows how this notion can be read directly from an $MRF$: given a random variable in the node $V_k$, two random variables in nodes $V_i$ and $V_j$ are conditionally independent if after removing all the edges connected to $V_k$ there is no path connecting $V_i$ to $V_j$.

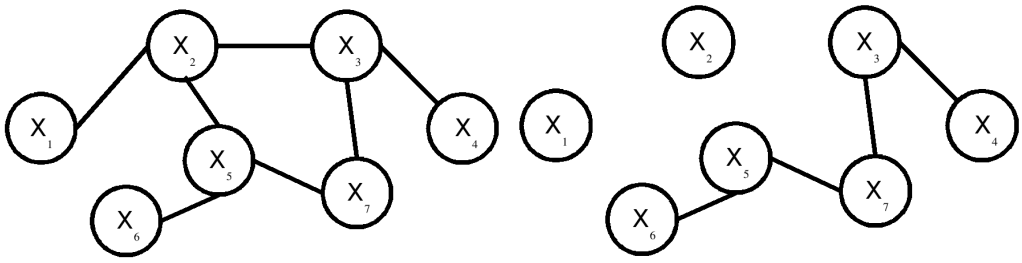


**Figure 2.18:** Example showing how to determine conditional independence encoded in the $MRF$. Left: original graph over seven random variables. Right: removing all the edges connected to $X_2$ informs us that $X_1$ is conditionally independent of any other random variable given $X_2$ since now it is isolated in the new graph. However, given $X_2$ no other random variable contains other conditional independences since for each $X_i$ with $i = 3, \ldots, 7$ there is still a path to connect to any $X_j$ with $j = 3, \ldots, 7$.

One more advantage of $MRF$ is the amount of research that has been devoted to solve typical inference queries in joint probability distributions that factorize as shown in Eq. (2.28). We refer to inference in the graphical model when we perform any query on the joint probability distribution. One common inference task is to find the joint Maximum A

Posteriori (MAP) assignment, which can be formulated as solving the following problem:

$$\underset{X_1,\ldots,X_N}{\operatorname{argmax}} P(X_1,\ldots,X_N) \tag{2.31}$$

Even with the factorization form in Eq. (2.28), finding the MAP assignment is still NP-hard for a general $MRF$ containing loops in the graph. However, efficient methods to find approximate solutions have been used in practice with great success. Reviewing all of them is beyond the scope of this background chapter, but some common methods include the max-product variants of message-passing algorithms [YFW03], graph-cuts [KZ04, BVZ01, BK04] and linear programming relaxations [SMG$^+$08, KKT09]. In the case of graphs with tree-like structures, such as the Markov chain in Fig. 2.17, these methods return the exact solution.

Another important inference task is to find the marginal distribution for each random variable $X_i$. In this case, the problem can be formulated as follows:

$$P(X_i) = \sum_{X_j;j\neq i} P(X_1,\ldots,X_N) \quad i = 1,\ldots N \tag{2.32}$$

Again, even with the factorization form in Eq. (2.28), calculating marginal distributions is still an NP-hard problem for a general $MRF$ containing loops in the graph. However, the sum-product variants of message-passing algorithms have proved very successful in different real-world applications and it is exact in graphs containing a tree-like structure. Notice that computing marginals is a very different type of query than finding the MAP assignment. One could think of computing all the marginals and then selecting the assignment with highest probability for each random variable. This selection can result in a very different answer than a MAP query, since the latter computes a global assignment that maximizes the joint probability for all the random variables at once. Therefore, it is important understand which information we want to extract from the distribution.

**Related work for alignment**

Once we have the MRF framework that provides and the relationship between graph and factors of the joint probability distribution, and the tools to query these probability distributions, we need to define three aspects that are problem-specific: the structure of the graph $G$ that encodes the conditional independence relations between random variables; the factors $\psi_C$ that encode the constraints or preferred assignments for different subsets of random variables, and the queries that we want answer from the joint probability distribution. As mentioned above, in this thesis we use an $MRF$ framework to solve alignment problems similar to the one used by Anguelov et al. [AScPK04] and Coughlan and Ferreira [CF02], which we will describe in the following paragraphs.

Anguelov et al. present an unsupervised algorithm for registering $3D$ surface scans of body shapes undergoing significant deformations. The objective of the task is to find point-to-point correspondences between two triangulated meshes computed from the surface scans. Each random variable $X_i$ for $i = 1, \ldots, N$ represents a point in mesh A and the set of possible assignment $X_i = \{x_1, \ldots, x_M\}$ represents all the points in mesh B. They optimize the joint probability $P(X_1, \ldots, X_N)$ to find the correct correspondence. Coughlan and Ferreira present a similar approach to detect and localize deformable shapes in $2D$ grayscale images. In particular, they show results for hand writing recognition. There, each random variable $X_i = (p_i, \theta_i)$ encodes a point $(p_i)$ and a normal orientation $(\theta_i)$ in a prior shape learned from training examples. The set of possible assignments $X_i = \{x_1, \ldots, x_M\}$ represents all the pixels in the gray scale image where $X_i$ could be assigned. One of the advantages of framing alignment as an optimization problem over a joint probability distribution is that we are performing a more global search over the set of possible correspondences than the RANSAC greedy approach described above.

Both [AScPK04] and [CF02] use an MRF model that contains only singleton and pairwise factors to model the joint distribution. This subset of $MRF$ is found in many contexts because of its simplicity and the practical ability to capture sufficient dependencies to solve

the problem. In this case, Eq. (2.28) reduces to:

$$P(X_1, \ldots, X_N) = \frac{1}{Z} \prod_{i=1}^{N} \psi_i(X_i) \prod_{(i,j) \in E} \psi_{ij}(X_i, X_j) \tag{2.33}$$

Also, both approaches use variants of the sum-product algorithm to perform approximate inference and calculate the marginals for each random variable. In both cases, they show empirically that selecting the highest probability assignment for each marginal suffices to obtain good results. The main difference between the two methods are in the definition of the singleton and pairwise potentials. We will describe them in the following paragraphs in more detail since they contain ideas that inspired the work presented in this thesis.

Anguelov et al. approach is unsupervised. Therefore, the singleton potentials try to capture whether a local region around a point in mesh A is similar to a local region in a point in mesh B. For this purpose the authors use spin images [JH99], since they have proved very effective for this kind of local discriminative description in $3D$ meshes. However, singleton potentials alone are not discriminative enough to establish correct correspondences. For example, one can imagine that the shape of the surface around the chin is very different from the one around the shoulder but similar to the one around the nose. The pairwise potentials enforce geometric constraints between adjacent points to improve correspondence. Edges in the MRF are only added between points that are close by in the human body. One of the key insights from Anguelov et al. to improve alignment results was using geodesic distances along the body surface instead of Euclidean distances to prune possible correspondence assignments, which it is easy to encode using the $MRF$ framework. For example, even if a hand is close to the hips, they are far apart going along the surface and the framework encodes that information removing edges between locations in the hands and locations in the hips. Once we have the edges, the values of the pairwise potentials are based on local deformations: if in mesh A we need a certain amount of stretching and twisting to fit the local surface coordinates between two adjacent

points, the amount of stretching and twisting needed in the pair of corresponding points in image B should be very similar. These pairwise constraints are similar to the ones use in Chapter 4.

The approach by Coughlan and Ferreira is very similar but adapted to $2D$ deformable shapes. Again, edges in the MRF are only considered between landmarks in the shape that are close to each other. The singleton potentials values are calculated from an edge map for $p_i$ and from an orientation map for $\theta_i$. Briefly, pixels belonging to an edge are more likely to be part of handwritten character, and certain orientations are also more likely to belong to certain characters. As in Anguelov et al., the pairwise potentials impose consistency constraints between adjacent points in the shape. In particular, the relative orientation and the relative position between two points has to be similar to the one estimated in the training set. These pairwise constraints are similar to the ones presented in Chapter 3.

We have seen that the definitions of singletons and pairwise factors is different for each problem and it is one of the key problems to address to obtain good solutions. In other words, since the factors define the probability distribution, if they do not contain meaningful information, finding the high probability assignments of the joint distribution is not going to return any meaningful solution. In the following chapters, we will present how we can use the $MRF$ framework for alignment of $CET$ images, one of the bottlenecks in the current $CET$ pipeline. In particular, the main challenge is to define appropriate singleton and pairwise factors targeted to solve the problem in $2D$ (Chapter 3) and $3D$ (Chapter 4). Along the text, but especially in Chapter 5, we will present results to show the strengths and weaknesses of such a probabilistic framework in the context of $CET$.

# Chapter 3

# Automatic alignment of CET tilt series

Reviewing the $CET$ pipeline in Fig. 2.12, the current bottleneck for real high-throughput tomography is the alignment process. We need to automate the alignment of $2D$ projections using gold fiducial markers if we want to increase the number of datasets available for data analysis.

Alignment is the process that obtains the projection maps from the $3D$ sample space to the $2D$ image space. These projection maps are necessary to reconstruct $3D$ density volumes [Law92, PMBF95] since they represent how each point in the sample has been projected in each image. Typical models solve for shifts, rotations and magnification changes that occurred during the data acquisition phase. Once we have the projection maps we can estimate the $3D$ density volume that generated all the projections. The reconstruction quality can only be as good as the alignment accuracy permits. Thus, it is important to perform the best possible alignment.

The main idea behind alignment is to identify and track common features across different projections. Once we have these correspondences, we can estimate the parameters of the projection maps by solving a fitting problem via different optimization techniques. Reconstructing a $3D$ scenario from multiple $2D$ projections is a common problem in vision

algorithms. Most of this literature can be found by searching for structure from motion algorithms in computer vision [HZ04, TV98, MSKS03]. However, the low signal-to-noise ratio ($SNR$) present in $CET$ samples presents a challenge for tracking methods used in this field.

In this chapter, we describe a new approach using Markov Random Fields ($MRF$) to automatically align $CET$ tilt series with low $SNR$. We apply this approach to align images using gold fiducial markers as high-contrast point-like features to guide the tracking, which is the most widespread type of alignment in $CET$ samples. Since gold particles are spherical, their projection looks like a disk from any angle, and their center always represents the same $3D$ point in the sample, which makes them ideal features to track. The topics presented in this chapter are based on content published in [AMC+08] and incorporated into the RAPTOR software package.[1]

Fig. 3.1 shows a block diagram describing in more detail the basic alignment pipeline. Here is a brief description of each step:

**Pre-processing:** Some basic image pre-processing usually includes hot pixel deletion from the charge-coupled device ($CCD$) and denoising to help detecting features. The denoised images are only used during the alignment but not for reconstruction.

**Coarse alignment:** Automatic data acquisition packages have become very reliable. However, raw tilt series can still contain large shifts and rotations. In general, it is advisable to correct for shifts and rotations using cross correlation between adjacent images [Guc82]. After this step, we can assume that markers have a smooth trajectory from image to image. For example, in the common single-axis tilt acquisition, markers travel roughly on straight lines perpendicular to the tilt axis.

**Detect features of interest:** Once we have templates of the features we want to find, we need to detect those features in multiple images. The most common technique

[1] RAPTOR is joint work with Farshid Moussavi and it is presented here as a whole to keep ideas in perspective. Even if it is hard to separate the work load in two parts, the methodology presented in Section 3.3 was led by Farshid Moussavi and Section 3.4 was led by Fernando Amat.

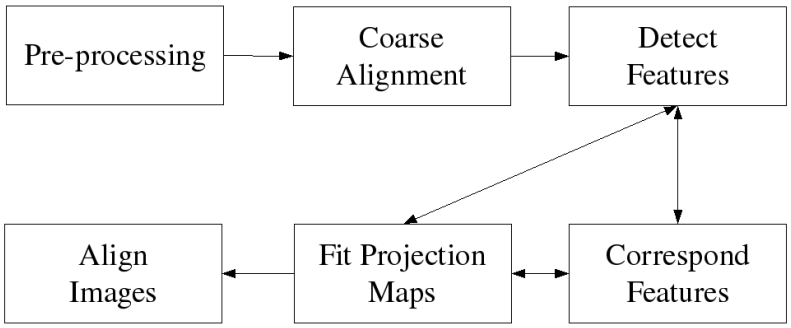


**Figure 3.1:** Alignment pipeline. There are three main steps to align $CET$ samples: detection of features of interest, correspondence of those features across different images and fitting of projection maps. Typical approaches iterate these three steps to improve results.

is to take a patch around the feature and perform the search in other images using cross-correlation. Due to high noise in $CET$ images it is not uncommon to have many false detections. Contextual information is then used to sort out true from false positives.

**Correspond features between images:** This is a key step in the alignment process and will be one of the main focuses of this chapter. Robust methods are critical here to increase success rate in the process.

**Fit projection maps:** Given the $2D$ location of each feature and the correspondence, we can fit a model of how points in $3D$ are projected into $2D$ images. Usually, this step is iterated with the correspondence step to add robustness.

## 3.1 Notation

Chapter 5 by Frank [Fra06] contains a very detailed description of the possible projection maps that can be used for alignment. We will follow its notation for consistency. Here is a summary of the most relevant elements: the nominal beam axis of the electron microscope is the z-axis. $3D$ coordinates are notated as $r_j = [x_j, y_j, z_j]^T$ and $2D$ image coordinates

in the projections as $p_{ij} = [u_{ij}, v_{ij}]^T$, $i = 1 \dots n_P$, where $n_P$ is the number of images and $j = 1 \dots n_T$ is the number of features identified in the sample.

Rotations are define using Euler angles as follows:

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{3.1}$$

$$R_y(\beta) = \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix} \tag{3.2}$$

$$R_z(\gamma) = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.3}$$

Following this notation, the projection map for a single image that approximates the electron microscope as a linear structure projector [Fra06]-ch.3) can be written as:

$$p_{ij} = m_i \; P \; R_z(\gamma_i) \; R_y(\beta_i) \; R_x(\alpha_i) \begin{bmatrix} x_j \\ y_j \\ z_j \end{bmatrix} + \begin{bmatrix} \Delta u_i \\ \Delta v_i \end{bmatrix} \tag{3.4}$$

where $m_i$ is a scaling factor to account for changes in magnification between images, $d_i = [\Delta u_i, \Delta v_i]^T$ represents the image shift with respect to a reference center in the $CCD$ and $P = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right]$ is a projection matrix.

The most common acquisition geometry in $CET$ is a single-axis tilt series: $\gamma$ is taken to be a constant in-plane rotation of the tilt axis and the user changes $\beta$ incrementally to generate the tilt series. $\alpha$ is non-zero only if we suspect that the tilt axis is not perpendicular to the optical axis. Unless stated otherwise, the descriptions presented in this chapter refer to data acquired following single-axis schemes.

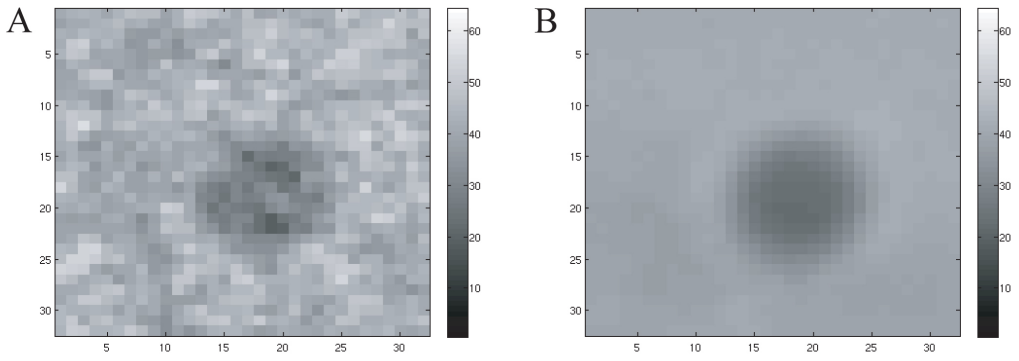


**Figure 3.2:** (A) 10nm diameter gold marker from $CET$ sample. (B) Template for a gold bead after averaging 100 markers. The template is used for marker detection with $NCC$.

## 3.2 Challenges in CET projections of tilt series

When the sample contains fiducial markers, we are looking for high contrast disks of a particular diameter. We generate the gold bead template by averaging a few of them selected from the $CET$ sample itself (Fig. 3.2), since all gold beads should look similar from any projection. Then, normalized cross-correlation ($NCC$) is used for the template matching to avoid errors due to change of illumination from image to image. As a rule of thumb, it is desirable that markers are at least 8 pixels in diameter in each image. Otherwise, detection can be difficult due to high noise in $CET$ images.

There are several problems that need to be taken into account during the marker detection process:

**Contrast reduces with thickness:** Fig. 3.3 shows a clear example of how sample thickness affects contrast in $CET$ images. As explained in Section 2.3, in single-axis tilt acquisition, thickness increases as $1/\cos(\beta)$. At high tilt angles, marker detection can be difficult even for the human eye.

**Markers tend to cluster:** Fig. 3.3 shows a typical example of a cluster of markers. In that situation it is difficult to distinguish the center of each individual marker and to avoid mistakes when corresponding markers from image to image. As a rule of

thumb, it is not advisable to try to track clusters with more than three markers on them.

**Markers might have low $NCC$ scores:** marker appearance can change from one image to the next and even from different regions in the same image. High $NCC$ scores always correspond to real markers, but low $NCC$ scores might be false positives or real markers. Therefore, using only the $NCC$ score does not give a clear indication of good matching. We need extra information (context) to discern good matches.

**Occlusion:** markers can disappear in some images due to occlusion from another object in the sample or because they are out of the field of view after tilting the sample. Therefore, the system has to account for the fact that some markers might not be found in some images.

Adding contextual information to the template matching is critical to make the detection process robust. The main idea is that because marker projections come from approximately a rigid body motion, their spatial location in contiguous images is highly correlated. The next section describes how we add this information in a probabilistic framework.

## 3.3 Pairwise image correspondence with MRF

Our task is to find a marker in a image that correspond to each marker in the current image. To show how local context helps, one only needs to think about mapping stars between two different nights (Fig. 3.4). At one extreme, we could attempt to solve the correspondence of all the stars simultaneously. This approach is very complex since we need to think about all the stars at once. The other extreme, trying to locate each star in isolation is also difficult because stars look too similar to each other (Fig. 3.4A). A good compromise solution is to break the sky up into constellations, locate each one for each night, and then build the global map based on these local mappings (Fig 3.4B- it is difficult to mistakenly find another big dipper). Intuitively, once given a stars constellation,

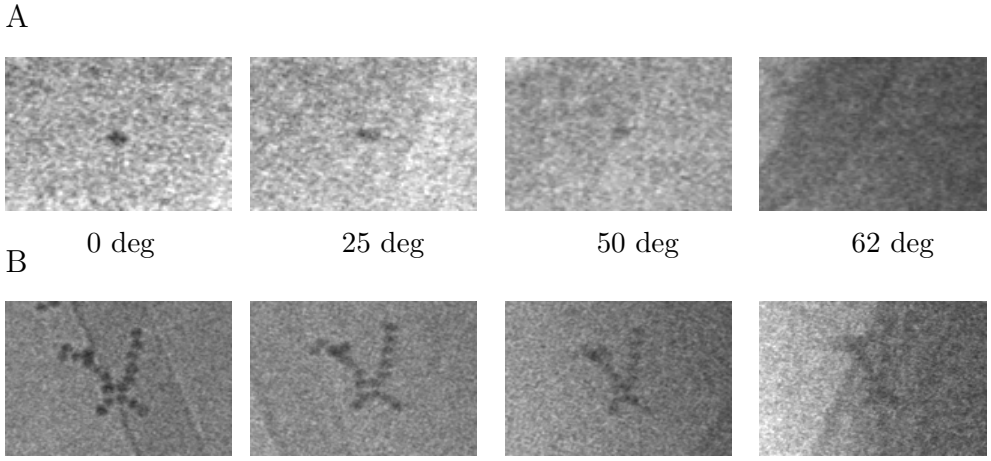


**Figure 3.3:** (A) Example of the same gold bead at different tilt projections. The contrast declines in higher tilt angles due thickness increment. (B) Example of gold markers clumping together. This makes it difficult to find all the centers and correspond each individual marker from image to image. Gold beads are 10nm in diameter. Image by Luis R. Comolli at Lawrence Berkeley Laboratory, California.

there is little information to be gained about that star from all the other stars in the sky. This intuition can be mathematically captured using $MRF$, which is the underlying mechanism of the method presented here. Actually, we will show that only looking at all possible pairwise interactions between nearby stars, is sufficient in our case to obtain a good solution to the correspondence problem.

### 3.3.1 Probabilistic framework

Let $I_1$ be the reference images and $I_2$ be the candidate image. Also, let $A = \{A_1, \ldots, A_M\}$ be the set of markers detected in $I_1$ and let $B = \{B_1, \ldots, B_K\}$ be the set of markers detected in $I_2$. We have $M$ markers $A_i$ for $i = 1 \ldots M$ in the reference image to correspond to a list of $K$ candidate markers $B_j$ for $j = 1 \ldots K$ in the candidate image. Typically, $K$ is taken to be 3 to 4 times larger than $M$ to avoid missing possible correspondences. The markers $A_i$ can be either from raw or prealigned images, and can also incorporate any coarse projection model information if that is available.

We treat each $A_i$ as a random variable that can take a value in the set $B \cup \{\emptyset\}$, where $\emptyset$ indicates that marker $A_i$ has no correspondence in the second image. This detail is very important to tolerate outliers. Our goal is to define the joint distribution

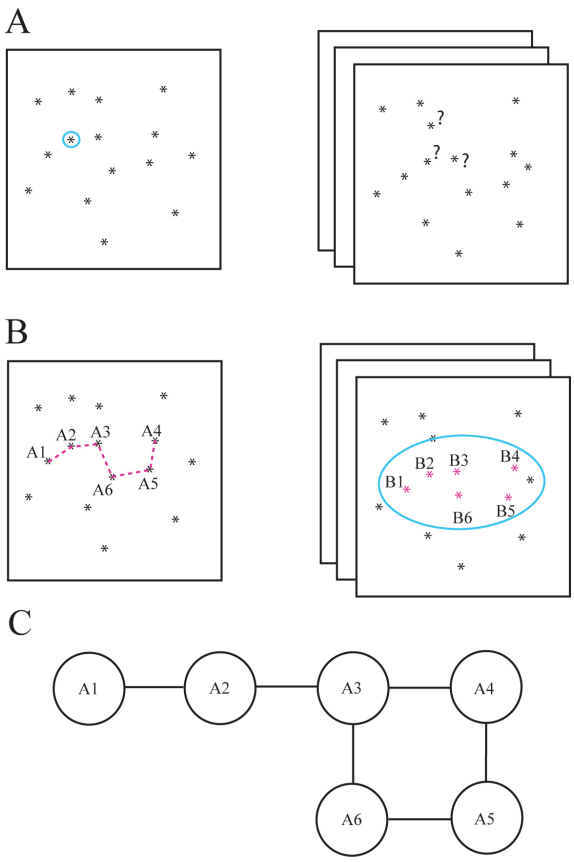


**Figure 3.4:** Local correspondence for one pair of images. (A) Finding the correspondence by looking at a single marker is difficult since it requires identifying a marker in the reference image and choosing one of several (quite similar) candidates in a nearby neighborhood in the candidate image. (B) On the other hand, if a group of markers is considered at a time, the contextual information (e.g. the Big Dipper pattern) can be exploited to find a more reliable correspondence. (C) To represent this context when finding correspondence, we represent local relationships as a dependency graph for MRFs. One possible dependency graph is shown here for the $MRF$ formed by the 6 markers in panel (B).

$P(A_1, \ldots, A_M | I_1, I_2)$ that assigns a probability value for each possible assignment of the markers in $A$. In particular, high probability assignments should indicate good correspondence between elements of $A$ and $B \cup \{\emptyset\}$. Therefore, we want to find the assignment that maximizes the following joint distribution:

$$A^* = \underset{A_i}{\operatorname{argmax}}\, P(A_1, \ldots, A_M | I_1, I_2) \tag{3.5}$$

We use the $MRF$ formalism presented in Section 2.4.3 to model the conditional probability between random variables in $A$: a pair of markers depend on each other directly, i.e., have an edge in the graph, if they are within some distance in the $2D$ projection. A graphical example of such a construction is shown in Fig. 3.4C and 3.5. To further contain computational complexity, we limit our factors to be over at most two markers. This construction factorizes the joint probability as follows:

$$P(A_1, \ldots, A_M | I_1, I_2) = \prod_{i=1}^{M} \psi_i(A_i) \prod_{(i,j)\in\varepsilon} \psi_{ij}(A_i, A_j) \tag{3.6}$$

where $\psi_i$ is the singleton potential or data term, and encodes how likely the candidate marker $B_r$ matches the reference $A_i$ using information only about the reference itself. $\psi_{ij}$, the pairwise potential, encodes how likely it is that a pair of linked reference markers $A_i$ and $A_j$ match a pair of markers $B_r$ and $B_s$ in the candidate image. These potentials are described in the coming subsections.

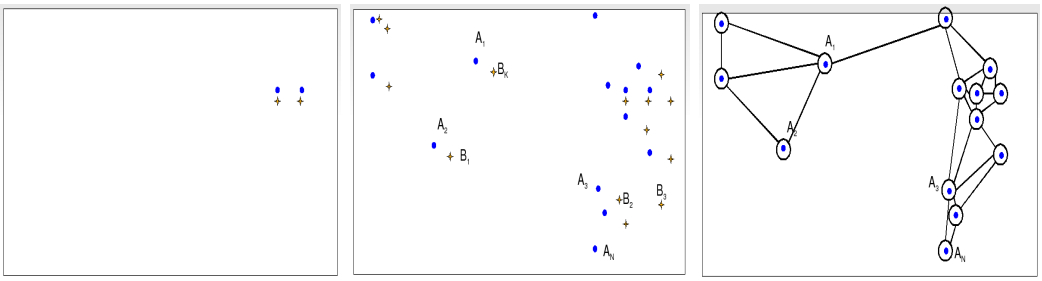


**Figure 3.5:** Blue points indicate putative location of markers in image A. Yellow stars indicate the same for image B. Left: when only looking at a local neighborhood it is easy to make correspondence mistakes. Middle: correspondence is clear looking at the two datasets. Right: construction of the $MRF$ graph to perform inference.

Even with the factorized distribution, exact inference of marginal and conditional

probabilities needed for correspondence is NP-hard. In this thesis, we use a variant of Loopy Belief Propagation (LBP) [EMK06] that has fast convergence properties to solve the problem. LBP is an approximate inference method that has been shown to converge to good results for $MRF$ problems in imaging applications [MWJ99] and in correspondence problems in particular [TKR08]. For a detailed description of LBP the reader is referred to [Bis07, YFW05].

### 3.3.2 Singleton potentials

The singleton potential is based on physical similarity, and is calculated by cross-correlating a square patch of fixed size centered at a marker extracted from the reference image with the same size patch centered at the candidate marker (Fig. 3.6). This score is then multiplied by an exponential roll-off factor which is a function of distance from the center of marker $A_i$ to candidate marker $B_j$. More formally:

$$\begin{aligned} \psi_i(A_i = B_r) &= NCC(A_i, B_r) \cdot f_{\sigma_1}(x_i, x_r) \\ f_\sigma(\overrightarrow{x}, \overrightarrow{y}) &= \exp\{-\frac{1}{2}\frac{\| \overrightarrow{x} - \overrightarrow{y} \|^2}{\sigma^2}\} \end{aligned} \tag{3.7}$$

where $NCC$ is the normalized cross-correlation between the two patches around the markers, $x_i$ is the position vector of the marker represented by $A_i$ and $x_r$ is the position of the marker represented by $B_r$. The exponential roll-off is just a function to disregard impossible correspondences between markers that are in opposite sides of the image. $\sigma_1$ measures the expected maximum shift between adjacent images: if images are perfectly prealigned, then $\sigma_1$ should correspond to the maximum displacement of a marker in a tilt series ($\sigma_1 = 0.5 \cdot 2048 \cdot (\cos(65°) - \cos(63°)) = 32$ pixels for a 2Kx2K $CCD$ with maximum tilt of 65°). In practice, as we discuss below, $\sigma_1$ is not a critical value and a value equal to 1/4 of the image size for raw datasets and 1/8 for prealigned images delivers good results.

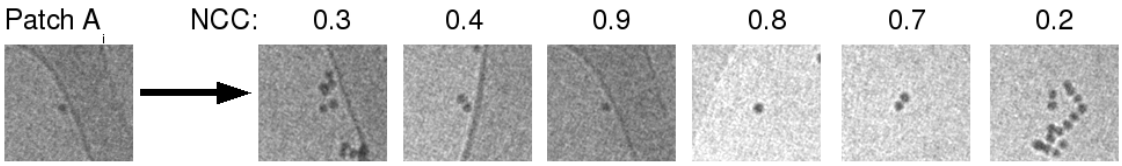


**Figure 3.6:** Example of computation of singleton potentials using $NCC$. The template matching helps differentiating between clustered markers and isolated markers to guide the approximate inference.

### 3.3.3 Pairwise potentials

The pairwise potential is based on relative orientation and position between two markers (Fig. 3.7). Defining vector $\overrightarrow{v_{ij}} = x_i - x_j$, and vector $\overrightarrow{v_{rs}} = x_r - x_s$, the difference vector, $\overrightarrow{v_{ij}} - \overrightarrow{v_{rs}}$ is a measure of the relative motion between marker $i$ and marker $j$ between the reference image and the candidate image. Formally, we can define this as follows:

$$\psi_{ij}(A_i = B_r, A_j = B_s) = f_{\sigma_2}(v_{ij}, v_{rs}) \cdot f_{\sigma_1}(x_i, x_r) \cdot f_{\sigma_1}(x_j, x_s) \tag{3.8}$$

The main assumption in Eq. (3.8) is that two adjacent images represent similar projections. Therefore, the relative position between neighboring markers should be similar. This spatial constraints are applicable to different acquisition schemes besides single-axis tilt, since they make no assumptions about the projection model.

As in Eq. (3.7), $f_{\sigma_1}$ disregards impossible correspondences. The reason to include this effect in both the singleton and pairwise potentials is empirical. During the development of RAPTOR, we tested correspondence results for different pairs of images from the datasets shown in Fig. 3.14, including $f_{\sigma_1}$ only in the singleton potentials or double counting the roll-off factor by including it in the pairwise potentials. We found that if we do not double count $f_{\sigma_1}$, then the parameter $\sigma_1$ needs to be adjusted for different datasets to deliver optimum results. However, if we double count $f_{\sigma_1}$, markers with more neighbors around them will have a tighter roll-off constraint, because they have higher number of edges in the $MRF$. In practice, this effect allows to set $\sigma_1$ constant for all datasets without

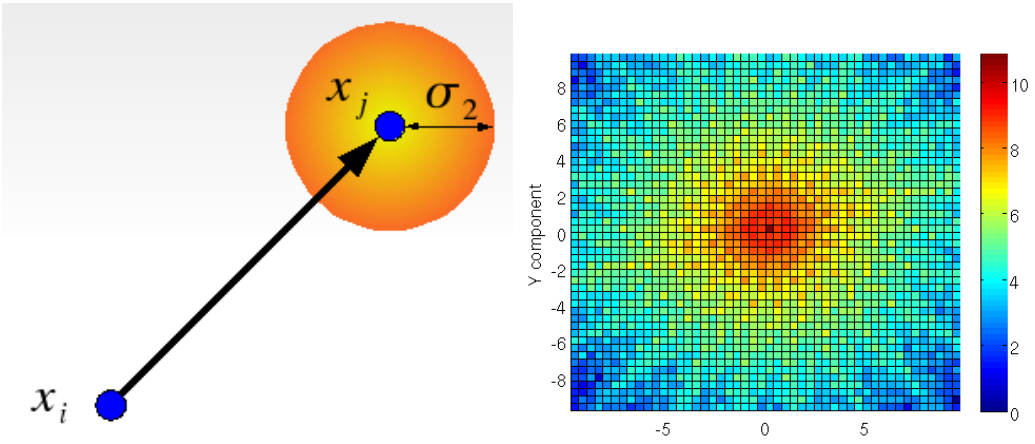


**Figure 3.7:** Left: example of pairwise potential using relative position constraints. If marker $A_i$ and marker $A_j$ have positions $x_i$ and $x_j$ respectively, we can define the vector $\overrightarrow{v_{ij}} = x_i - x_j$. Then the vector between marker $B_r$ and marker $B_s$ should be within the orange ball if $A_i$ to $B_r$ and $A_j$ to $B_s$ generates a good correspondence. Right: $2D$ histogram of vectors $\overrightarrow{v_{ij}} - \overrightarrow{v_{rs}}$ obtained from 10 tilt series aligned by hand using IMOD [KMM96]. Colormap represented in logarithmic scale due to large number of samples. Statistics provide estimates for $\sigma_2$ parameter in 2Kx2K images.

sacrificing performance.

The only thing left is to set a value for $\sigma_2$. This parameter represents a stiffness constraint if we understand the $MRF$ as a graph with springs attached between neighboring pixels. In other words, it models how much deformation we are willing to tolerate between images. We can use previously aligned tilt series to estimate the value of $\sigma_2$. Fig. 3.7 shows statistics taken from 10 datasets from a 2Kx2K $CCD$ camera at 20,000X magnification. We can measure the covariance matrix from the histogram $\Sigma = \left[\begin{smallmatrix} 11.477 & 0.4928 \\ 0.4928 & 13.021 \end{smallmatrix}\right]$ and set $\sigma_2 = 0.5\sqrt{11.47 + 13.02} = 3.5$. It only needs to be changed if we deal with images of different sizes and different magnification. However, the $CCD$ of most available manufacturers follows powers of two, so we just need to calibrate $\sigma_2$ for 1024, 2048 and 4096 sizes and a few standard magnifications used in $CET$.

## 3.4 Global correspondence

While the pairwise correspondence now makes use of all the markers in the two images, it still only considers two images at a time. To generate marker trajectories more robustly across the tilt series, we correspond each image with several of its following images. We use the redundant information both to allow trajectories to skip over a missing marker in some images, and to reduce the probability that a bad correspondence will affect the global trajectories.

### 3.4.1 Global trajectories

The first step towards global correspondence is to build trajectories. To do this we use a multi-level approach. The first level contains correspondences between the $i$-th image and the $i+1$-st image. The second level contains the correspondences between the $i$-th and the $i+2$-th image. We use up to 3 levels in images spaced up to 2 degrees apart because after that it is difficult to find correspondences due to the tilt difference between projections. We initialize the first trajectory as the first point in image 1 and level 1. We call this point $p_{11}$. Local correspondence gives us the location of this trajectory in image 2, i.e. $p_{12}$. Next, we try to find $p_{12}$ in the local correspondence between image 2 and image 3. If we find it, that returns $p_{13}$ and we repeat the process for this new location to find $p_{14}$ (Fig. 3.8).

We use the extra levels in two ways. First, if a local correspondence for a track is missing due to occlusion or being outside the field of view, we look in the next level of the global correspondence to complete the track. For example, if we cannot find $p_{13}$ from $p_{12}$ in the first level, we try to find $p_{14}$ from $p_{12}$ in the second level, which contains local correspondence between image 2 and image 4. If we find $p_{14}$, we continue the process for this new location to try to extend the trajectory (Fig. 3.9). If we do not find it, we try to find $p_{15}$ from $p_{12}$ in the third level, which contains local correspondence between image 2 and image 5. We iterate this process until the first trajectory can not be extended any longer, either because we reach the last image (full trajectory) or because one location can

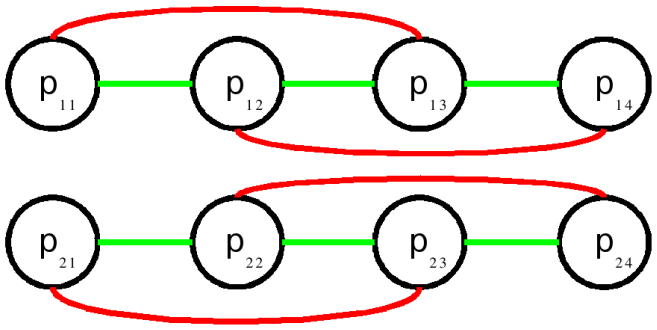


**Figure 3.8:** Schematic with two markers and four images to exemplify how RAPTOR builds global trajectories from pairwise correspondences. $p_{ij}$ represents the $i$-th in image $j$. Green edges indicate pairwise correspondences found between adjacent images (level 1). Red edges indicate pairwise correspondences found between the $i$-th image and the $i+1$-st image (level 2). Level 3 is not shown for clarity. In the diagram presented in this figure, the first global trajectory would be $\{p_{11}, p_{12}, p_{13}, p_{14}\}$ and the second would be $\{p_{21}, p_{22}, p_{23}, p_{24}\}$, since all the correspondences were found at level 1 and all level 2 correspondences verify the trajectory.

not be found in any level (partial contour). To create new trajectories, we use points in each level that have not been used in previous trajectories, since they are potential seeds of new trajectories.

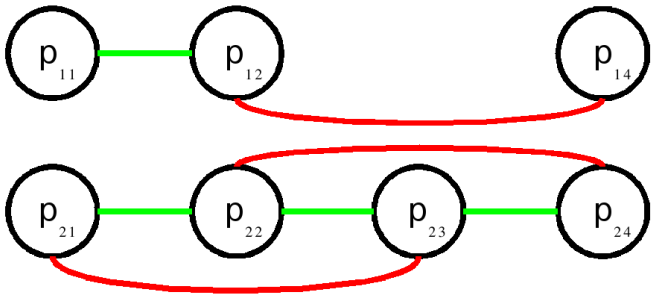


**Figure 3.9:** Notation is the same as in Fig. 3.8. In the diagram presented in this figure, the global trajectory starting with $p_{11}$ is not straight forward because we are missing the marker at $p_{13}$ due to occlusion. However, we can still build the following trajectory $\{p_{11}, p_{12}, \varnothing, p_{14}\}$
using the correspondences at level 2

Second, we also use the extra levels to add confidence in the trajectories, by using the redundant information contained in the local correspondences that have not been used in the first step. For example, suppose we found $p_{12}$, $p_{13}$ and $p_{14}$ using in the first step. We

can check this path against the local correspondence for $p_{12}$ in image 4, which is contained in the second level (Fig. 3.10). Then, we count the number of times such comparisons fail in each trajectory and if this occurs more than 20% of the time of this particular trajectory, we drop the trajectory. Otherwise, we retain the consistent part of the trajectory.

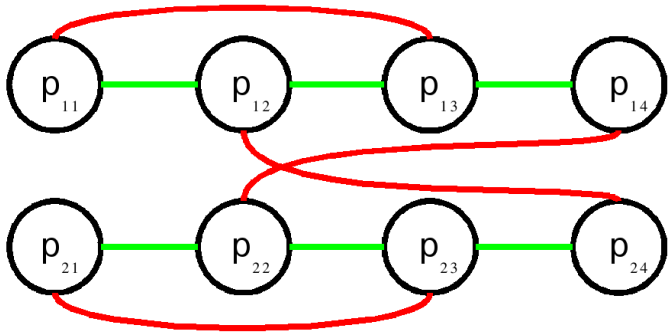


**Figure 3.10:** Notation is the same as in Fig. 3.8. In the diagram presented in this figure, the level 2 pairwise correspondences for $p_{12}$ and $p_{22}$ do not agree with the trajectories built using level 1 correspondences. In this case $p_{14}$ and $p_{24}$ are marked as an uncertain point in the trajectory.

Finally, only trajectories of certain length are kept to estimate the projection model. The length is selected automatically to assure that we have enough points in each image to estimate the projection model. As explained in the next section, the projection estimation is able to handle partial trajectories of any length across the tilt series. This is extremely useful at high tilt angle, where many markers are present in only a few images before disappearing.

### 3.4.2 Robust structure from motion

Once we have recovered trajectories across the whole tilt series, a linear projection model can be estimated to represent the transformation between the $3D$ location of the markers and their $2D$ projection in each image. The projection model finds a common origin of coordinates for all the images in order to align them.

We use convex optimization techniques [BV04] to estimate the linear projection model in Eq. (3.4). In particular, we solve the following minimization problem:

$$\underset{G_i,\Delta u_i,\Delta v_i,x_j,y_j,z_j}{\operatorname{argmin}} \sum_{j=1}^{n_T}\sum_{i=1}^{n_P} w_{ij}\ \rho\left(p'_{ij} - G_i \begin{bmatrix} x_j \\ y_j \\ z_j \end{bmatrix} + \begin{bmatrix} \Delta u_i \\ \Delta v_i \end{bmatrix}\right) + \lambda \|G_i - Go_i\|_2^2 \tag{3.9}$$

where $p'_{ij}$ are the marker coordinates found by correspondence, $G_i$ is a 2x3 matrix projection matrix and $\Delta u_i$, $\Delta v_i$ are image shifts. The unknowns are the parameters of the microscope $(G_i, \Delta u_i, \Delta v_i)$ and the $3D$ locations of the markers $(x_j, y_j, z_j)$. The equation is clearly non-linear and non-convex in those unknowns. However, if we fix $G_i$, Eq. (3.9) is linear and convex in the remaining unknowns, and vice versa. We can use an iterative scheme to solve only linear equations. First we solve for $G_i$, and then we solve for $x_j, y_j, z_j$, and $\Delta u_i, \Delta v_i$. We iterate this process until convergence. This method has been known for a long time in the optimization literature as alternate convex programming [KK05] and it gives similar performance as non-linear optimization methods with the advantage that is computationally simpler to solve.

The term $w_{ij}$ in Eq. (3.9) contains the weights for each residual of the re-projection model to allow the incorporation of partial trajectories in the re-projection model estimation. If a point is missing in the j-th trajectory for the i-th image, it is assigned a weight $w_{ij} = 0$. Otherwise, we set $w_{ij} = 1$.

In order to keep the alternative convex programming as a linear optimization problem, we model $G_i$ as a 2x3 matrix with no rotation restrictions and we impose regularization with respect to the ideal microscope pose. The term $\lambda\|G_i - Go_i\|_2^2$ in Eq. (3.9) performs the regularization. It indicates that the matrix $G_i$ should not be very different from a given matrix $Go_i$ which is set by the geometry of the microscope. The parameter $\lambda$ indicates how much we penalize the fact that $G_i$ deviates from $Go_i$. $\lambda$ has been tuned using cross-validation with different datasets and it is a fixed parameter.

The matrix $Go_i$ that is used as the reference is $Go_i = R_z(\gamma)\cdot R_y(\beta_i)$. This regularization allows the automatic alignment of images even in tilt series where very few markers are present, since it imposes constraints on the shape of $G_i$. These constraints indicate that $G_i$

should be close to a composition of rotations (at least for an ideal microscope). However, it does allow some flexibility to incorporate affine effects such as changes of magnification between different projections or skewing.

Finally, a crucial part of Eq. (3.9) is the cost function $\rho(r)$, which measures the error between the trajectory given by global correspondence and our re-projection model. The definition of $\rho(r)$ has a great effect in the estimation of the re-projection model. All the existing alignment methods use $\rho(r) = r^2$ [LBPE06, BHE01, KMM96] except for [BZ06]. This quadratic error measure establishes that the solution to Eq. (3.9) is the well-known least-squares problem. However, it is also well-known in the optimization literature that least-squares methods are not robust to outliers. Even a single error in a trajectory can bias the entire re-projection model. Thus, a more robust penalty function is used in this thesis. We use what is known as Huber penalty function [Hub81] in robust statistics and is defined as follows:

$$\rho(r) = \begin{cases} \frac{1}{2}r^2 & \text{for } |r| \leq \delta, \\ \delta\left(|r| - \delta/2\right) & \text{otherwise.} \end{cases} \tag{3.10}$$

This penalty function uses a combination of linear and quadratic penalties: for small residuals the penalty is quadratic, and small residuals are thus fitted as least squares. Big residuals, however, are penalized using a linear function, which results in large deviations (outliers) not affecting the re-projection estimation model. The fact that large deviations are tolerated and isolated in the linear region makes it easier to detect and correct them automatically without manual intervention. Notice that the Huber penalty is still convex, so we avoid the problem of finding a local minimum as a solution to our problem. Since we know the gold bead diameter and the pixel size for each tilt series, we set $\delta$ equal to half the diameter of a gold bead (in pixels) because we expect errors locating each bead center due to noise in the images.

### 3.4.3 Iterative refinement

As shown in Fig. 3.1, it is typical for alignment algorithms to iterate the following steps in order to improve results: marker detection, marker correspondence and fitting projection maps. In RAPTOR, once we have global trajectories and the parameters of the projection maps from Eq. (3.9), we iterate over the following steps to improve alignment results:

1. Remove outliers automatically
2. Fill missing markers in global trajectories using the projection maps and singleton potentials
3. Refit parameters of projection maps using filled trajectories

As mentioned in the previous section, the robust penalty allows us to fit the projection model even in the presence of outliers. Even more, we can analyze the residual values for each marker in each image to automatically detect outliers. In RAPTOR, we use the following *ad hoc* procedure to solve this task: first, any residual above two times the diameter of the gold bead is considered an outlier. Second, we analyze residuals for each trajectory separately because the fitting procedure tends to overfit areas of the image containing more tracked gold beads. Thus, we want to separate residuals from different trajectories in order to avoid considering isolated markers as possible outliers. Defining the set of residuals from the $j$-th trajectory as $r_j = r_{1j}, \ldots, r_{n_P j}$, we can sort them in ascending order to generate the vector $\tilde{r}_j = (\tilde{r}_{1j}, \ldots, \tilde{r}_{n_P j})$, where $\tilde{r}_{ij} < \tilde{r}_{i+1j}$ for $i = 1, \ldots, n_P$. Since generally the higher the tilt angle the larger the residual due to image quality degradation, the difference between two consecutive residuals $d_i = \tilde{r}_{i+1j} - \tilde{r}_{ij}$ tends to be a positive constant. However, in the presence of an outlier, the finite difference $d_i$ contains a sudden jump. We detect this by calculating the median of $d_i$ for $i = 1, \ldots, n_P$ and considering an outlier any value above three times the median.

Once we have removed outliers, we want to fill missing markers in the global trajectories. At this point, we have a a very good estimation where those missing markers should be located since we have a complete projection model. We use the $NCC$ score described for

the singleton potentials in Eq. (3.7) in combination with the projection model as follows: given the reprojection point $p'_{ij}$ for the $i$-th marker in the $j$-th trajectory, we calculate $NCC$ using a patch around the point from the previous image $p_{i-1j}$ as a template. If we obtain an $NCC$ score above 0.5 closer to $p'_{ij}$ than twice the gold bead diameter we add that point as the new location of $p_{ij}$. The steps of outlier detection, filling missing markers and recalculating the parameters of the projection maps are iterated until no more new markers are found and until no more outliers are detected.

## 3.5 Results

All the steps described above have been incorporated into the software package Robust Alignment and Projection Estimation for Tomographic Reconstruction (RAPTOR). It has been freely available since 2006 and it is being used in multiple laboratories such as Lawrence Berkeley Laboratory and California Institute of Technology, allowing us to test RAPTOR on hundreds of different datasets and to refine the methodology. Moreover, in combination with other automations in the $CET$ pipeline, it has enabled real high-throughput tomography: a recent review paper [LJ09] shows how it is possible to collect and reconstruct 650 $CET$ volumes in one month. Five years ago, this number of tomograms would have been impossible even if one spent a year on the task.

This section presents RAPTOR tests performed on thirteen $CET$ datasets of thick bacteria whole cells of *Deinoccocus grandis* and *Caulobacter crescentus* to illustrate the main strengths and limitations of this alignment approach. These datasets were chosen from a much larger set of tilt series acquired by Luis R. Comolli and the total dose of at most 100 $e^-/Å^2$, which results in very noisy individual images.

The robustness of RAPTOR predicting marker trajectories is illustrated in Fig. 3.11. Panel (A) shows the position of a marker that will intersect a whole group of markers as the tilt series proceeds. In panel (B) the projected positions of all markers are tightly clustered. At higher angles, the two groups are again resolved, as shown in panel (C). The algorithm successfully tracks the position of markers in difficult situations such as this

one, which can occur quite frequently. The robustness in these situations is a consequence of our incorporation of contextual information in the correspondence. Panels (D) to (F) illustrate the robustness built into RAPTOR using robust convex optimization techniques. Panel (D) shows the position of a chosen marker which is placed in the wrong location by global correspondence in a subsequent image, panel (E). At a subsequent step, however, the trajectory is recovered, as shown in panel (F). This example shows that the projection model is not biased by the presence of outliers. This feature allows the automatic removal of outliers and their re-assignment of their correct location. Methods based on least squares fail in such cases.

Fig. 3.12 illustrates the overall performance of the algorithm in the alignment of a difficult data set of *Caulobacter crescentus*. This data set was acquired with a Helium-cooled stage, and as a consequence the relative position of the markers was significantly more variable than is normally the case with liquid nitrogen-cooled stages [CD05]. Panel (A) of Fig. 3.12 shows the projection acquired at -58 deg. Panel (B) shows the set of trajectories recovered by RAPTOR. There are complete trajectories across the whole tilt series as well as partial trajectories starting from both ends of the tilt series, disappearing at lower angles. The meniscus formed by the cell and bars of the cryo-grid may occlude, at high tilt angles, markers chosen at low angles. At specific angles throughout the tilt series, the overlapped projection of groups of markers which belong to different planes within the cryo-grid may cause gaps in their tracked trajectories. Many gaps also represent locations where the probability of making a correct marker assignment is not sufficiently high. Finally, in the high angle projections there are many markers with ideal $SNR$ which are left out of the view in the projections at lower angles. All these events are represented in the diagram plotted in Fig. 3.12B.

The physical trajectories of all chosen markers in image space (x-y coordinates in pixel number), in the raw data and the aligned data, are shown in Fig. 3.12C and 3.12D. The inset in each panel shows one single trajectory. No pre-alignment of the raw data set was performed before obtaining the automatically aligned one, as is readily obvious from the

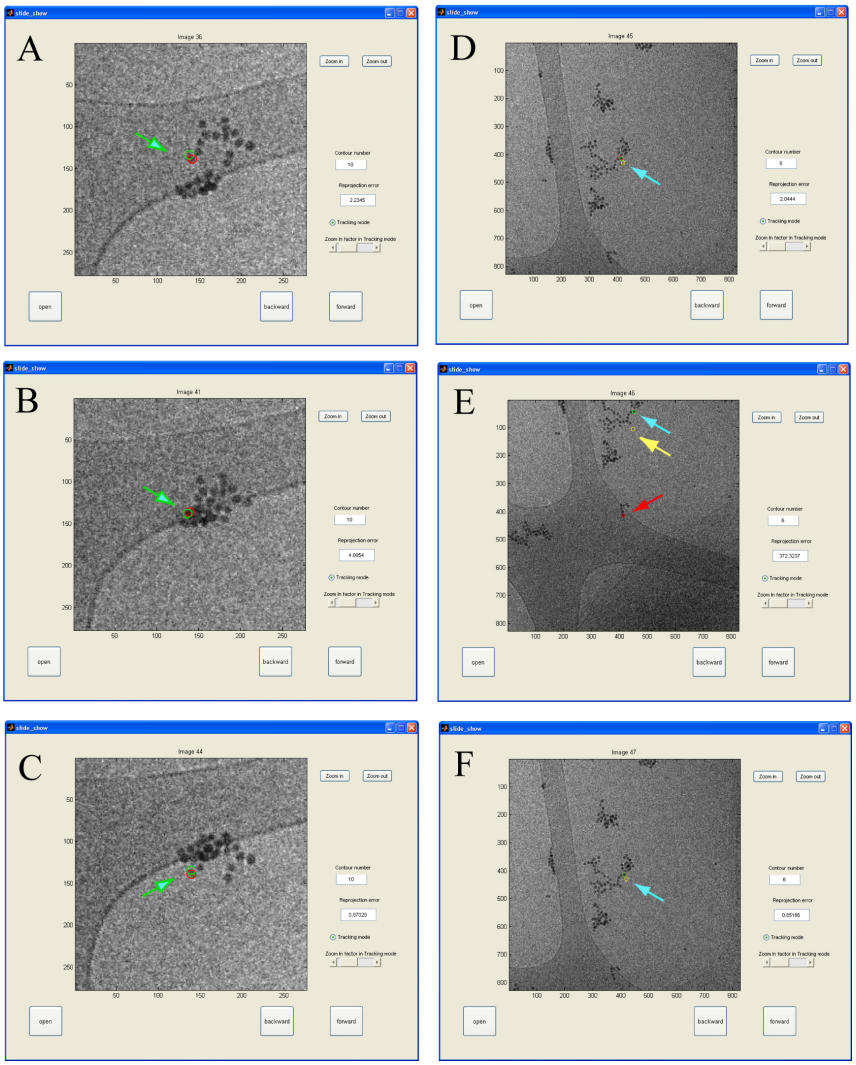


**Figure 3.11:** Robustness of the local and global correspondence. All images are magnified in sections from images in the actual tilt series. (A), (B), and (C) show the tracking of a chosen marker whose trajectory in $2D$ projections passes through a group of similar markers throughout the tilt series. This situation is particularly difficult for methods which track single markers at a time. The $2D$ projections (A), (B), and (C) are spaced by angular increments of more than 6 deg, for better visualization. Green arrows point to the reprojection position. Panel (D) shows the position of a chosen marker whose trajectory is lost in (E) (red arrow) , and subsequently recovered in (F). Yellow arrow points to reprojection position using least-squares. Blue arrow points to reprojection position using Huber penalty.

cloud shape of marker trajectories in the raw data (Fig. 3.12C).

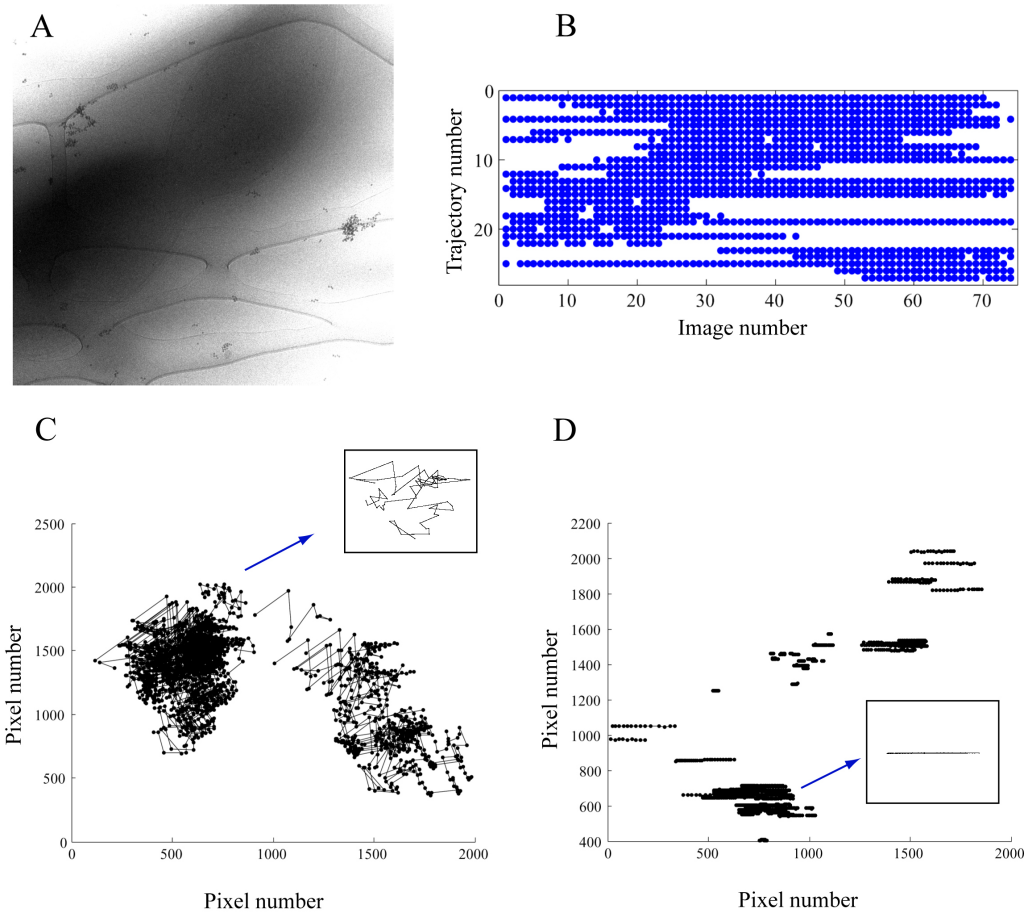


**Figure 3.12:** RAPTOR performance tracking markers in a *Caulobacter crescentus* dataset. (A) The -58 deg $2D$ original projection from the tilt series. The size of the image is 2048 pixels on edge. The cell forms a meniscus of considerable thickness. Both the cell and the grid bars may occlude markers at high tilts. The wider area projected into the images at these angles contains many markers with appropriate $SNR$, far from the cell, which disappear at lower angles. (B). Trajectories recovered from the tilt series. (C) Overlay of several marker locations in image space (x-y coordinates in pixel number) throughout the raw datasets illustrates the trajectories in the raw data set, (D) trajectories of same markers after RAPTOR automatic alignment. The insets in (C) and (D) show the trajectory of a single marker.

Once we have assessed that RAPTOR tracks markers correctly, we want to use a quantitative resolution metric to compare alignment precision between RAPTOR and an expert user of IMOD [KMM96], which is one of the most commonly used software

packages to align CET images. For this purpose, we use a standard method in *CET* called Noise-compensated Leave-One-Out (NLOO) attributed to Cardone et al. [CGS05]. NLOO estimates alignment resolution by comparing the original 2D projection in the raw data with the reprojection of the tomogram calculated from all the other projections. For example, if image number 10 was obtained at $+20°$, we obtain a tomogram with all the projections available except for number 10, and then we reproject that tomogram at $+20°$. If the alignment is correct, the original image obtained from the microscope and the reprojected one should be very similar. In order to compare 2D images we use a standard metric in electron microscopy named Fourier ring correlation (FRC), which can be formulated as follows:

$$FRC_{F,G}(k) = \frac{\sum_{m,n \in R(k)} Re\{F_{m,n}\bar{G}_{m,n}\}}{\left(\left(\sum_{m,n \in R(k)} \|F_{m,n}\|^2\right)\left(\sum_{m,n \in R(k)} \|G_{m,n}\|^2\right)\right)^2} \tag{3.11}$$

where $F$ and $G$ are the Fourier transform of each image, $m, n \in R(k)$ are all the Fourier coefficients with radius $k$ and $\bar{G}$ is the conjugate value of G. In words, we are calculating a normalized cross correlation for each spatial resolution $k$. Therefore, in graphs such as the one in Fig. 3.13, the best possible curve would be one with a constant value of 1 at all spatial frequencies. However, it is typical that as we go to higher frequency rings, the correlation coefficient decays. In particular, researchers tend to look at two points in the curve to estimate resolution: when the FRC reaches a score of 0.5 and of 0.3.

Two final points remain to fully describe NLOO: how to account for the noise in CET images in order not to bias the estimation and how to compare all the projections in one tomogram. The first one is relatively straight forward since we just need to normalize the FRC score in Eq. (3.11) by the comparison between the original 2D raw data with the reprojection using a tomogram reconstructed from all the slices. In this case, the difference between reprojection and raw data can only come from the noise statistics. The second point is also straightforward: since we can not choose one projection over the other, we

compute the FRC score with noise normalization leaving one image out at a time and then averaging the results over all the images. Fig. 3.13 panels (C) and (F) show the resolution plots based on the NLOO metric for two datasets. The plots show how RAPTOR obtains the same resolution as the manual reconstruction in both cases.

Fig. 3.14 shows the different resolutions using the 0.3 and 0.5 thresholds for NLOO metric after aligning 13 datasets automatically with RAPTOR and with an expert IMOD user. The datasets use different angular steps between projections, as well as different marker diameters and a different number of markers present in each image. RAPTOR resolution is always comparable, or even better in some datasets, than these obtained by an expert user with extensive manual intervention, demonstrating that RAPTOR obtains full-precision alignment in a variety of difficult scenarios.

The $SNR$ numbers in Fig 3.14 provides an estimate of the feature visibility in the raw datasets. We chose a signal to noise measure that reflects the visibility of the gold particles in the images as defined in [Fra96]:

$$SNR(marker) = \frac{\sum_{i,j} |\overline{m_{ij}}|^2}{\sum_{i,j} \sum_k |m_{kij} - \overline{m_{ij}}|^2} \tag{3.12}$$

The measure of the $SNR$ is based on the average of all the gold bead images as shown in Fig. 3.2. The signal power is the variance of the average gold image, and the noise power is the variance of all the gold images with respect to this average image. $m_{kij}$ is the value of the (i,j)-th pixel in the k-th image, and $\overline{m_{ij}}$ is the value of the (i,j)-th pixel in the zero mean average image. Notice how all the datasets have $SNR$ well below 0dB.

For all data sets, we also compared visually the quality of the tomographic reconstructions obtained after RAPTOR alignment with those obtained after alignment by an expert IMOD user. Two such comparisons are shown in Fig. 3.13. Panels (A) and (B) show, side by side, a one-pixel-thick slice through a tomographic reconstruction of a dividing *Cc* cell obtained after manual alignment and after RAPTOR automatic alignment, respectively. The quality of both is equivalent. Another example is shown in Fig. 3.13 panels (D) and

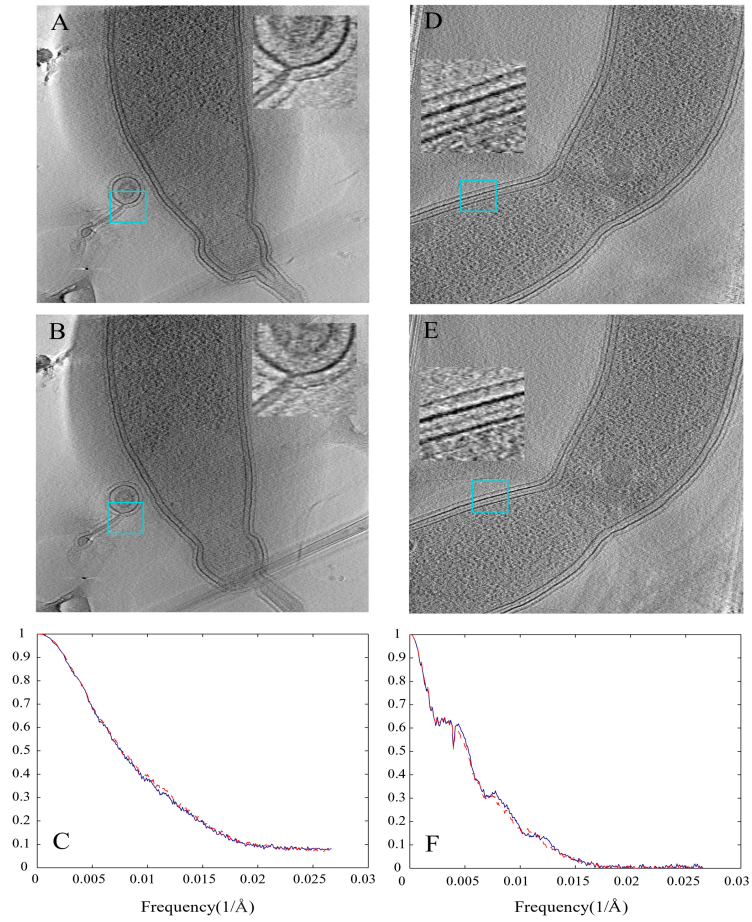


**Figure 3.13:** Tomographic reconstructions based on RAPTOR and manually aligned data sets. (A), (B), one pixel-thick slices from tomographic reconstructions of a Caulobacter crescentus. polar mutant, obtained from a manually aligned data set and from an automatically aligned data set, respectively. (D), a one pixel-thick slice from a tomographic reconstruction of a dividing Caulobacter crescentus. cell after manually aligning the tilt series, and (E), after RAPTOR automatic alignment. Reconstruction after marker tracking by hand and human-directed alignment, (A) and (B), and after automatic alignment, (D) and (E), results in equivalent final quality. The images are sections, 900 pixels on edge, from tomographic reconstructions binned from 2048 to 1024 pixels on edge. The boxes shown in the insets measure 84 pixels on edge in the binned images. Panels (C) and (F) show quantitative resolution comparisons for tomographic reconstructions of the two datasets above using noise-compensated leave-one-out method in [CGS05]. The resolution curves obtained with RAPTOR, in red, and with human-directed alignment, in blue, are equivalent.

(E). The features observed in a one pixel-thick slice through a tomographic reconstruction are comparable in definition and resolution, whether the data set was aligned with human intervention or automatically by RAPTOR. A better comparison is established by careful inspection of regular features, as shown in the insets. These reconstructions were binned by two, from an image size of 2048 pixels on edge to 1024, and each box measures 84 pixels on edge.[2]

## 3.6 Discussion and limitations

Making RAPTOR freely available has allowed us to test it in almost any possible tomography dataset, from cryo-samples to plastic sections or soft X-rays, and from samples with only three markers to samples with hundreds of them. This data has provided information about areas where RAPTOR could be improved.

The first area of improvement is in samples where there are more than 100 markers in the image, where RAPTOR tends to track short overlapping trajectories that do not provide high accuracy alignment. Fig. 3.15 exemplifies the main issue in these situations showing an extreme case with a sample containing over 400 markers. If we try to track 400 markers, only 25% of the marginal probabilities returned by correspondence contain a clear assignment, so RAPTOR only retains 130 markers for this pairwise image. On the other hand, if we try to track 80 markers, most correspondence assignments are clear winners, but we are missing many markers because we were not looking for them in the first place. In both situations the end result is the same[3]: RAPTOR recovers short partial trajectories because in each pair of images the correspondence returns clear winners for a different set of markers, which lowers the final alignment accuracy because the trajectories are not globally consistent. This is especially important in plastic embedded sections, where global distortions are more common than in *CET* samples. Incorporating the projection

[2]Notice that it is impossible to obtain exactly equivalent slices from these reconstructions because the orientation of the reconstructed volume is not identical in data sets aligned by hand or automatically. The quality comparison of the reconstructions seems clear and straightforward nonetheless.

[3]In practice, there is higher chances of success setting $M = 80$ than $M = 400$.

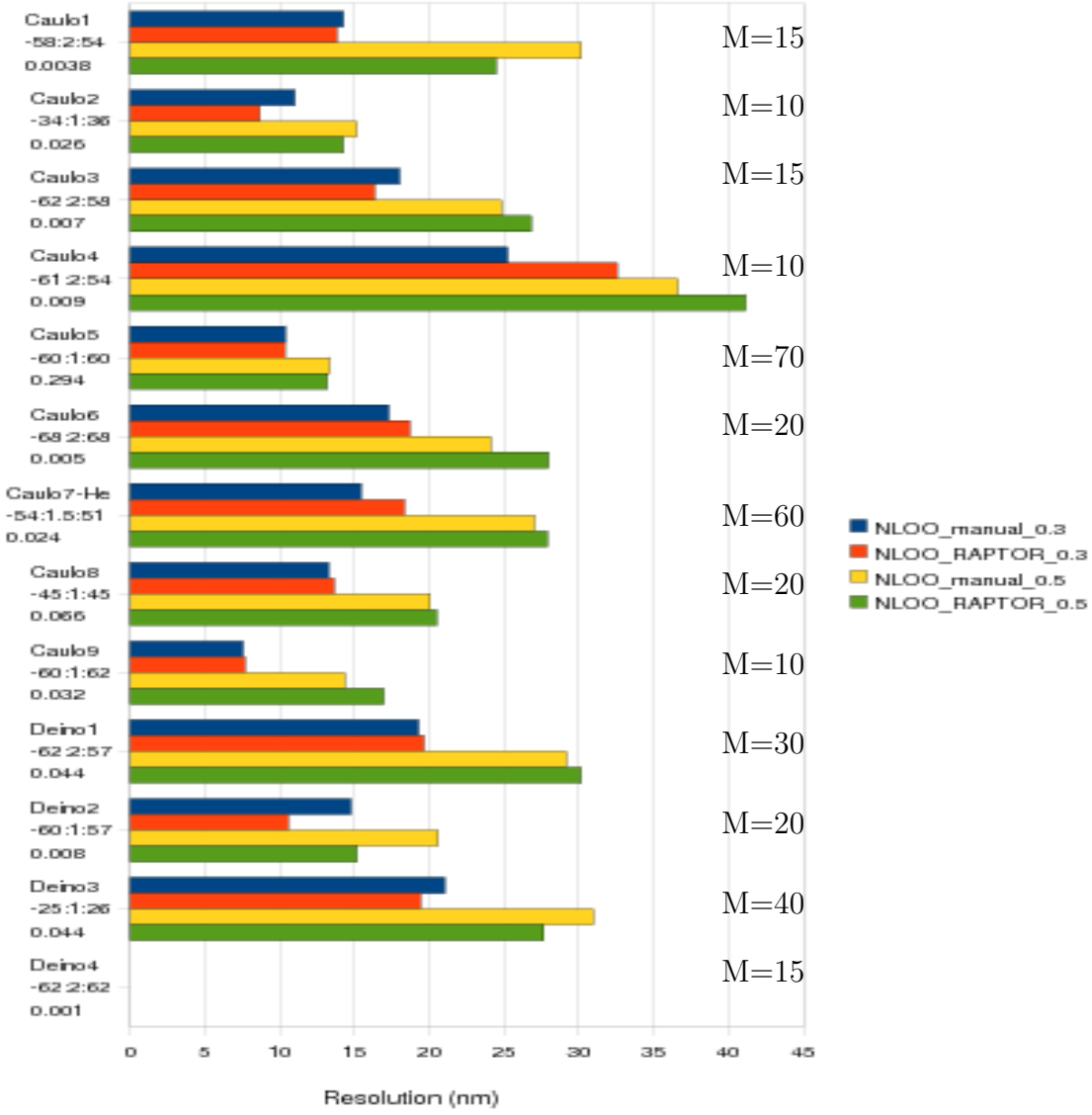


**Figure 3.14:** Summary of the datasets tested to assess the performance of RAPTOR. Bars represent resolution estimated by thresholds 0.3 and 0.5 in NLOO score comparing a human-directed alignment versus the automatic full-precision alignment presented in this thesis. The smaller the bar, the better alignment. At the left of each group of 4 bars we have the dataset name, the spacing between projections (in degrees) and the $SNR$ of the gold beads. At the right of each group of 4 bars we have the number of markers to track (M). Deino4 had such a low $SNR$ that was not aligned by any of the two methods, showing that there are still images where tracking is not possible.

model into the $MRF$ framework instead of waiting to have all pairwise correspondences to fit the model should help to deal with the uncertainties created in crowded samples.

Fig. 3.15 also shows how the estimated number of markers $M$ is the most critical parameter in order for RAPTOR to succeed. Right now it is considered an input from the user and it has to be close to the true number of markers in the sample, except in the case of samples with more than 100 markers. If the input $M$ is lower, RAPTOR will track short overlapping trajectories (instead of global ones across all the images). If the input $M$ is larger, RAPTOR will detect too many false positive markers per image that will throw off the correspondence. Future work will address this issues including the development of a module to automatically estimate the best $M$, so users do not have to look at the dataset before the microscope sends the tilt series directly to RAPTOR.

The second limitation is the fact that $NCC$ is not the best template matching score for all possible samples. $NCC$ is a good choice for $CET$ datasets because they have very low contrast features. However, X-ray or plastic embedded samples have much higher contrast and edges in stained surfaces receive high scores from the gold bead template shown in Fig. 3.2. Therefore, marker detection needs to be adapted to the type of sample that is being aligned.

Finally, even if in general $CET$ samples do not present non-linear distortions, they need to be taken into account in plastic sections and large field tomograms. Moreover, we need to account for different acquisition schemes such as dual tilt or conical sections to extend the applicability of RAPTOR in electron tomography. Methods like the one presented in [LBPE06] are extending the projection models in Eq. (3.4) up to third-order in order to improve the resolution reconstruction. However, there is a trade-off between the number of tracked markers and the number of model parameters to be fitted. The more non-linear distortions we want to capture, the more markers we need to track to avoid overfitting. Therefore, before extending the structure from motion approach explained in Section 3.4.2 we need to make sure that we can track over a hundred markers reliably.

We will see in Chapter 5 that RAPTOR has an approximate success rate of 70% in

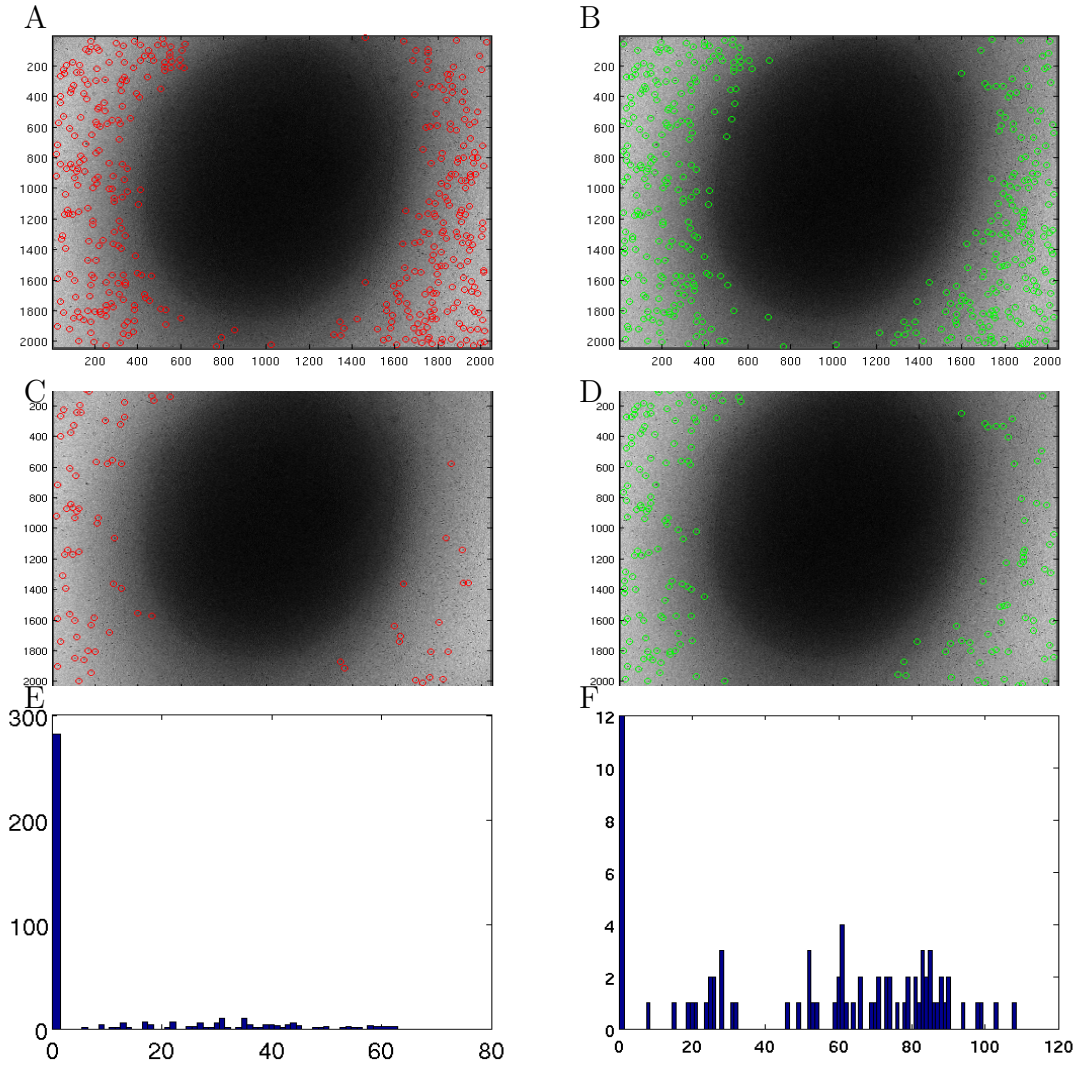


**Figure 3.15:** Results of pairwise correspondence in a sample with over 400 markers. (A) Markers detected (red circles) targeting $M = 400$ in the reference image. (B) Markers detected (green circles) in the candidate image. $K = 400$ in this case (C) Markers detected (red circles) targeting $M = 80$ in the reference image. (D) Markers detected (green circles) in the candidate image. $K = 180$ in this case. (E) Histogram showing the difference in the marginals in log-space (x-axis) returned by LBP between the maximum probability assignment and the second best choice for each marker between image (A) and (B). For 66% of the markers correspondence does not return a clear assignment. (F) Same as in (E) for pair of images in (C) and (D). For 15% of the markers correspondence does not return a clear assignment. *CET* images from Grant Jensen laboratory at Caltech.

$CET$ samples with gold beads. Even with the limitations and future work presented in this section, we have shown that we can reliably infere global geometry (projection maps) starting from local spatial correlations ($MRF$), which in combination with automation of other blocks of the $CET$ pipeline, has enabled a dramatic increment of the number of tomograms that can be acquired and reconstructed. The next chapter presents how we can take advantage of such a large number of tomograms in order to obtain higher resolution $3D$ images. In particular, we will extend to $3D$ the idea of using local spatial correlations to guide the alignment of pairs of images.

# Chapter 4

# Subtomogram alignment

Now that tomograms can be acquired, aligned and reconstructed in a high-throughput pipeline we need to extract and analyze structural patterns in the $3D$ volumes. In this chapter, we extend to $3D$ the probabilistic framework presented in the previous chapter to align $CET$ images and describe how we can obtain higher resolution structures by averaging thousands of aligned subvolumes. Fig. 4.1 shows the basic idea: by combining two similar images that have been blurred in different directions by the missing wedge, we can obtain a better reconstruction of the original image. This procedure is known in the field as subtomogram averaging and it has shown impressive results in many recent papers [LBB$^+$08, FMZ$^+$05, BRG$^+$09, NSP$^+$06]. Resolutions close to 20Å have been reported using subtomogram averaging, and continuing progress in the field makes higher resolutions expected. The end goal is to achieve near-atomic resolution of biological complexes close to their native state. Bartesaghi and Subramaniam [BS09] contains an excellent review on this topic.

Fig. 4.2 shows the typical work flow to perform subtomogram averaging. The first step is to approximately detect the position and orientation for each subtomogram. For example, imagine we are studying the HIV virus, then we need to find each virus present in each tomogram before we start the alignment and averaging procedure. By far, the most common method used in CET to perform this task is manual selection. The user

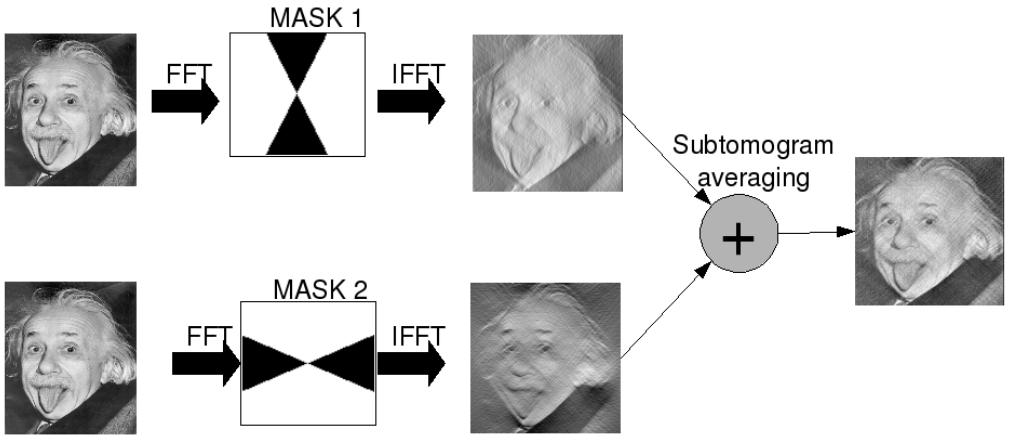


**Figure 4.1:** Basic idea behind subtomogram averaging: by combining two similar images that have been blurred in different directions by the missing wedge, we can obtain a better reconstruction of the original image.

will annotate each volume with locations and orientations for each object of interest. If the object under study contains preferred axis, it is very common to include that in the annotation, so we have all the structures prealigned before hand.

Since CET image have very low $SNR$ and the $SNR$ of an averaged structure increases as the square root of the total number of averaged objects, it is common to require thousands of examples in typical subtomogram averaging projects. Thus, automatic object detection would be desired for the initial selection step. Unfortunately, the low $SNR$ and the missing wedge effect that blurs features in different orientations, have prevented the adoption of standard object detection algorithms from computer vision. To date, the only attempts to use automatic object detection in $CET$ volumes were reported by Ortiz et al. [OFK$^+$06] performing exhaustive search in volumes by cross-correlating a template of the searched structure. As discussed in Chapter 6, automatic object detection in $CET$ volumes is still a very open field and its incorporation into the data analysis pipeline would be a great contribution.

Once we have selected the subtomograms in each volume, we have a list of positions and orientations for each particle.[1] Now we need to classify and align all the particles in order

[1]Sometimes we refer to a single subtomogram as a particle in analogy with the single particle terminology

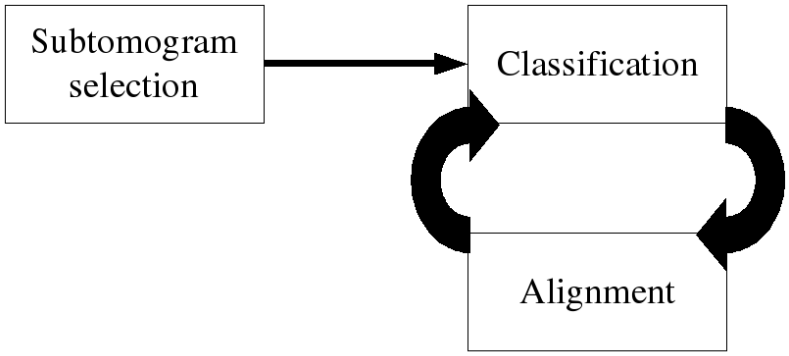


**Figure 4.2:** Block diagram showing the basic steps in the process of subtomogram averaging. First we need to select the subtomogram location and orientation in each $3D$ volume and then we iterate the steps of classification and alignment to find the underlying structures common to all the subtomograms.

to discover the underlying common structures. As in RAPTOR, we require high precision alignment, since the final resolution will only be as good as the precision in our alignment. The classification and alignment steps are generally iterated until convergence, and they have borrowed a lot of insights from the field of single particles described in Section 2.2. In fact, subtomogram averaging can be seen as an extension to $3D$ of the single particle concepts. However, the addition of the missing wedge effect on top of the low $SNR$ makes the problem particularly challenging.

Most existing approaches [BSL$^+$08, FPSF08, Win07, SPK$^+$06, WLR$^+$09, NSP$^+$06] solve the classification and alignment step by defining a metric, usually called dissimilarity score, between two subtomograms that measures how similar they are. In $CET$, the metric needs to be as insensitive to noise and missing wedge as possible. Once we have a metric that assigns scores between pairs of subtomograms, we can calculate the distance between all pairs to perform a classification step through standard clustering methods such as hierarchical clustering or spectral clustering. At the same time, using the centroids of each cluster, we can calculate a new alignment by minimizing the following equation:

described in Section 2.2.

$$\hat{R}^k, \hat{T}^k = \underset{R^k, T^k}{\operatorname{argmin}} \quad d(F^0, R^k \circ F^k + T^k) \qquad k = 1, \dots B \tag{4.1}$$

where $B$ is the total number of subtomograms and $d$ is the dissimilarity score of choice. $F^0$ represents the average obtained in the classification step and $R \circ F^k + T$ represents the rotation and translation of the subtomogram $F^k$. In all of the work published thus far, each individual particle is considered a rigid body and different *ad hoc* methods have been presented to deal with the problem of having multiple conformations of the same underlying structure to perform a multi-class alignment. A common strategy is also to perform a coarse-to-fine search for the parameters of $R$ and $T$.

In the recent years, subtomogram averaging has emerged as an important tool in the data analysis pipeline for $CET$, and we have seen above there are still many possibilities to be explored. For example, most of the work has focused on thin samples such as viruses, where missing wedge effects are more important than noise in registering subtomograms. In order to use the same averaging and classification tools in flexible macromolecules contained in thicker samples such as whole cells, we need to make the techniques more robust against noise. Higher tolerance to noise also allows one to acquire tomograms with lower defocus, effectively pushing the trade-off with respect to the first zero of the CTF to higher frequencies, which can also increase resolution.

In this chapter we focus on improving the alignment step to make it more robust against noise by defining a new dissimilarity score and incorporating local spatial correlations between different subtomograms. Section 4.2 below explains how we can use the $MRF$ framework explained in Chapter 3 to incorporate these two contributions at once. The main idea is that in many cases particles are part of a larger biological structure, so if we know the location and orientation of one subtomogram we can infer $T$ and $R$ for other subtomograms.

## 4.1 Notation

Before formalizing the ideas, we need to define some notation used in this chapter. $f^k = (f_1^k, f_2^k, \ldots, f_N^k)$ with $k = 1, \ldots, B$ represents the intensity values of the k-th subtomogram. We have $B$ subtomograms with $N$ voxels each. $F^k = (F_1^k, F_2^k, \ldots, F_N^k)$ represent the Fast Fourier Transformation (FFT) coefficients of each subtomogram. We want to define a metric or dissimilarity score between any two subtomograms based on a weighted Euclidean distance between Fourier coefficients for reasons explained later in this chapter:

$$d(F^{k_1}, F^{k_2})_K = \sum_{i=1}^{N} \|F_i^{k_1} - F_i^{k_2}\|_2^2 \cdot K_i \tag{4.2}$$

where $K = (K_1, K_2, \ldots, K_N)$ is a kernel function to be able to weight each coefficient differently. The only condition on $K$ is that $K_i \geq 0$ for $i = 1 \ldots N$.

For each $d(F^{k_1}, F^{k_2})_K$ we can associate an inner product between two subtomograms as follows:

$$\langle F^{k_1}, F^{k_2} \rangle_K = \sum_{i=1}^{N} F_i^{k_1} \cdot \bar{F}_i^{k_2} \cdot K_i \tag{4.3}$$

where $\bar{F}_i^{k_2}$ is the conjugate of the complex value $F_i^{k_2}$. The relation between Eq. (4.2) and (4.3) is:

$$d(F^{k_1}, F^{k_2})_K = \langle F^{k_1}, F^{k_1} \rangle_K + \langle F^{k_2}, F^{k_2} \rangle_K - 2 \cdot Re\{\langle F^{k_1}, F^{k_2} \rangle_K\} \tag{4.4}$$

where the operator $Re\{\}$ takes the real part of the inner product. Often it is convenient to normalize the subtomograms before comparing them in order to be resilient to illumination changes. Thus, we can normalize the subtomogram as:

$$\widetilde{F_K}_i^k = \frac{F_i}{\sqrt{\langle F^k, F^k \rangle_K}} \; ; \; i = 1 \ldots N \tag{4.5}$$

Then, Eq. (4.4) reduces to:

$$d(\widetilde{F_K}^{k_1}, \widetilde{F_K}^{k_2})_K = 2 \cdot (1 - Re\{\langle \widetilde{F_K}^{k_1}, \widetilde{F_K}^{k_1} \rangle_K\})$$

Therefore, minimizing $d(\widetilde{F_K}^{k_1}, \widetilde{F_K}^{k_2})_K$ is equivalent to maximizing $Re\{\langle \widetilde{F_K}^{k_1}, \widetilde{F_K}^{k_1} \rangle_K\}$.

## 4.2 Pairwise subtomogram correspondence with MRF

Corresponding two $3D$ volumes is in many ways analogous to corresponding groups of markers in $2D$: biological structures can be seen as large deformable structures composed by small rigid parts connected with each other (Fig. 4.3). This analogy has already been used in many applications such as the one for human pose estimation described in Chapter 2. In this framework, every rigid part is equivalent to a fiducial marker in $2D$ and we want to find its correspondence imposing relative spatial constrains between adjacent parts. As mentioned in the introduction, we assumed the initial step of subtomogram detection has been performed, so we have a list of positions and orientations for each rigid part.

The conceptual analogy allows us to use the same basic framework we used before: each rigid part in the first volume defines a random variable $A = \{A_1, \ldots, A_M\}$ and each possible location and orientation for the rigid parts in the second volume defines a set of possible assignments $B = \{B_1, \ldots, B_K\}$ for the correspondence problem. Defining each volume as $I_1$ and $I_2$, the solution to the correspondence problem is the assignment that returns the highest joint probability:

$$A^* = \underset{A_i}{\operatorname{argmax}}\, P(A_1, \ldots, A_M | I_1, I_2) \tag{4.6}$$

Using again the $MRF$ formalism, we can decompose the joint probability as:

$$P(A_1, \ldots, A_M | I_1, I_2) = \prod_{i=1}^{M} \psi_i(A_i) \prod_{(i,j)\in\varepsilon} \psi_{ij}(A_i, A_j) \tag{4.7}$$

One main difference with the methodology presented in Chapter 3 is the meaning of the random variables $A_i$. In RAPTOR, each random variable encodes a predetermined set of discrete hypothesis defined by the location of fiducial markers. In this chapter, each $A_i$ represents the possible rigid transformation ($R^k$ and $T^k$) applied to each subtomogram to align it with the reference image. In particular, the set of possible assignments $B_r$ is a discretization of the space of possible rotations and translations for each subtomogram.

All the effort in this chapter is devoted to describe the appropriate singletons ($\psi_i$) and pairwise ($\psi_{ij}$) potentials to define the meaning of high probability solution for the alignment problem in $3D$. While in $3D$ different rigid parts can look different, so it can be easier to prune corresponding candidates, the objects are not high-contrast point-like features any more, which makes the definition of singleton potentials more challenging. Moreover, even if we only consider rigid parts, in $3D$ we have 6 degrees of freedom (3 for translations and 3 for rotations), which makes the cardinality of the set $B$ much larger. Fortunately, we will see how most structures can be prealigned in certain directions using strong low resolution features, which will effectively reduce the search space.

### 4.2.1 Singleton potentials

This section presents a dissimilarity score robust to noise and missing wedge effects in $CET$ subtomograms, called Thresholded Constrained Cross-Correlation ($TCCC$) for reasons

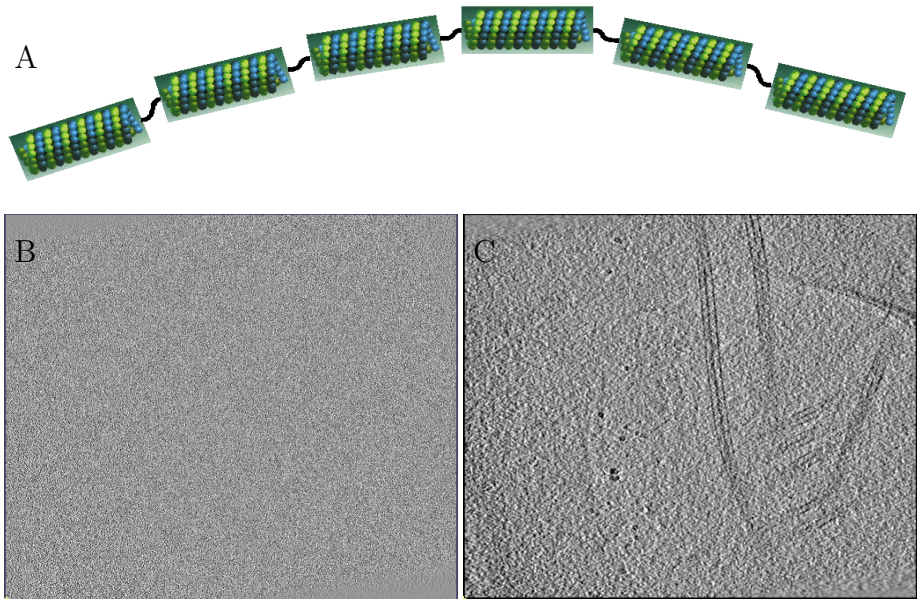


**Figure 4.3:** (A) Schematic showing how large biological structures can be thought as a composite of small rigid parts connected to each other by deformable edges. This particular schematic represents microtubules shown in Section 4.4. Microtubule schematic from [KF01]. (B) XY slice of a tomogram containing microtubules used in Section 4.4. (C) Same slice as in (B) using low pass filter to show the microtubules in the sample.

that will become obvious later. This dissimilarity score will replace the Normalized Cross-Correlation ($NCC$) used for $2D$ images. Our starting point will be the metric defined in Eq. (2) in [BSL$^+$08] since it provides "a formal framework for the generalization of dissimilarity measures previously proposed in the literature." Briefly, Eq. (2) in [BSL$^+$08] compares Fourier coefficients that are not masked by the missing wedge in either of the two volumes. They normalize this constrained cross-correlation by the number of coefficients in the comparison to avoid missing wedge bias. The mathematical expression is given in Eq. (4.8). For each possible translation and rotation $B_r$, we compare Fourier coefficients of the transformed volume instead of pixel intensities because the missing wedge is separable in Fourier space but not in image space.

**Thresholded Constrained Cross-Correlation**

The main idea behind $TCCC$ is that a subtomogram with $N$ voxels, with $N$ in the order of $10^6$ or above, can be represented by a small percentage of the required $N/2$ Fourier coefficients, especially in low $SNR$ regimes like in $CET$ (Fig. 4.4). The rest of the coefficients are either not necessary or are overwhelmed by noise. Below we show how to select a threshold that indicates which coefficients are useful to compare between subtomograms and how this modification makes the new metric more robust to noise.

The concept of thresholding in a transformed space where most of the energy is concentrated on a few coefficients is not new. For example, work by Donoho and Johnstone [DJJ93, Don95] over a decade ago sparked intensive research on natural images denoising via optimal thresholding of wavelet coefficients. There are many great reviews on the topic [Mal99, Fle08, FK03] and it has proved to be an excellent technique for denoising under additive white Gaussian noise models. Sorzano et al. [SOLR06] show an application of such denoising techniques to single particle $EM$ datasets. The metric presented here borrows some of those ideas and takes advantage of two facts: first, the number of voxels $N$ in subtomograms is much larger than the number of pixels in two-dimensional images, which is known to improve sparse representations. Second, instead of a single image, we have multiple copies of the same object, so we can obtain more reliable noise statistics that facilitates the threshold selection.

The main difference between cross-correlation-like measures is the kernel function, $K$, defined in the notation section. Using the above notation we can represent the metric in Eq. (2) in [BSL$^+$08] as:

$$d_B(F^k, F^l) = \frac{d(\widetilde{F_{K^B}}^k, \widetilde{F_{K^B}}^l)_{K^B}}{\sum_{l=1}^{N} K_i^B} \tag{4.8}$$

$$K_i^B = M_i^k \cdot M_i^l \cdot H_i \ ; \ i = 1 \ldots N$$

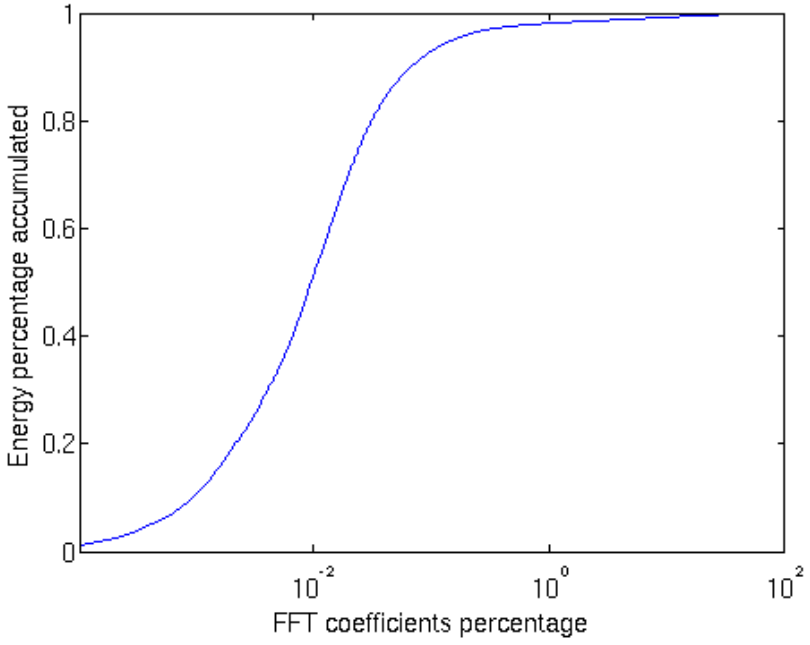


**Figure 4.4:** Cumulative magnitude distribution of sorted Fourier coefficients for EMBD model 1581 shown in Fig. 4.7. Most of the energy is concentrated in a small percentage of the total number of coefficients.

where $H_i$ is the coefficient of a band pass filter and $M^k$ and $M^l$ the binary mask representing the missing wedge. $M^k = 0$ if we do not have information in that particular coefficient or $M^k = 1$ if we have information. Given the orientation of the subtomogram and the tilt angles used for acquisition, it is possible to calculate $M^k$ for any subtomogram thanks to the Fourier Projection-Slice Theorem described in Chapter 2. We define the missing wedge overlap between two subtomograms as the Fourier coefficients where either $M^k$ or $M^l$ are equal to zero.

The authors in [BSL+08] also developed on top of this dissimilarity score an efficient method to align subtomograms based on Spherical Harmonics. In order to do that, they transform each subtomogram into a two-dimensional image by averaging the magnitude of Fourier coefficients along different rays in a sphere. This dimensionality reduction helps improve the SNR, so it can lead to good and fast coarse alignments even in low SNR

conditions. However, the simulations in Fig. 4.8 indicate that high precision alignment refinement using Eq. (4.8) might not be accurate. The main problem with this metric is that a large number of coefficients which are extremely small in $\tilde{F}^k$ are overwhelmed by noise in low $SNR$ settings. Those coefficients are not helpful when comparing two subtomograms. Thus, in those situations, the normalization factor using the amount of missing wedge overlap dominates the metric and biases the alignment toward orientations that maximize the size of the overlap between $M^k$ and $M^l$.

Inspired by the ideas on image denoising by thresholding wavelet coefficients we modify Eq. (4.8) by comparing only the strongest $C$ coefficients in $\tilde{F}^k$ and $\tilde{F}^l$. The new proposed metric can be expressed as:

$$\begin{aligned} d_C(F^k, F^l) &= \frac{d(\widetilde{F_{K^B}}^k, \widetilde{F_{K^B}}^l)_{K^C}}{\sum_{l=1}^{N} K_i^C} \\ &= \frac{d(\tilde{F}^k, \tilde{F}^l)_{K^C}}{C} \\ K_i^C &= \mathbb{1}_{(i \in \mathcal{C})} \end{aligned} \tag{4.9}$$

Where $\mathbb{1}$ is the indicator function[2] and $\mathcal{C}$ is the set of coefficients with $C$ largest FFT magnitude coefficients in the first or the second subtomogram outside the overlap between $M^k$ and $M^l$. In order to calculate $\mathcal{C}$ we compute the array $F_i^M = \max(\|\widetilde{F_{K^B}}^k{}_i\|, \|\widetilde{F_{K^B}}^l{}_i\|)$ for $i = 1 \dots N$. Then we sort this array based on descending magnitude order and the highest $C$ coefficients form the set $\mathcal{C}$. In other words, we select the coefficients that have high energy in subtomogram $F^k$, $F^l$ or both. Notice that we compute $\widetilde{F_{K^B}}^k$ as in Eq. (4.8) so we do not need $\mathcal{C}$ during the normalization step. In practice, we also apply a small high-pass filter ($H_i$ in Eq. (4.8)) to avoid selecting coefficients with very low frequency that can affect the alignment.

Eq. (4.9) has two main advantages: first, given the threshold $C$ has been chosen appropriately, it only considers coefficients that are not overwhelmed by noise to compare

[2]The indicator function $\mathbb{1}_{(i \in \mathcal{C})}$ is 1 if element $i$ belongs to the set $\mathcal{C}$ and 0 otherwise.

two subtomograms. That makes the metric more robust in typical low $SNR$ electron microscope subtomograms. Second, we can expect the value of $C$ to not change across a set of subtomograms obtained from similar datasets. That makes the normalization factor in Eq. (4.9) be the same for all possible pairs of comparisons. Thus, the new metric does not favor alignments with small overlap between missing wedges (like in [FPSF08] or with large overlap (like in [BSL$^+$08]).

The main assumption is that FFT coefficients where the energy is concentrated can not all be masked by a single missing wedge,[3] so even if two subtomograms have different missing wedges we can still align a portion of the high energy coefficients. Moreover, Fig. 4.5 shows how there is a range of $C$ values that return similar $TCCC$ scores. Therefore, even if the optimal choice of $C$ might change for two pairs of subtomograms with different relative orientation, assuming a constant $C$ for a set of subtomograms delivers close to optimal results as long as the main assumption is satisfied.

Most of the literature on denoising by thresholding uses various wavelet bases instead of Fourier transform coefficients because wavelets encode all the information in even fewer coefficients. We decided to use Fourier transform for two reasons: first, the interpretation and incorporation of the missing wedge in Fourier space is straightforward. It is easy to separate missing data from useful information in each subtomogram. Second, theoretical studies on sparse representations via Fourier or wavelet analysis [DJJ93] show that thresholding techniques perform better if $N$ is large (asymptotic case). Typically $N$ is on the order of $10^6$ or larger in our subtomograms, making $N$ much larger than typical two dimensional natural images reported in the literature. For such a large $N$, the choice of wavelets versus Fourier transform is not critical in order to obtain a sparse representation.

### Adaptive threshold selection using maximum likelihood

The number of coefficients $C$ is the main parameter of the metric presented above. We will take advantage of the fact that in subtomogram averaging we have many copies of

[3] This is an implicit assumption in all cross-correlation like approaches.

the same object we can use to develop a maximum likelihood ($ML$) approach to estimate $C$ from the data itself. Notice that if we set $C$ equal to the total number of coefficients available outside the overlap of missing wedges, then we recover Eq. (4.8).

Appendix B shows that we can approximate the noise present in $CET$ volumes as follows:

$$f^k = x + \varepsilon^k \ ; \ k = 1, \ldots, B \tag{4.10}$$

where $x$ is the true underlying signal and $\varepsilon^k$ is additive noise with Gaussian distribution $\varepsilon^k \sim N(\mathbf{0}, \Sigma)$ for $k = 1, \ldots, B$ and $\Sigma = diag(\sigma_1^2, \ldots, \sigma_N^2)$. In other words, the noise has zero mean and a diagonal covariance matrix $\Sigma$ calculated over a set of $B$ subtomograms with different noise power for each voxel.

Based on the noise model, we can determine the optimal value of $C$ from the data itself. In order to do that, first we need to estimate two parameters, $\mu$ nd $\Sigma$, assuming we do not threshold subtomograms. Since each subtomogram in a class contains the same underlying object $x$, we can estimate $\mu$ and $\Sigma$ with the usual mean and variance unbiased estimator for each voxel. Formally:

$$\mu_i = \frac{1}{B} \sum_{k=1}^{B} f_i^k \quad i = 1, \ldots, N \tag{4.11}$$

$$\sigma_i^2 = \frac{1}{B-1} \sum_{k=1}^{B} (f_i^k - \mu_i)^2 \quad i = 1, \ldots, N \tag{4.12}$$

We have made two main assumptions: First, the subtomograms are aligned so all the variance is due to noise. Second, the subtomograms have similar orientations so the blur caused by the missing wedge does not bias the statistics in real image space. The first assumption is acceptable if boxes are roughly aligned. We can do that by doing a first pass with a guessed value[4] for $C$ and estimate $C$ after the first iteration with this noise model.

[4]In general a value of $C$ representing 10 to 20% of the Fourier coefficients is a safe choice for an initial alignment.

While estimating $C$ we use only the half of subtomograms with lowest dissimilarity score to ensure that all the variance is due to noise. The second assumption can be handled because we know the missing wedge orientation of each subtomogram. Thus, we only consider the largest subgroup of subtomograms with similar orientation when estimating $C$.

Once we have estimated $\mu$ and $\Sigma$ we can estimate $C$ using an $ML$ approach. The underlying assumption is that the signal $x$ can be generated by a small percentage of the total number of Fourier coefficients (Fig. 4.4). Thus, $C$ can be understood as the number of non-zero Fourier coefficients that are needed to describe the underlying signal $x$. Given that our noise is additive Gaussian, we can formulate the log-likelihood for the parameter $C$ given subtomograms $f^1, \ldots, f^B$ as:

$$\begin{aligned}
\mathcal{L}(f^1, \ldots, f^B; C) &= \log \prod_{k=1}^{B} \mathcal{L}(f^k; C) \\
&= \log \prod_{k=1}^{B} A \; e^{-\frac{1}{2}(g(f^k;C)-\mu)^T \;\; \Sigma^{-1} \;\; (g(f^k;C)-\mu)} \\
&= -\frac{1}{2} \sum_{k=1}^{B} \sum_{i=1}^{N} \left( \frac{g(f^k;C)_i - \mu_i}{\sigma_i} \right)^2 + B \log A
\end{aligned} \tag{4.13}$$

where $A$ is a normalization constant, $\mathcal{L}$ is the likelihood function (Gaussian in our case), and $g(f^k; C)$ is the thresholding operation on the k-th subtomogram. The operation $g(f^k; C)$ computes the FFT of $f^k$, selects the $C$ coefficients with the largest magnitude, sets the rest to zero, and computes the inverse FFT. $\Sigma$ just weights each voxel according to its uncertainty, so we compute a weighted Euclidean distance between the average of all the subtomograms and each thresholded subtomogram. From Eq. (4.13) we see that the $ML$ estimator for $C$ is given by:

$$C^* = \underset{C}{\operatorname{argmax}} -\frac{1}{2} \sum_{k=1}^{B} \sum_{i=1}^{N} \left( \frac{g(f^k;C)_i - \mu_i}{\sigma_i} \right)^2 \tag{4.14}$$

Clearly, it is difficult to obtain an analytical expression for the above equation and tests in different datasets (Fig. 4.5) show the likelihood function is smooth but non-convex as a function of $C$. Thus, gradient descend methods might not find the global minimum. However, Eq. (4.13) can be evaluated very efficiently for multiple values of $C$ at the same time since we only need to compute an extra inverse FFT for each value of $C$. In practice, we evaluate Eq. (4.13) for multiple values of $C$ for all the boxes and choose the $C$ that returns a higher likelihood score. Fig. 4.5 shows curves for the function to be maximized in Eq. (4.14) and how the maximum changes in different $SNR$ situations.

We can imagine two extreme cases for $C^*$. If $C^* = 0$, we would be ignoring completely any underlying signal $x$ and we would have a large bias representing $x$. On the other extreme, $C^* = N/2$, we would be considering all the available Fourier coefficients, which will give us a large variance on the representation of $x$ due to overfitting the noise in Eq. (4.10). The purpose of the ML estimation procedure is to find the value for $C$ in between these two extremes given the observed data.

### 4.2.2 Pairwise potentials

In Section 3.3 we defined the pairwise potentials $\psi_{ij}$ as a Gaussian relaxation of the relative position between neighboring fiducial markers. Moreover, since the gold beads are spherical, we did not have to consider possible in-plane rotations of the features. This concept is straight forward to extend to $3D$ coordinates, but now we also have to account for possible changes in the relative orientation between neighboring rigid parts. In this section we will show how to represent the space of possible $3D$ orientations to incorporate that constraint in our pairwise potentials.

The space of possible orientations in $3D$, $SO(3)$, is formally defined as the set of 3x3 orthogonal matrices that have determinant equal to 1. There are many parametrizations of $SO(3)$, such as Euler angles, rotation matrices, axis-angle or quaternions, and each presents different advantages and disadvantages depending on the problem at hand. For

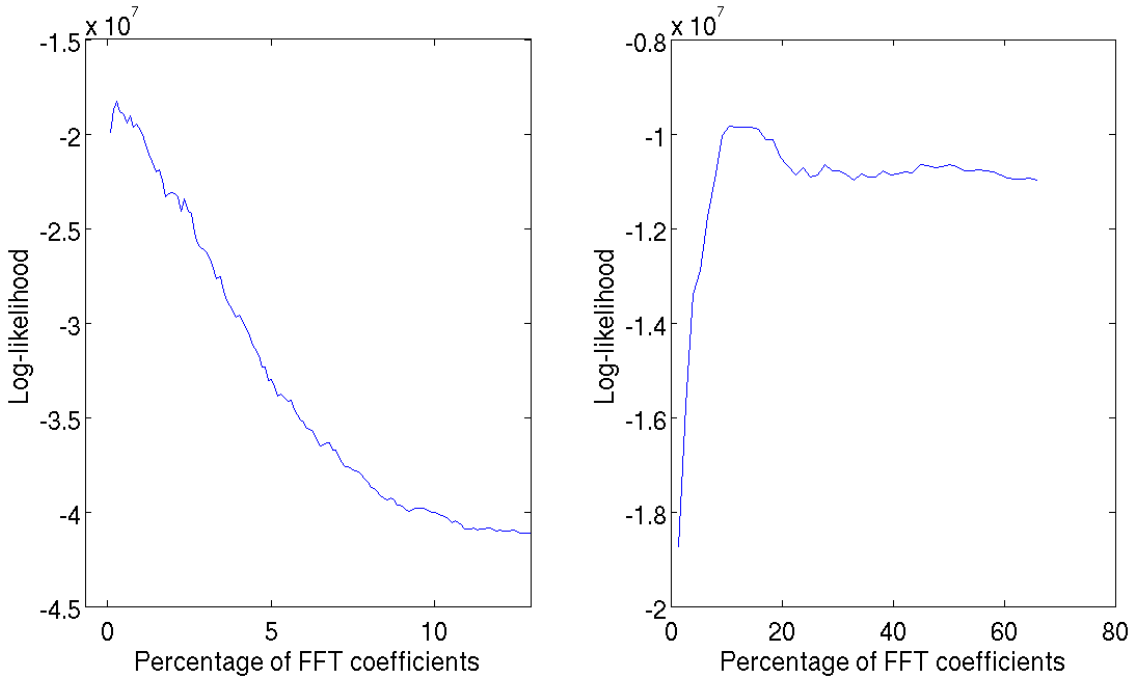


**Figure 4.5:** Likelihood score as a function of $C$ for different $SNR$. Left: likelihood estimated from 300 aligned subtomograms from the real data described in Appendix A where $SNR$ is very low. $C^*$ is very small. Right: likelihood estimated from 150 aligned subtomograms from the phantom described in Section 4.4 with $SNR = 10$. $C^*$ is analogous to selecting all the non-zero coefficients according to Fig. 4.4.

example, Euler angles are conceptually easier to understand but they present singularities. If we want to compare relative distances between orientations of different objects, continuity is an important property to have. Quaternions are a mathematical construct containing four elements that have proved very helpful to parametrize and operate in the set of 3D orientations avoiding singularities like the Gimbal lock [Hoa63] in Euler angles. In the next section, we present a brief introduction to the quaternion formalism in order to understand how we define the pairwise potentials.

### Quaternions

Quaternions can be approached from different points of view since they have found applications in different fields such as robotics [FP88], aeronautics [Kui98], computer graphics

[Sho85] or molecular modeling [Kar07]. Since we are dealing with rotations, we will approach them from their relationship with the axis-angle parametrization of $SO(3)$: any rotation in 3D can be defined by an axis $\overrightarrow{v} \in \mathbb{R}^3$ with unit norm and an angle $\theta$ defining the rotation around that axis. More formally, $SO(3)$ is locally equivalent to the product of $\mathbb{S}^2$x$\mathbb{S}^1$, where $\mathbb{S}^N = \{(x_1, \ldots, x_{N+1}) \mid x_1^2 + \cdots + x_{N+1}^2 = 1\}$ , so we parametrize each point in the sphere $\mathbb{S}^2$ by $\overrightarrow{v}$, but we still have one degree of freedom for an in-plane rotation defined by $\theta$ that represents $\mathbb{S}^1$. This parametrization of $SO(3)$ is captured using unit quaternions by defining a four element vector $q = [q_0\ q_1\ q_2\ q_3] \in \mathbb{S}^3$ as follows:

$$q = [\cos(\theta/2) \quad \overrightarrow{v} \cdot \sin(\theta/2)] \tag{4.15}$$

The advantage of this formulation is that quaternions have a well defined set of operations that we can use to operate in $SO(3)$. In particular, they can be seen as an extension of complex numbers, where $q = q_0 + q_1 i + q_2 j + q_3 k$ has one real part containing the information related to $\theta$ and three imaginary parts containing information about $\overrightarrow{v}$. The rules to add and multiply quaternions are the same as for complex numbers with $i^2 = j^2 = k^2 = ijk = -1$. Therefore, if we multiply two unit quaternions $q^A$ and $q^B$ we obtain another unit quaternion equivalent the composition of the two rotations represented by $q^A$ and $q^B$. This is analogous to the case of rotations in $2D$ represented by complex numbers with norm one and phase $\theta^A$ and $\theta^B$.

Looking at the definition in Eq. (4.15) it seems that the set of orientations in $3D$ is equivalent to the hypersphere $\mathbb{S}^3$. However, this statement is not entirely precise. Using again the axis-angle representation, we realize that the rotation around $\overrightarrow{v}$ and angle $\theta$ is equivalent to the rotation around $-\overrightarrow{v}$ and angle $2\pi - \theta$. Substituting this in Eq. (4.15), we obtain that $q$ and $-q$ represent the same orientation. Therefore, we need to identify every antipodal point $q$ and $-q$ in $\mathbb{S}^3$ to enforce uniqueness. An intuitive way to see this is to fix $\overrightarrow{v}$ and start moving $\theta$: at the beginning we move away from the unit rotation but after $\theta > \pi$ we actually start getting closer again until we reach the opposite point in

the hypersphere. The formal definition of this equivalence between unit quaternions and $SO(3)$ is formulated as the following quotient space [Kui98]:

$$SO(3) = \{\mathbb{S}^3/\{q, -q\} \;\; | \;\; q = [q_0 \; q_1 \; q_2 \; q_3] \in \mathbb{R}^4; \|q\|_2 = 1\} \tag{4.16}$$

where the operator / denotes the quotient space given by the identification of $q - q$ in $\mathbb{S}^3$. This representation of $SO(3)$ defines a natural metric between any two orientations $a, b \in SO(3)$ as the length of the shortest arc between $q^A$ and $q^B$ on the hypersphere: $dist(q^A, q^B) = \arccos(|q^A \cdot q^B|)$. Remember that having a metric to compare distances in $SO(3)$ is one the properties we need to define our pairwise potentials.

However, $dist(q^A, q^B)$ is not Euclidean since it is defined along the hypersphere, and we can not use it with a standard Gaussian relaxation. In order to do that, we use again the axis-angle representation of quaternions and the Lambertian equal-volume projection into a unit $3D$ ball to have an Euclidean metric [Kar07]. Imagine the north hemisphere of the Earth: it is easy to visualize how to flatten it into a two-dimensional disc on the plane preserving the area (Fig. 4.6). This is exactly the Lambertian equal-area projection used in world maps to represent the earth. We can use the same concept to "flatten" $SO(3)$ into a unit ball in $3D$ by scaling axis $\overrightarrow{v} \in \mathbb{R}^3$ with a magnitude proportional to the in-plane rotation $\theta$. The proportionality function is defined in order to preserve volumes in the projected space. More formally:

$$\tilde{q} = \lambda \cdot \overrightarrow{v} \;\; \in \mathbb{R}^3 \tag{4.17}$$

$$\lambda = \left( \frac{|\theta| - \sin|\theta|}{\pi} \right)^{1/3} \;\; \in \mathbb{R} \tag{4.18}$$

It is clear in Fig. 4.6 that we preserve areas by distortioning the shape (or angles) of the elements on the sphere. However, points around the center suffer minimal distortion.

Since the pairwise potentials measure relative changes in orientation between neighboring elements, our observations lie close to the center of the unit ball, which allow us to use Gaussian statistics in $\tilde{q}$ to measure relative changes in orientations for the pairwise potentials.

**Figure 4.6:** Example of equal-area Lambertian projection to "flatten" the North hemisphere into a disk. In order to preserve area the projection map has to distort angles, which affects the shapes close to the equator.

Before we explain in detail the computation of pairwise potentials, we should point to the reader why three parameters are never enough to obtain continuous parametrizations of the space of orientations $SO(3)$. Again, we will use the analogy with $\mathbb{S}^2$ to gain intuition. Informally, the popularly called "hairy ball" theorem from algebraic topology [EG79], states that it is impossible to brush a hairy tennis ball in a way such that the vector field created by the hairs is continuous everywhere. There will always be a discontinuity or a singularity. A corollary of this statement is that there is no global parametrization of the whole sphere with only two parameters without discontinuities or singularities. We need to embed the sphere in $3D$ with the implicit definition $x^2 + y^2 + z^2 = 1$ in order to avoid degenerations even if the sphere only has two degrees of freedom. The situation in $\mathbb{S}^3$ is analogous which is why we need quaternions if we want a parametrization without discontinuities.

### Computation of pairwise potentials

Now that we have the notion of how quaternions can be used to parametrize orientations, it is relatively straight forward to define the pairwise potentials for alignment of $3D$ volumes. The mathematical expression for $\psi_{ij}$ is the following:

$$\psi_{ij}(A_i = B_r, A_j = B_s) = f_{\sigma_1}(v_{ij}, v_{rs}) \cdot f_{\sigma_2}(\tilde{q}_{ij}, \tilde{q}_{rs}) \tag{4.19}$$
$$f_{\sigma}(\overrightarrow{x}, \overrightarrow{y}) = \exp\{-\frac{1}{2}\frac{\| \overrightarrow{x} - \overrightarrow{y} \|^2}{\sigma^2}\}$$

where, as in Section 3.3, $v_{ij}$ and $v_{rs}$ represent the vector connecting points $A_i$ to $A_j$ and $B_r$ to $B_s$ respectively. $\tilde{q}_{ij}$ and $\tilde{q}_{rs}$ represents the change in orientation between $A_i$ to $A_j$ and $B_r$ to $B_s$ respectively, in a space with a metric locally equivalent to Euclidean, so we can still use a Gaussian relaxation to impose the constrains.

Defining $q_i$, $q_j$, $q_r$, $q_s$ as the quaternions representing the orientations of $A_i$, $A_j$, $B_r$ and $B_s$ respectively, we can define the terms in Eq. (4.19) as $q_{ij} = q_i/q_j$ and $q_{rs} = q_r/q_s$, where / represents the division of two quaternions. In other words, $q_{ij}$ is the rotation necessary to move the structure in $A_i$ to the structure $A_j$. Once we have $q_{ij}$ and $q_{rs}$ we can project them to $\mathbb{R}^3$ using Eq. (4.17) and evaluate $f_{\sigma_2}(\tilde{q}_{ij}, \tilde{q}_{rs})$.

The only thing left is how to choose the value of parameters $\sigma_1$ and $\sigma_2$. Those parameters represent how much we allow the rigid components to deform relative to each other in terms of position ($\sigma_1$) and orientation ($\sigma_2$). Unfortunately, in the $3D$ case we do not have manually aligned data as in the $2D$ case to estimate those parameters. Usually we label one dataset manually to obtain a first estimate and use as a reference. Another option is to use different sets of values for the parameters $\sigma_1$ and $\sigma_2$ and compare which ones return better results in terms of final resolution using Fourier shell correlation ($FSC$) curves.

## 4.3 Alignment refinement

Once we have an initial location for each particle we would like to refine its position and orientation in a final alignment step in order to increase the final resolution of the average structure. The key point here is to have a good initial estimate of $R^k$ and $T^k$ so we can apply an iterative local refinement method without being trapped in a local minimum. We use a standard algorithm that can be described in terms of expectation-maximization [DLR77] or iterative closest points (ICP) [BM92] for $3D$ volumes. In the field of subtomogram alignment, this approach was first described by Forster et al. [FMZ$^+$05].

The alignment refinement is performed iteratively in two steps: first, we calculate the average structure given by the positions and orientations returned by the inference in the $MRF$. Second, given the average structure, we perform a local search to refine the location and orientation of each individual subtomogram independently to minimize the $TCCC$ dissimilarity score between the average structure and each subtomogram. The process is repeated until convergence.

In our implementation, the maximization step finds the rotation $\hat{R}^k$ and the translation $\hat{T}^k$ such that:

$$\hat{R}^k, \hat{T}^k = \underset{R^k, T^k}{\operatorname{argmin}} \quad d_C(F^0, R^k \circ F^k + T^k) \qquad k = 1, \dots B \tag{4.20}$$

where $F^0$ represents the template obtained in the expectation step and $R \circ F^k + T$ represents the rotation and translation of the subtomogram $F^k$. The code applies a gradient descent search of translations and rotations over a range of values specified by the user with a certain step size. For each calculation of $d_C$ we normalize the transformed subtomogram as described in Eq. (4.5) to recalculate the $C$ highest coefficients. Once we have aligned each subtomogram, we calculate a new template by averaging all of the aligned subtomograms in Fourier space in order to appropriately weigh each coefficient by the number of contributing subtomograms. Averaging in Fourier space is critical to avoid mixing regions with missing

information with regions of useful information from different subtomograms.

Also, after each iteration we can calculate a distribution of number of contributing subtomograms for each Fourier coefficient in order to check if the template contains also a missing wedge. Ideally, different subtomograms will have different orientations to contribute information to different parts of the Fourier domain and recover a structure without missing information (Fig. 4.1). However, this is not always the case. Since $d_C$ incorporates the missing wedge effect for both $F^0$ and $F^k$, it is straight forward to handle a missing wedge in the template itself. Also, before each alignment cycle we can re-estimate $C$ if necessary as explained in Section 4.2.1.

## 4.4 Results

The alignment refinement methodology using $TCCC$ has been implemented in $C++$ and made publicly available. The code can run in single thread machines or large clusters in order to take advantage of the simple parallel structure of the algorithm. Most of the results presented in this chapter would not have been possible without leveraging on parallel computing, since we aligned thousands of subtomograms to generate them.

### 4.4.1 Alignment refinement accuracy in synthetic data

First, we tested the $TCCC$ dissimilarity score against other existing methods to verify that is more robust against noise and missing wedge effects. We performed exactly the same tests (same data, same search procedure, etc.) but testing three different metrics: Constrained Cross-Correlation as presented in [FPSF08], Eq. (2) from [BSL$^+$08] and the new metric $TCCC$ presented here in Eq. (4.9). All the metrics are implemented using the FFT libraries in FFTW [FJ05] and the same missing wedge mask, windowing and band pass for each subtomogram. Thus, the only difference in the code is the metric itself.

First, we generate synthetic data as realistic as possible using the following steps:

1. Download a model publicly available from Electron Microscopy Database (EMDB)

[TNC+02]. In particular we download entry 1581 showing dynein's microtubule-binding domain (Fig. 4.7A). We bin the model by two so each particle fits in a box of $96^3$ voxels considering the necessary windowing to avoid edge effects in the FFT.

2. Generate six tomograms of $512 \times 512 \times 512$ in size with 25 particles in each of them at random orientations and locations. We record the location and orientation of each particle to have ground truth. In total we have 150 particles.

3. For a given tilt range $m$, we generate a file with tilt angles from $-m$ to $+m$ with 1 degree increments. For a given $SNR$ we scale the tomogram so when introducing Poisson noise in each projection we will have the desired $SNR$. Finally, we use the *xyzproj* program from IMOD [KMM96] to obtain projections at different angles. Our procedure to define and estimate $SNR$ follows the formula in [Fra96] to reflect the visibility of different biological structures in the image:

$$SNR = \frac{\sum_{i=1}^{N} \hat{f}_i^2}{\sum_{i=1}^{N} \frac{1}{B-1} \sum_{k=1}^{B} (f_i^k - \hat{f}_i)^2}$$
$$\hat{f}_i = \frac{1}{B} \sum_{k=1}^{B} f_i^k \ ; \ i = 1 \ldots N$$

4. Introduce Poisson noise in each projection using the function *poissrnd* from MATLAB$^{TM}$ software. If the i-th pixel has intensity value $p_i$, then the new value is generated following a Poisson distribution with mean $p_i$.

5. Generate a random shift for each projection to simulate missalignment. We draw the shift from a uniform probability distribution in x,y of $[-3, 3] \times [-3, 3]$ pixels.

6. Reconstruct the projections using the weighted back-projection (WBP) algorithm in IMOD [Mas97]. We choose RADIAL filter parameters 0.5 and 0.0 to avoid any low pass filtering that limits resolution.

To test alignment accuracy, we add random alterations to the ground truth in order to give misplaced initial points and then we run the alignment refinement step explained in Section 4.3. The alterations are drawn from a uniform distribution in (x,y,z) of $[-10, 10] \times [-10, 10] \times [-10, 10]$ voxels and a uniform distribution in ($\alpha$,$\beta$,$\gamma$) (the three Euler angles) of $[-15, 15] \times [-15, 15] \times [-15, 15]$ degrees. We generate tomograms following the steps above for any combination of $SNR = \{10, 1, 0.1, 0.01, 0.001\}$ and tilt range= $\{50, 60, 70, 80, 90\}$ degrees. Fig. 4.7B and C show examples of the same particle with different $SNR$ and tilt range. The initial template is obtained with the perturbed initial points from the tomogram with $SNR = 0.001$ and tilt range = 50. We use the lowest quality template as an initial template for the alignment refinement to demonstrate that we do not need a clear template for the algorithm to converge to a good alignment.

Fig. 4.8 presents the root mean square error (RMSE) for each metric. Here we follow the error metric described in [BSL$^+$08]. We define 4x4 matrix with the rotation and transformation needed to transform each final aligned subtomogram with its original ground truth location and orientation. We calculate the RMSE of all the terms in each of those matrices. For $SNR < 0.1$ there is a clear gap between previous metrics and the new $TCCC$ presented in this thesis. This is expected since at very low $SNR$ the useful information in most of the coefficients is occluded by noise. Thresholding makes the metric more robust against noise, which is a main issue in $ET$. In higher $SNR$ environments, most coefficients carry useful information and so all the metrics perform similarly.

Table 4.1 shows how the maximum likelihood estimation for $C$ is able to adjust that parameter correctly. If we had used a small value of $C$ in high $SNR$ images, the alignment precision would have been worse than the other two dissimilarity scores. For example, we try alignment for $SNR = 10$ and no missing wedge using only 1.4% of the coefficients instead of 14%, and the alignment RMSE jumped from 0.1 to 0.5. Notice that for high $SNR$ regimes, the percentage of selected coefficients in Table 4.1 agrees with the point in Fig. 4.4 where the cumulative sum reaches almost one. In other words, we are comparing all the non-zero Fourier coefficients. This result does not hold for $SNR = 10$ and tilt range

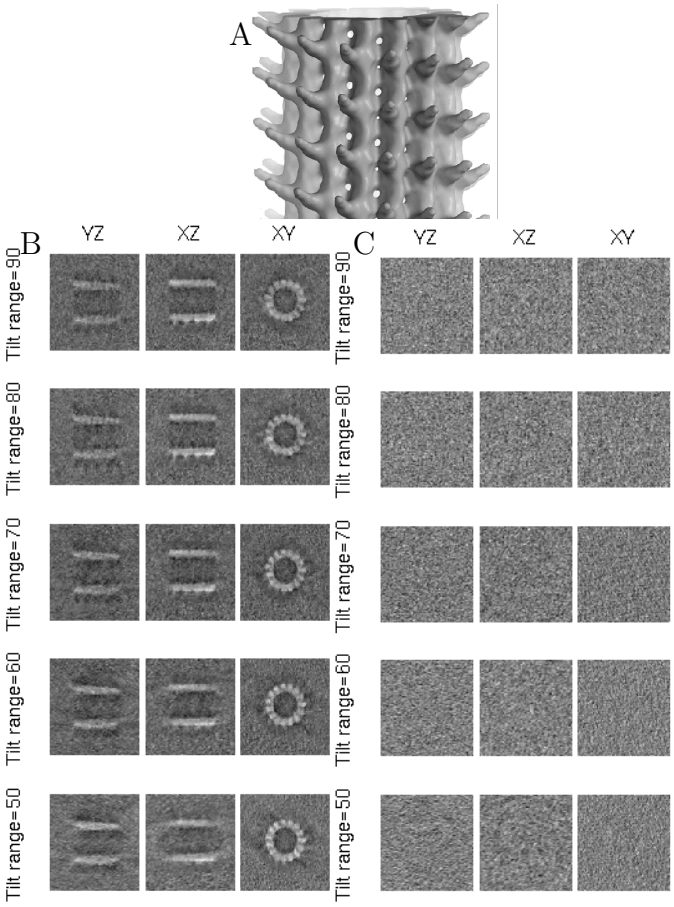


**Figure 4.7:** (A) Visualization of dynein's microtubule-binding domain from EMDB entry 1581. (B) Phantom generated from (A) at $SNR$ equal to 1 for different tilt ranges. (C) Phantom generated from (A) at $SNR$ equal to 0.001 for different tilt ranges.

= 50 because the large missing wedge is affecting the estimation of noise statistics. With only 150 particles, the largest subgroup with similar orientations was not large enough to get a good estimate of $C$ and we had to include all the particles, which affects the estimation because blurring caused by the missing wedge is confused with noise across the samples.

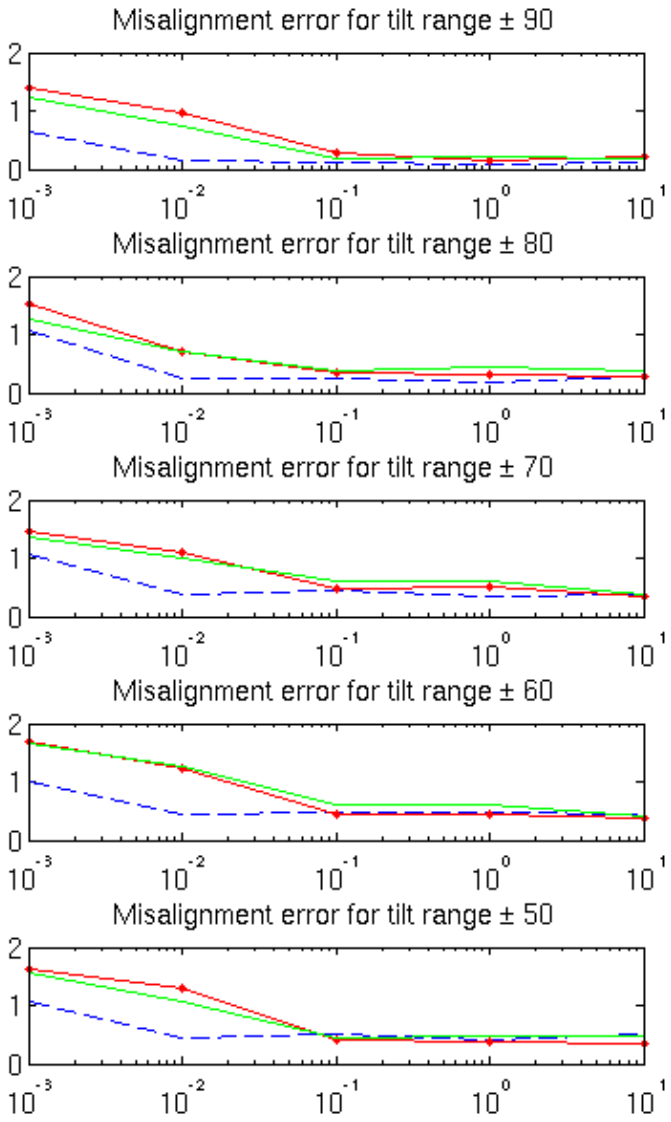


**Figure 4.8:** RMSE alignment comparison between three different metrics for synthetic data: *TCCC* (dashed blue), constrained cross-correlation from [FPSF08] (dot-continuous red) and Eq. (2) from [BSL$^+$08] (continuous green). Each graph shows results for a different tilt range and in each graph we plot different SNR ($x$-axis in log-scale) versus misalignment error ($y$-axis). Lower RMSE indicates better alignment.

### 4.4.2 Alignment accuracy in real data

In this section we present results in real *CET* images of microtubules. Fig. 4.3 shows a *CET* image of a microtubule and how it can be decomposed in rigid pieces. Basically, the

| $\frac{TR}{SNR}$ | ±90 | ±80 | ±70 | ±60 | ±50 |
|---|---|---|---|---|---|
| 10 | 14.24 | 14.73 | 11.76 | 10.78 | 4.84 |
| 1 | 4.34 | 3.85 | 3.85 | 3.35 | 1.95 |
| 0.1 | 1.87 | 2.36 | 1.87 | 1.37 | 0.88 |
| 0.01 | 0.39 | 0.88 | 0.39 | 0.39 | 0.26 |
| 0.001 | 0.39 | 0.39 | 0.32 | 0.32 | 0.32 |

**Table 4.1:** Percentage of Fourier coefficients selected in phantom for different $SNR$ and tilt range (TR) configurations.

repetitive rigid parts form a Markov chain where we can apply the probabilistic framework described above.

### Resolution criteria

Since we do not have ground truth for the structures we are resolving we need a resolution criteria to compare different alignments. We rely on Fourier Shell Correlation ($FSC$) curves [vHS05], which is an extension to $3D$ of the Fourier Ring Correlation (FRC) defined in Eq. (3.11). In this case, instead of using rings of Fourier coefficients, we use shells because we are in 3D. Mathematically, FSC between two subtomograms can be formulate as follows:

$$FRC_{F^k,F^l}(k) = \frac{\sum_{m,n,s\in R(k)} Re\{F^k_{m,n,s}\bar{F}^l_{m,n,s}\}}{\left(\left(\sum_{m,n,s\in R(k)} \|F^k_{m,n,s}\|^2\right)\left(\sum_{m,n,s\in R(k)} \|F^l_{m,n,s}\|^2\right)\right)^2} \tag{4.21}$$

$m, n, s \in R(k)$ are all the Fourier coefficients with radius $k$. In words, we are calculating a normalized cross correlation for each spatial resolution $k$. Therefore, in graphs such as the one in Fig. 4.9, the best possible curve would be one with a constant value of 1 for each spatial frequency. The main challenge when using FSC to compare subtomogram alignment accuracy is to satisfy the main hypothesis underlying FSC comparisons that two structures are completely independent. Arguably, the "perfect" FSC would be calculated between two structures taken from different samples, imaged on different microscopes and

processed with different software by different people. In this scenario, any correlation shown between samples will represent true similarities in the structures. In order to be as close as possible to the "perfect" FSC, we process two sets of subtomograms, which are selected from different datasets, independently from the beginning. Moreover, the initial template used for the alignment will also be different. All the FSC curves shown in this paper are calculated using *bresolve* from Bsoft [HCWS08].

Fig. 4.9 shows a simulation of how the effects discussed in Section 2.3.4 about a spatially variant CTF in tilted projections affects the resolution of subtomogram averaging. The visual figure is important to have as a reference so we can determine when our resolution is limited by the CTF or by other effects such as misalignment errors, noise, heterogeneity or electron beam incoherence. Unless raw $2D$ projections are corrected for CTF effects[5] the backprojection will be adding data from images with different CTFs. Thus, even if the alignment between subtomograms is perfect, we will still have the effect shown in Fig. 4.9: resolution is affected in the spatial frequency bands around zeros of the CTF.

### Microtubule doublets

Microtubules serve as structural components within cells and have a crucial role in the life of all eukaryotic cells, since they are involved in key processes such as organelle movement, separation of chromosomes during cell division and maintenance of cell shape. They form an elongated macromolecular machinery as the one shown in Fig. 4.3 which is composed mainly of tubulin protein. In this section, we show examples of microtubule doublet structures, which are components of axonemes that contain a number of proteins besides tubulin, which we would like to identify. Axonemes are usually found in arrays of nine doublets arranged around two singlet microtubules. Adjacent doublets coordinate sliding moves in order to produces periodic beating movements of the axoneme. We will show how using the methodology explained above we can obtain resolutions close to 35Å, which is sufficient to accurately find the location of several proteins within the microtubule doublet

[5]CTF correction in *CET* samples was not done for samples in this thesis.

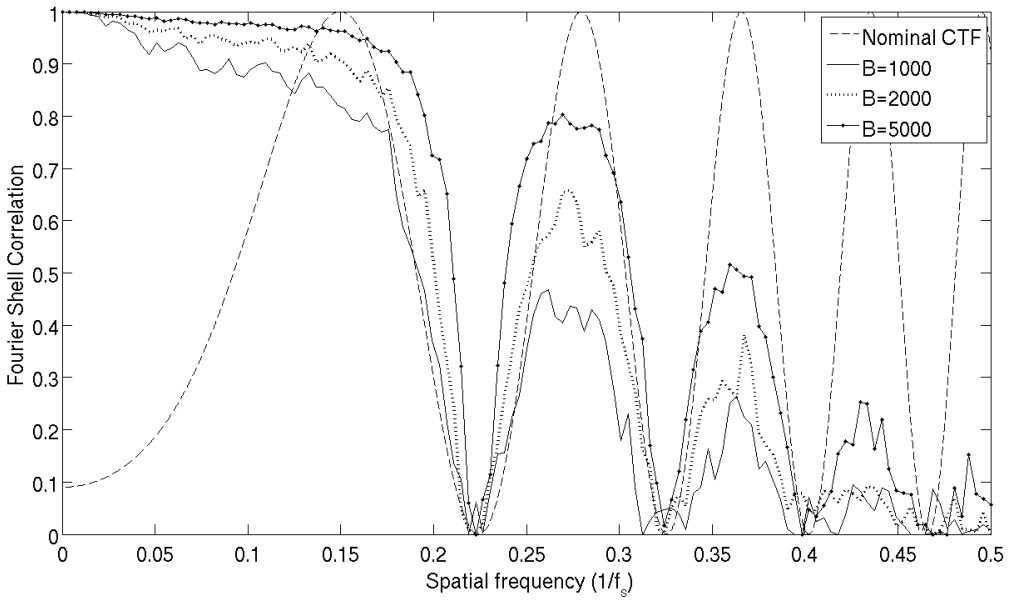


**Figure 4.9:** Simulations of how CTF variations between tomographic projections affects subtomogram averaging resolution. We averaged $B$ particles simulating a CTF effect with nominal defocus -10$\mu m$ plus random uniform deviations in $\pm 1.5\mu m$ range. Each particle was corrupted by Poisson noise to have $SNR = 0.01$. Dashed line is theoretical CTF at -10$\mu m$, continuous line has $B = 1000$, dotted line has $B = 2000$ and dot-dash line has $B = 5000$. Because we use synthetic tomogram, the decrease in resolution is due to noise and CTF variability, but not to misalignment.

structure and if known, dock their crystal structure to obtain a near-atomic resolution map [SD06].

Our final goal is to be able to compare microtubule structures from many different kinds of cells. In order to do that, we need to set up a methodology to efficiently process dozens of datasets. In this section, we present results for microtubule doublets of Chlamydomona flagellum to test the following pipeline:

1. Click ten points along each microtubule mid-axis to coarsely locate the center of each box using cubic splines. Because microtubules do not have high curvature ten points were enough to sample the mid-axis. In total 14 microtubules were extracted using this manual selection.

2. The data was split in two groups (A and B) of 7 microtubules each to perform

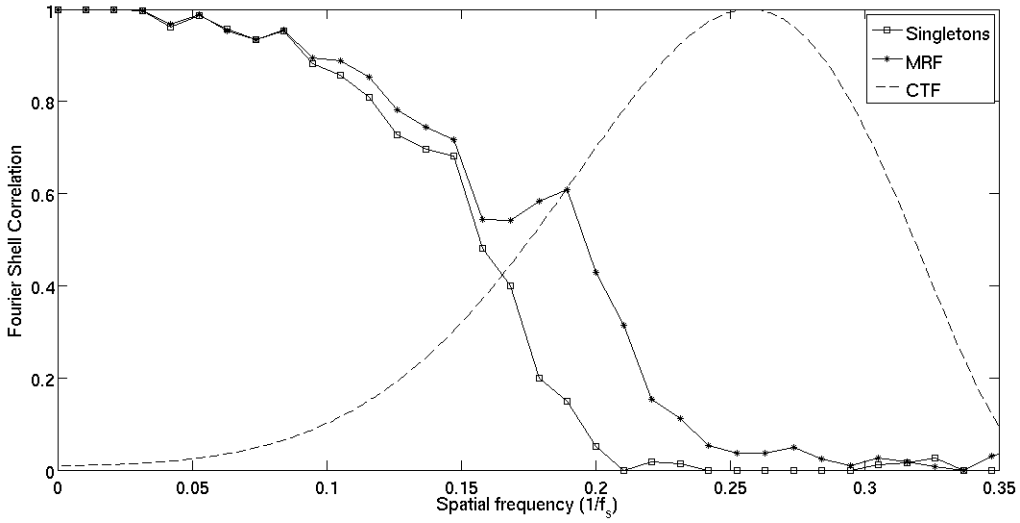


**Figure 4.10:** FSC curves comparing resolution obtained using only singleton potentials (squared line) and incorporating local spatial correlations with pairwise potentials (starred line). The spatial constraints improve the resolution of the average, although in this case we are not limited by the CTF (dashed line).

independent reconstructions to validate the resolved structures as shown in Fig. 4.10.

3. Each microtubule was split in sets of thirty consecutive boxes forming a Markov chain like the one shown in Fig. 4.3A. The separation between boxes within each set was 160Å, since it is one of the expected periodicities in the microtubule observed from previous results [SD06].

4. For each group A and B we choose a microtubule that would be used as a reference for the alignment using the probabilistic framework. The selection was made after visual inspection to select microtubules with the lowest curvature, so it is easy to perform the alignment automatically.

5. Singleton potentials were calculated using $TCCC$ with square boxes of size $70^3$nm and cylindrical mask around the microtubule. A band-pass filter between 3nm and and 20nm was also applied to alleviate noise effects. The continuous space of possible rotations and translations were discretized with a step size of $5^{circ}$ and 2.5 pixels

respectively. The search was only performed for 3 translations and a rotation around the mid-axis, since the spline curve removes two degrees of freedom. Therefore, due to the geometry of microtubules only 4 degrees of freedom were needed instead of 6, which reduces the cardinality of $B$.

6. Inference using max-product message passing as described in Elidan et al. [EMK06] was run to obtain the MAP assignment corresponding to Eq. (4.19)). Since microtubules form a Markov chain, inference is exact in this case.

7. Using the locations and positions returned for each box we run the alignment refinement procedure until convergence without any filtering to achieve the final averaging.

One advantage of the alignment this alignment work flow is that the initial reference can be obtained directly from the data itself instead of from an external template. This consideration is important since it is know in the subtomogram averaging literature that the initial template used in the local alignment refinement can bias the final result.

The final average for group A using only singleton potentials in Eq. (4.19) and incorporating the spatial correlation constraints with pairwise potentials can be seen in Fig. 4.11. The average obtained including spatial constraints is able to recover more structural details, since the MAP assignment delivers a better starting point for the refinement step. This visual perception is quantitatively confirmed when we compare FSC curves (Fig. 4.10) between group A and group B.

The reconstructions shown in Fig. 4.3B and C were obtained at -2$\mu m$ defocus, which provides very low contrast images. The estimated $SNR$ after alignment is 0.01, which belongs to the region where the use of $TCCC$ improves alignment according to Fig. 4.8. However, Fig. 4.12 shows the main reason why a dissimilarity score alone is not enough and we need pairwise potentials to improve the results in the case of microtubules. The value of the $TCCC$ score does not change abruptly along the direction of the mid-axis of the microtubule or rotating the reference around the mid-axis. In other words, the minimum in the $TCCC$ score is not clear along these two degrees of freedom due to

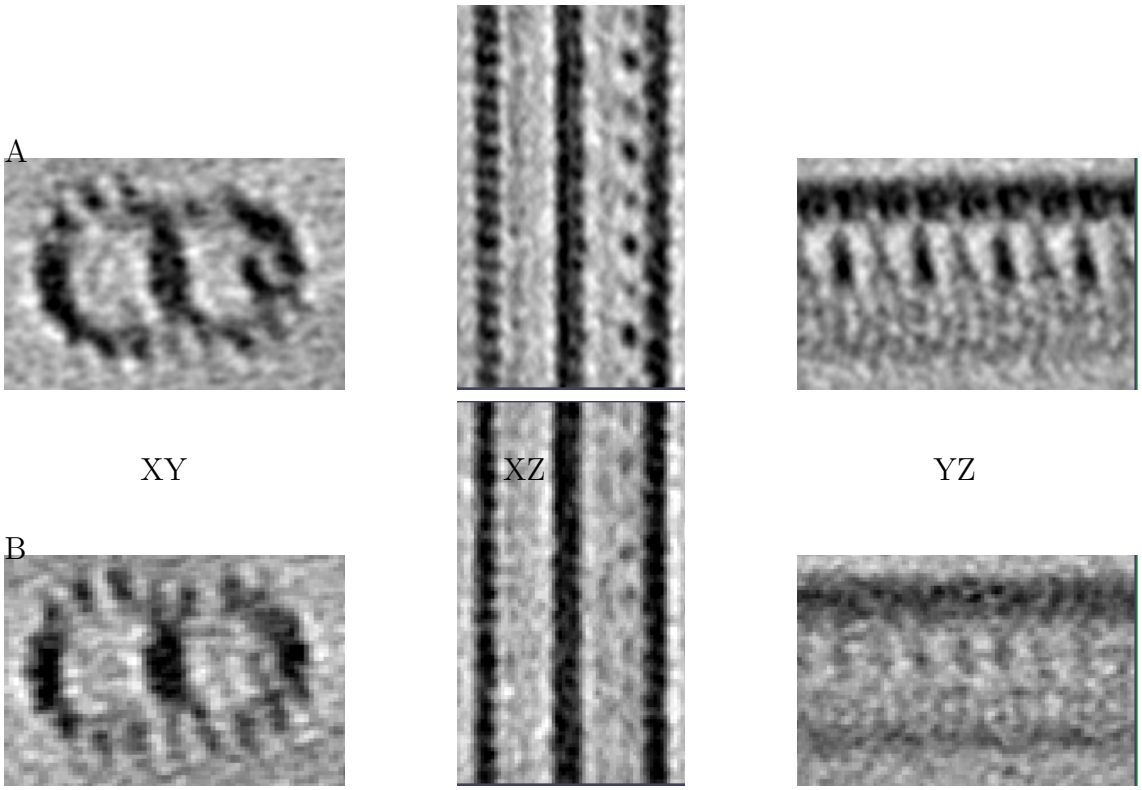


**Figure 4.11:** Comparison between the average structure obtained using only singletons (B) and incorporating pairwise potentials (A). Projection XZ visually shows a difference in the vertical spot-like motif repeated every 160Å on the right side tubule.

noise in the case of the translation and due to missing wedge in the case of the rotation. The pairwise approach helps impose spatial constraints to effectively select a location and orientation along the "plateau" of possibilities. Without these constraints the location along the microtubule direction by finding the maximum of the singleton is inconsistent and we can not resolve some of the periodicities present in the microtubules.

## 4.5 Discussion and limitations

In this chapter, we have described how to align hundreds of $3D$ images in order to obtain higher resolution structures combining the common information between them. The problem of subtomogram alignment is the same as $3D$ volume registration with the specific

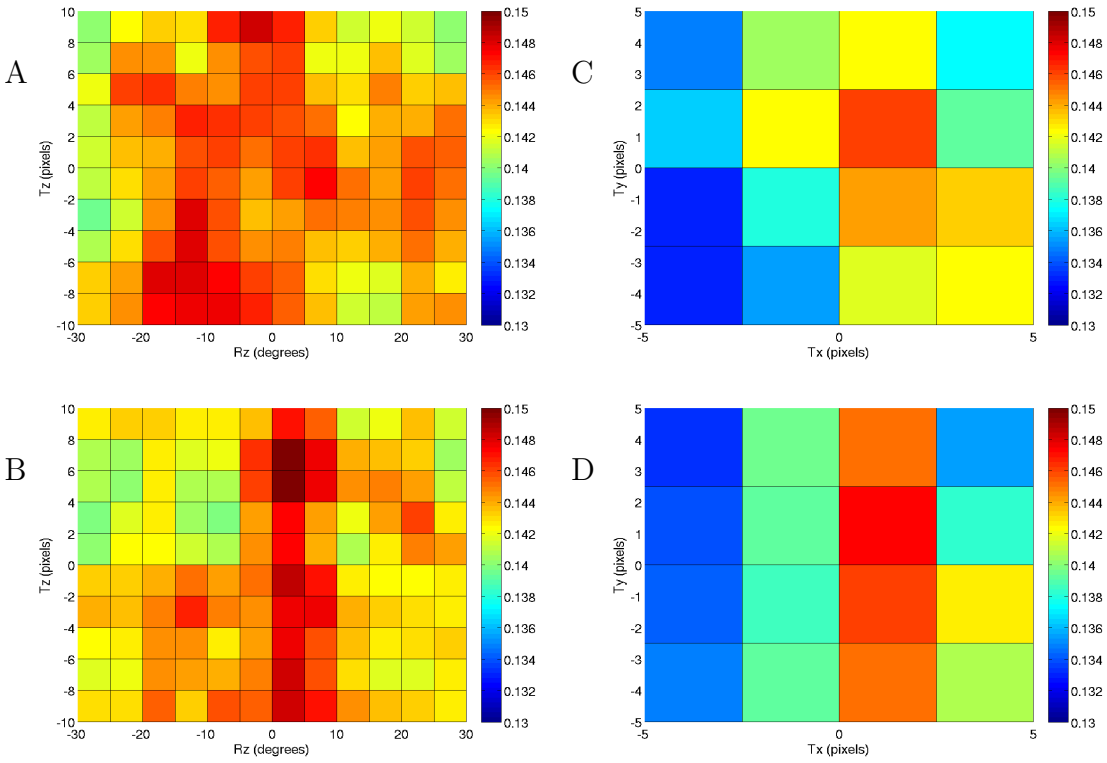


**Figure 4.12:** Grid search for $\psi_i$ around alignment location and orientation for two adjacent boxes along the microtubule. (A) Shows variation along the direction of the mid-axis of the microtubule ($T_z$) and rotating around the microtubule axis ($R_z$). The maximum of $\psi_i$ is unclear, with changes of up 15 degrees for $R_z$. (B) Same as (A) for an adjacent box. Here the maximum is more clear, which will help deciding the maximum in (A) using pairwise constraint $\psi_{ij}$. (C) Same as in (A) but translating the box perpendicular to the mid-axis of the microtubule ($T_x$ and $T_y$). (D) Same as in (C) for an adjacent box.

challenges of low $SNR$ and missing wedge from $CET$ images. We have shown how to use sparsity and local spatial correlations between rigid structures in the volume to improve the alignment results, which in turn improves resolved structures.

One of the advantages of the methodology is that it uses references from the data itself, avoiding the reference bias problem that occurs when external templates are used.

Moreover, results shown in Fig. 4.10 are obtained by processing each group of subtomograms completely independently, which validates the resolution estimates given for the averages. Finally, because the steps described in the previous section require minimal user intervention and we have a high-throughput pipeline to acquire tomograms as described in Chapter 3, we can consider the idea of obtaining averages for microtubule structures from different cells to analyze the differences at near-atomic resolution. The next chapter presents a case study to apply all these ideas to analyze the differences of the cell wall in *Caulobacter crescentus*, where dozens of tomograms and thousands of boxes need to be processed to obtain the results.

However, among the issues presented in Section 2.3 and in Fig. 4.2, this chapter has not addressed heterogeneity of biological structures. Fig. 4.13 shows how group A and B seem to present a subtle difference in the structure of the inner decoration of the microtubule. However, if we split the data into two different groups A' and B', with each new group A' and B' containing different microtubules than the original groups A and B, we obtain the averages shown in Fig. 4.14. With this new partition, no structural differences are shown in the inner decorations. Even if other studies in doublet microtubules have shown different inner decorations for microtubules belonging to the same axoneme [PS00], we still need to try to answer the question whether the differences we see represent real biological hetereogeneity in the data or if it is an artifact produced by noise. Unfortunately, the dissimilarity score $TCCC$ is too global to capture these small inner decorations in our low $SNR$ subtomograms when compared with the strong protofilaments features defining the edge of the microtubule. Thus, we can not distinguish if these subtle structural differences really exist in each individual subtomogram. In the next chapter, we will see how this problem with subtle structural features appears again and how we design an *ad hoc* solution for it. However, further research is necessary to develop metrics that can separate boxes in different classes based on local subtle structural features and to devise validation methods that separate variance from structural conformations and from noise.

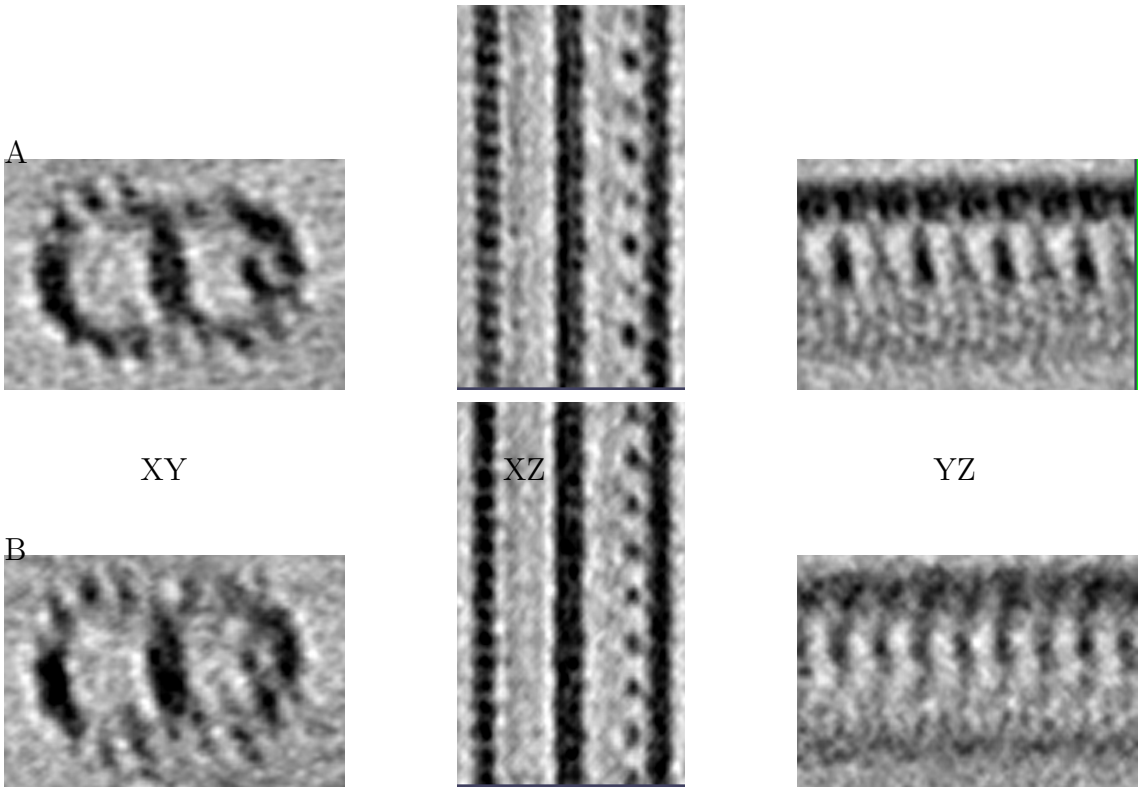


**Figure 4.13:** (A) Microtubule averaged structure shown in Fig. 4.11. Average was obtained using 650 boxes. (B) Microtubule averaged structure obtained using a different set of 500 boxes. (A) and (B) were used to compute the FSC shown in Fig. 4.10. The inner decoration of the right side microtubule presents structural differences clearly noticeable when comparing YZ projections: in (A) we have straight and tilted protuberances while in (B) they seem to have a vertical S-shape.

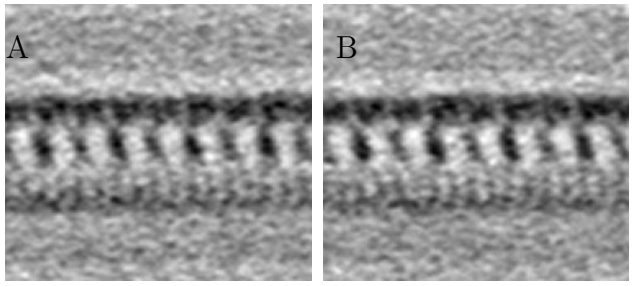


**Figure 4.14:** New YZ plane of microtubule averages after randomly mixing the two populations of microtubules shown in Fig. 4.13. (A) YZ projection of the average obtained with 575 boxes. (B) YZ projection of the average obtained with 575 boxes. In both averages the structure from group A in Fig. 4.13 seems to dominate.

# Chapter 5

# Analysis of Caulobacter crescentus S-layer

In this chapter, we present a case study of the surface layer (S-layer) structure of the *Caulobacter crescentus* bacterium using cryo-electron tomography of whole cells. In particular, we show how by combining all the methodology explained in previous chapters we obtain new insights. Among our findings are the presence of short range order; the coexistence of spatially localized regions with double and single S-layer formations in the same cell; and the determination of the N and C terminus in the RsaA protein combining nanogold labeling and subtomogram averaging. However, this chapter also shows the strengths and weaknesses of the $MRF$ probabilistic framework that we realized after acquiring, aligning, reconstructing and analyzing 26 tomograms in detail.

Surface layers are the outermost cell wall component in many archaea and bacteria and have been studied for more than 30 years [Sle78, Bev79]. Most S-layers are composed of a single protein or glycoprotein species that self-organizes into $2D$ lattices of different sizes with different symmetries, such as square or hexagonal (Fig. 5.1). This geometrical arrangement is almost the only commonality between species since sequence homology between proteins is low and functionality differs in many cases. In many archaea the S-layer is the only cell wall component, so they may have a role in shape determination.

However, in bacteria like Caulobacter crescentus, the role more likely related to protection against a variety of predatorial assaults [BPS$^+$97]. Understanding what is the role of the S-layer and its interaction with other cell wall elements and the outer media is still an open question.

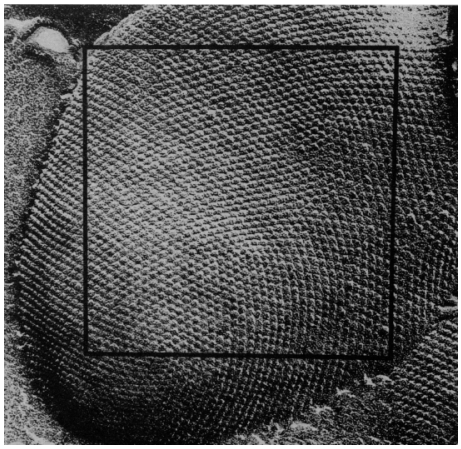

**Figure 5.1:** Example of S-layer image by freeze-etching electron microscope displaying the characteristic $2D$ lattice crystalline arrangement. Figure from [PMS91].

The densely packed $2D$ geometrical arrangement is not only aesthetically pleasant, but it has opened the door to multiple applications in nanotechnology [SSMP94, SHI$^+$07] and medical therapy such as anti-HIV microbicide development [NLL$^+$10] and cancer therapy [BAN$^+$06]. The main idea is to be able to display different proteins in specific points of the S-layer structure without breaking the self-assembly property between subunits, so we can create dense affinity arrays. Researchers have achieved this goal successfully in both *in vitro* and *in vivo* S-layers [Smi08].

The basic structure of the cell wall for gram negative bacteria like *Caulobacter crescentus* is shown in Fig. 5.2B and C. The main four elements ordered from external to internal are: S-layer, outer membrane ($OM$), peptidoglycan and inner membrane ($IM$). In *Cc.* each S-layer subunit is composed by six RsaA monomers [BNS97b] that form a characteristic hexagonal core with p6 symmetry and self-assemble to form a $2D$ lattice with p3 symmetry in junction points between hexagonal cores (Fig. 5.2A). The cores

attach to the *OM* by divalent calcium ions [WKRS94] that bridge RsaA protein with lipopolysaccharides (LPS) molecules.

Even if much about S-layer is known (the reader is referred to [BG91, CDE$^+$05, EP98, SH02] for different in depth reviews) there are still many open questions. For example, higher resolution structural information is needed to better understand which sites should be used to display different proteins and how the S-layer interacts with the cell-wall. To date, work by Pavkov et al. [PET$^+$08] and by Norville et al. [NKK$^+$07] represent the highest-resolution examples of S-layer proteins, although they were all achieved in *in vitro* studies. In fact, most of the structural studies have been performed in negative stained isolated S-layers. Quoting Engelhardt [Eng07]: "Functional aspects have usually been investigated with isolated S-layer sheets or proteins, which disregards the interactions between S-layers and the underlying cell envelope components."

In this chapter we present a quantitative analysis *in situ* of *Cc.* S-layer using *CET* of whole cells. In particular, we report two analysis: the first one studies global properties of the S-layer such as lattice regularity and periodicity using tomograms where the field of view contains the whole bacterium (Fig. 1.2). The second one tries to achieve atomic resolution information about the S-layer structure using subtomogram averaging on datasets with higher magnification, where the field of view contains only the stalk, which is the thinnest part of *Caulobacter crescentus*. However, subtomogram averaging is not enough to resolve primary structures in this case. Thus, we image mutant strains of *Cc.* with cysteines in different positions of the RsaA protein and deposit nanogold in the growing media, which should bind to the cysteines. At 1-2 nm in size the nanogold is not visible in individual CET images of whole cells. However, we will show that the nanogold can be seen after averaging hundreds of S-layer subunits due to the regularity of the S-layer structure, effectively increasing the resolution of *CET* imaging.

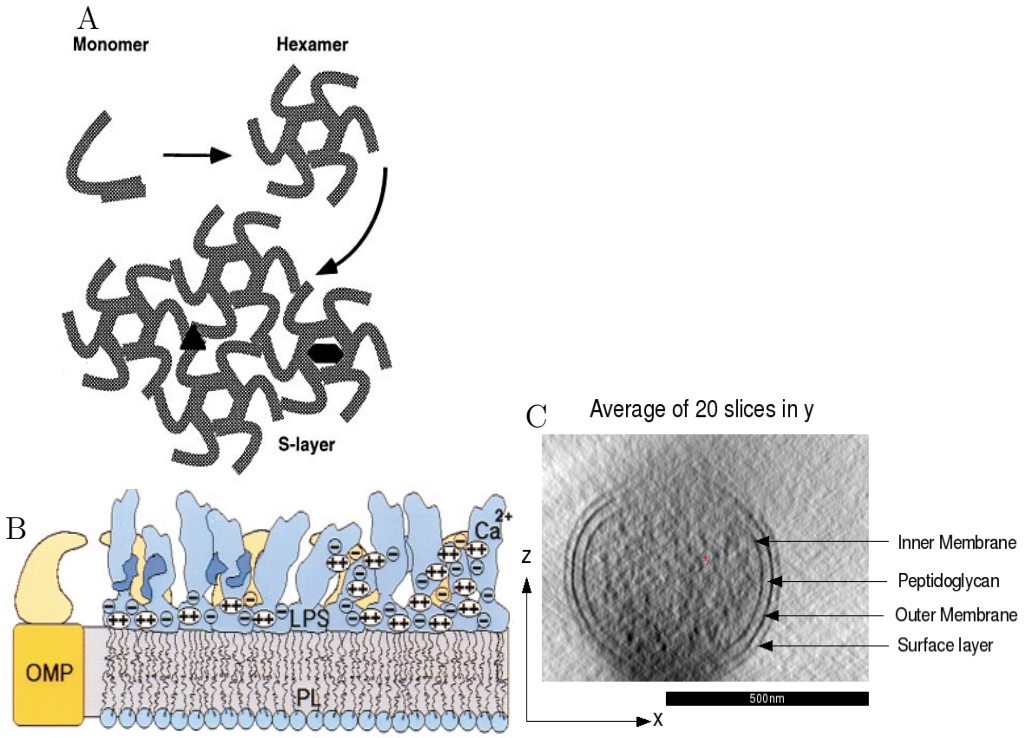


**Figure 5.2:** (A) Schematic from [BNS97b] showing how six RsaA monomers build a hexagonal S-layer subunit in *Caulobacter crescentus*. (B) Schematic from [EP98] showing cell wall structure of gram negative bacteria like *Caulobacter Crescentus*. The S-layer units attach to the outer membrane ($OM$) by divalent calcium ions that bridge RsaA protein with lipopolysaccharides (LPS) molecules. (C) Cross Section of a *Caulobacter crescentus* tomogram to show the cell wall components. The effect of the missing wedge blurring the features on the top and bottom of the cell is obvious. Black bar represents 500nm.

## 5.1 Image analysis pipeline

To generate the data presented in this chapter, a total of 26 tilt series were acquired, aligned and reconstructed. Appendix A contains the materials and methods related to these datasets. Briefly, 6 datasets were of *Cc.* whole cells with pixel size 12Å/pixel and -12$\mu$m defocus while 20 were of *Cc.* stalks at higher magnification mostly with pixel size 6.8Å/pixel and defocus between -3.8$\mu$m and -6.6$\mu$m. 19 out of the 26 datasets (70%) were successfully aligned with RAPTOR without any user intervention; 2 datasets (7%)

required some manual adjustments to the tracking provided by RAPTOR in high tilt images due to the low contrast of the gold beads and the remaining 6 datasets (23%) were aligned manually using IMOD [Mas97] because RAPTOR failed to provide good alignment. All the failures occurred in high magnification tomograms where less than 6 fiducials were present in the sample and they were all localized in a small region of the image. All the tomograms were reconstructed using weighted backprojection provided by IMOD and then screened by visual inspection. Only 16 out of 26 were used (61%) to report the results shown in this chapter. The criterion to disregard tomograms was the lack of apparent regular S-layer subunits on the surface of the cell wall. These percentages show how there are still many factors in the *CET* pipeline that can affect the final reconstruction quality in a real case study, and there is still work needed ahead to improve high-throughput *CET* techniques if we want to scale the pipeline to hundreds of thousands of tomograms.

To study global properties of the S-layer lattice, we analyze tomograms at 25,000X where the field of view contains the bacterium. For each tomogram, we need to perform the same two tasks described in Chapter 4 for microtubule doublets: First, detect the S-layer, and second, extract S-layer subunits from the detected surface. In the following paragraphs, we describe in detail how to perform these two tasks.

The first step towards finding the S-layer is to detect the cell wall in each tomogram. BLASTED [MHA$^+$10] enables a user to only label one cross section (like the one shown in Fig. 5.2C) and the algorithm sequentially tracks the *OM* with high accuracy by inferring boundary points and cell wall shape at the same time. Once we have the location of the *OM* in each tomogram we parametrize the surface using bicubic B-splines [MP87]. Because *Cc.* has a banana-like shape, cylindrical coordinates with a cubic axis (instead of a straight axis) is the best system of coordinates to parametrize the surface. The parametrization allows us to operate on the surface using differential geometry concepts like geodesics or tangent planes, which are necessary for the quantitative analysis presented in Section 5.2. Appendix C describes the parametrization in more detail.

Using the surface parametrization and the alignment algorithm presented in Chapter 4 we refine the location of S-layer subunits in each tomogram. Briefly, we divide the surface obtained above in small square patches that can contain up to two hexagons of the S-layer. We adjust each patch location and orientation using template matching to locate the characteristic hexagonal pattern of the S-layer. The $MRF$ imposes smoothness constraints on the surface between neighboring patches to avoid errors due to noisy detections. This step is particularly important for the top and bottom parts of the cell wall, where the missing wedge has blurred the cell wall features. However, we can still find the S-layer hexagons in the XY cross-sections using the Thresholded Contrained Cross-Correlation ($TCCC$) dissimilarity score described in Chapter 4.

Once we have refined the location of S-layer subunits we can combine them using subtomogram averaging to obtain better resolution of the S-layer subunits. We use $TCCC$ as a dissimilarity scores between subtomograms and the alignment refinement procedure described in Section 4.3. We also use a two steps coarse-to-fine procedure to improve alignment results. In particular, 16761 boxes of size 76nm x 76nm x 76nm were extracted from four different datasets of whole cells (Fig. 5.3). We obtained an initial low resolution template by averaging 200 boxes selected manually by visual inspection from the extracted surface in one dataset. The selection criterion was the overall appearance and location around the cell to uniformly cover missing wedge orientations. Then we crosscorrelated the template along the extracted S-layer surface using the $TCCC$ metric to select points where the dissimilarity score had a local minimum.

First, we ran a cycle of coarse alignment with a heavy band pass filter to obtain a rough orientation of all the particles. Only spatial frequencies between 12nm and 30nm were used for this alignment. After convergence, we selected the boxes where the outer membrane was visible as an anchor point and the $TCCC$ value[1] was above 0.25. The total number of boxes left for alignment was 3777. The dramatic decrease in the total number of subtomograms is due to two reasons: first, misalignment of the boxes due to noise, and

[1] $TCCC$ values range from -1 to +1 as other normalized cross-correlation like measures.

second, only boxes below 60 degrees from the XY plane contain the $OM$ and $IM$ due to missing wedge effects (Fig. 5.3). However, since the missing wedge is only from $\pm 60^{circ}$ and we are recovering boxes with orientations within 60 degrees of the XY plane, we still fill all the Fourier space after averaging the number of boxes left.

Before disregarding the boxes above 60 degrees from XY plane, we realized by inspecting averages from different tomograms that the most distinguishable structural feature was a "single" versus a "double" S-layer on top of the $OM$ (Fig. 5.4). Therefore, it is reasonable to disregard the boxes where the blurring due to missing wedge is perpendicular to the $OM$ since they do not contain relevant information with respect to the sought structural difference. However, as already happened in Section 4.5 for microtubules, the $TCCC$ dissimilarity score was not able to distinguish that subtle feature in low $SNR$ images in comparison with the strong $OM$ edge. In this particular scenario, we use an *ad hoc* solution by projecting each individual subtomogram along the X and Y axis in order to obtain 1D profiles of density along the normal direction with respect to the S-layer (Fig. 5.4). The projection aggregates densities, which effectively increases the $SNR$ in the 1D profiles to easily visualize one or two "dips" (Fig. 5.4B and C) and classify the particles according to single or double S-layer. Using this classification, 2089 boxes had double S-layer and 1688 had single S-layer.

As in Chapter 4, we split the subtomograms in two groups from different tomograms containing double S-layer to calculate $FSC$ resolution curves without introducing artificial correlations. The first group contains 1106 boxes from tomograms 1 and 2 and the second one contains 983 boxes from tomograms 3 and 4. For each subset of boxes we generated an initial template by averaging 200 particles randomly selected and we ran a separate second cycle of alignment without any band pass filter. We followed exactly the same procedure to align and average the 1688 boxes displaying single S-layer, although in this case we did not split the dataset in two because it was better to average all the boxes to obtain a better reconstruction instead of repeating the FSC calculations.

In the second part of this case study, we also need to detect, align and average S-layer,

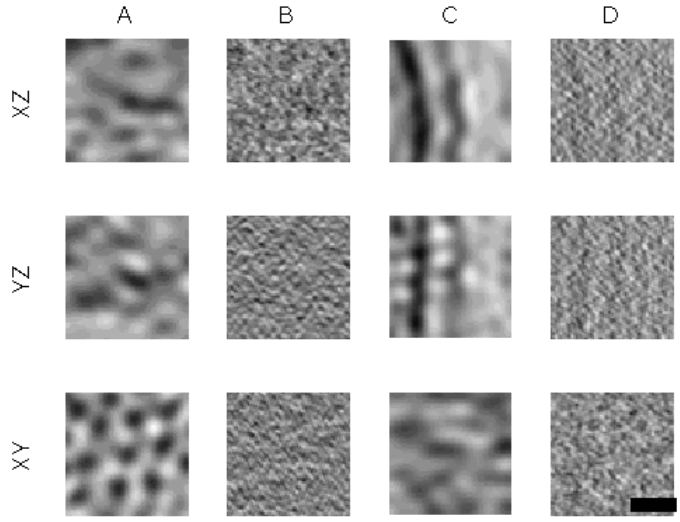


**Figure 5.3:** XY, XZ,YZ projections of two aligned boxes from raw data for *Caulobacter crescentus* S-layer. Bar represents 20nm. (A) Same subtomogram as in (B) but with a low-pass filter to visualize main features. The hexagonal pattern is preserved (XY) but the outer membrane feature is blurred out. (B) Box where missing wedge blurring affects perpendicular to the cell wall. Noise masks all the features. (C) Same subtomogram as in (D) but with a low-pass filter to visualize main features. Outer membrane is visible but hexagonal pattern is blurred out. (D) Box where missing wedge blurring affects tangential to the cell wall. Again noise masks all the features.

but for high magnification tomograms (40,000X) of *Cc.* stalks. For the most part, the steps in this case are very similar to the ones described above. However, the S-layer was manually selected using IMOD [KMM96] capabilities, since the segmentation of stalks is straightforward because they are short straight cylinders. Moreover, we did not need to disregard subtomograms above $60^{circ}$ because of the smaller diameter of the stalk. For each RsaA strain with a cysteine inserted in a different location of the protein chain, we manually selected 200 boxes to obtain an initial template from the data itself. Again, the selection criterion was the overall appearance and location around the cell to uniformly cover missing wedge orientations. At 40,000X , the exposure time at higher magnifications needs to be reduced in order to radiate the same total amount of dose, which increases the noise. To increase the $SNR$ we binned by two all the tomograms containing stalks after reconstruction. Once we have an initial template for each strain, we can search for

that template in the parametrized surface as done in whole cells. In total, we extracted 879 boxes for wild type datasets, 619 boxes for Cys944 datasets and 196 boxes for Cys277 tomograms. The reason for the lower amount of boxes in Cys277 tomograms is due to the fact that regular S-layer areas were not as abundant in this strain. Once we have all the locations for the initial boxes we run a two step coarse-to-fine alignment procedure as desc

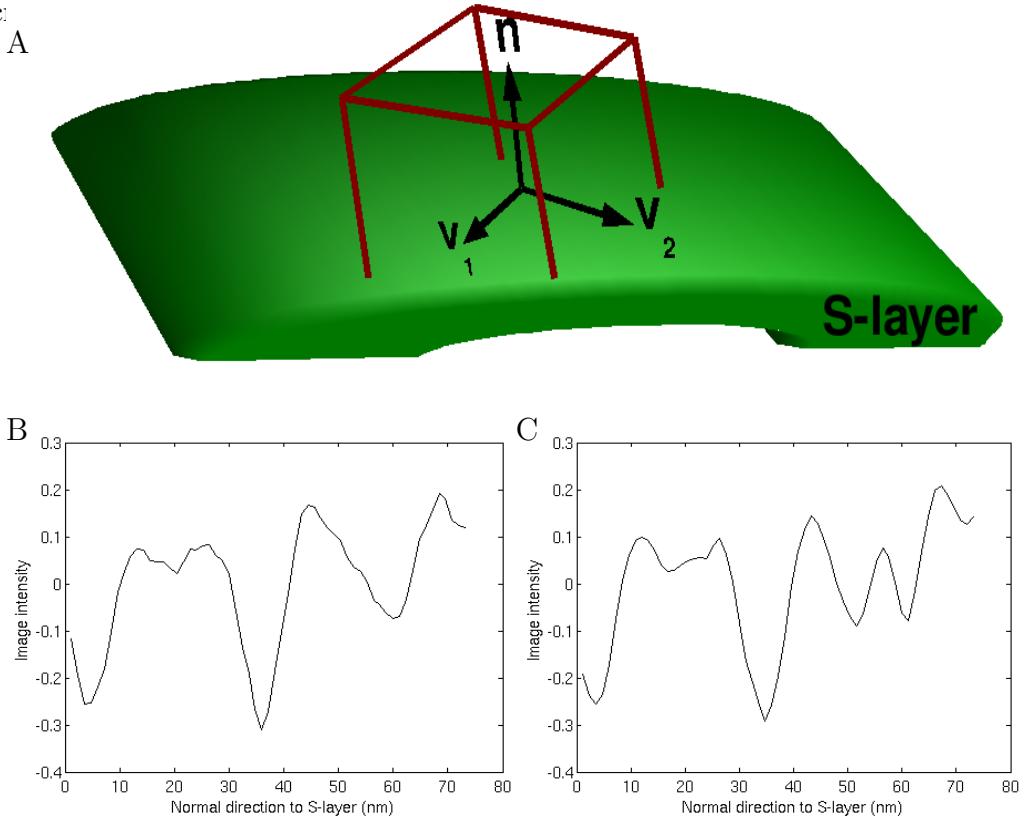


**Figure 5.4:** (A) Schematic showing local coordinates for boxes on S-layer surface. The profiles shown in (B) and (C) are the result of adding intensity values along $V_1$ and $V_2$ in a single subtomogram. In the profiles we can see the $IM$ "dip" at 4nm, the $OM$ "dip" at 35nm an the S-layer "dip" between 50nm and 65nm. (B) shows a single "dip" versus (C) that shows a double "dip". These profiles help classifying each subtomogram into single or double S-layer for subtomogram averaging.

One of the main problems to verify results is how to visualize the S-layer in each tomogram. The S-layer is a non-planar surface embedded in a $3D$ density map. Therefore, visualization of the whole S-layer at once by standard tools displaying slices through the

volume is not possible. We developed a customized code to generate Visualization Toolkit (VTK) files [SML97] that can be visualized in Paraview [Hen04]. The main idea is to triangulate the parametrized S-layer surface into triangles with area smaller than 1 voxel. Then, we can assign an interpolated value of the closest density in the $3D$ volume to each triangle and render the surface. Results of the rendering can be seen in Fig. 5.5. Such a visualization is very useful to select S-layer subunits locations and analyze the short and long-range order of the lattice. Moreover, it allows us to validate the results of the surface extraction algorithm (Fig. 5.6).

## 5.2 Results

In the first part of this study we focus on global properties of the S-layer lattice structure by obtaining tomograms of whole cells *in situ* and rendering the the S-layer. From previous papers [SA82] we know that *Cc.* S-layer is built by self-assembly of RsaA monomers. Thus, we compare two populations, wild type versus an RsaA mutant producing 25% more RsaA.

Fig. 5.5 shows a typical visualization of the rendered S-layer. Even though it is known that the S-layer covers all the bacterial surface [SA82], the hexagonal pattern is not present everywhere. Moreover, even in places where we see the hexagonal pattern, it is not as crystalline as shown in isolated S-layer [SEV$^+$92] sheets. In order to verify these results we completed two procedures: first, we rendered sequential extrusions of the surface (Fig.5.6), which confirmed we were not missing any part of the S-layer. Second, we created a heat map of the template matching score between each point on the surface and the expected hexagonal pattern (Fig. 5.5). The results quantify the fact that there are areas with the characteristic local hexagonal pattern and areas with no presence of it.

The second part of the analysis on global properties of the S-layer was to offer quantitative measurements of different parameters related to lattice regularity. Fig. 5.7A shows the main methodology to achieve that: we can draw intensity profiles along the axis of the p6 symmetry to measure periodicity and autocorrelation functions. The main advantage of our surface parametrization using cubic splines in cylindrical coordinates is that we

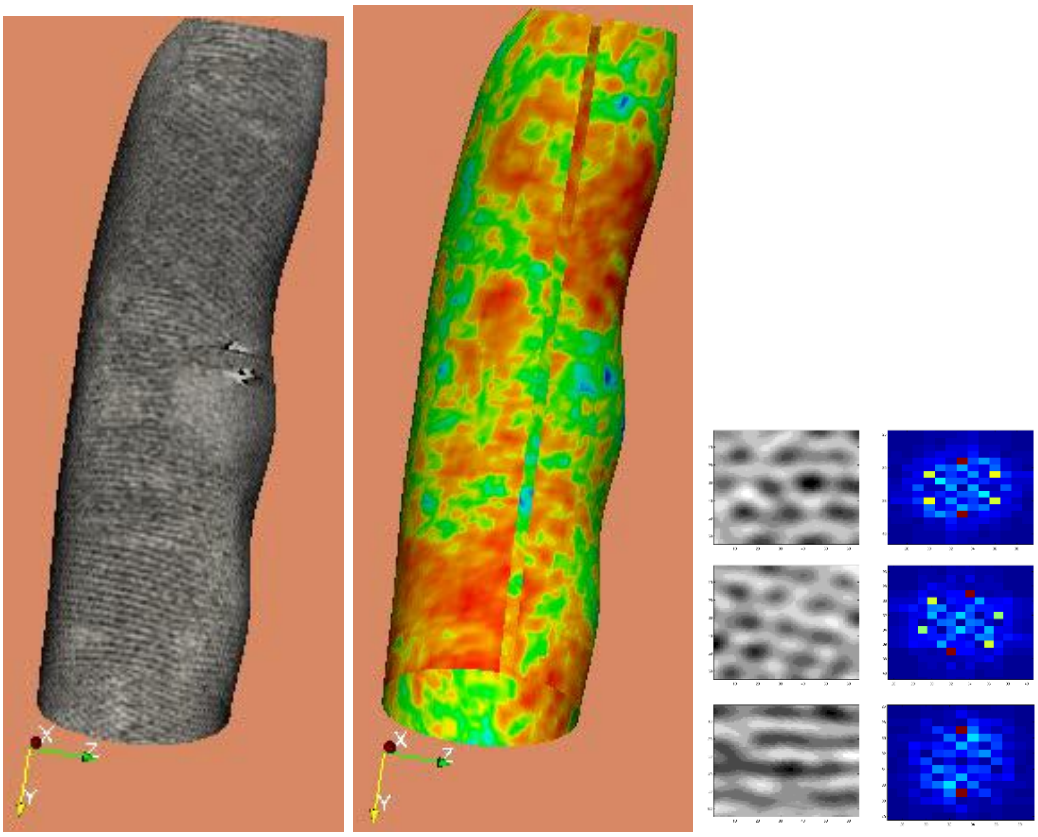

**Figure 5.5:** Left: rendering of S-layer surface from intersection of the surface with the $3D$ density values of the tomogram. Patches of S-layer are visible on the surface but no long-range coherence is apparent and hexagonal regularity is not found. Center: heat map of template matching scores on the same surface looking for a hexagonal pattern. Red color indicates high correlation and blue indicates low. Heat map agrees with the visualization to show that regular hexagonal pattern is not present in all the surface. Right: local tangential patches to the S-layer showing hexagonal pattern and the magnitude of their FFT. 6 peaks in Frequency domain represent the hexagonal pattern although the missing wedge blurs some of them.

can measure distances in $3D$ along paths on the surface, so we do not need to physically flatten the S-layer shape to perform quantitative analysis *in situ*.

Fig. 5.7B shows the averaged autocorrelation for 222 intensity profiles from three different wild type datasets. It shows a correlation with periodic ups and downs and an exponential decay of the envelope, which indicates short-range order. In particular, if we

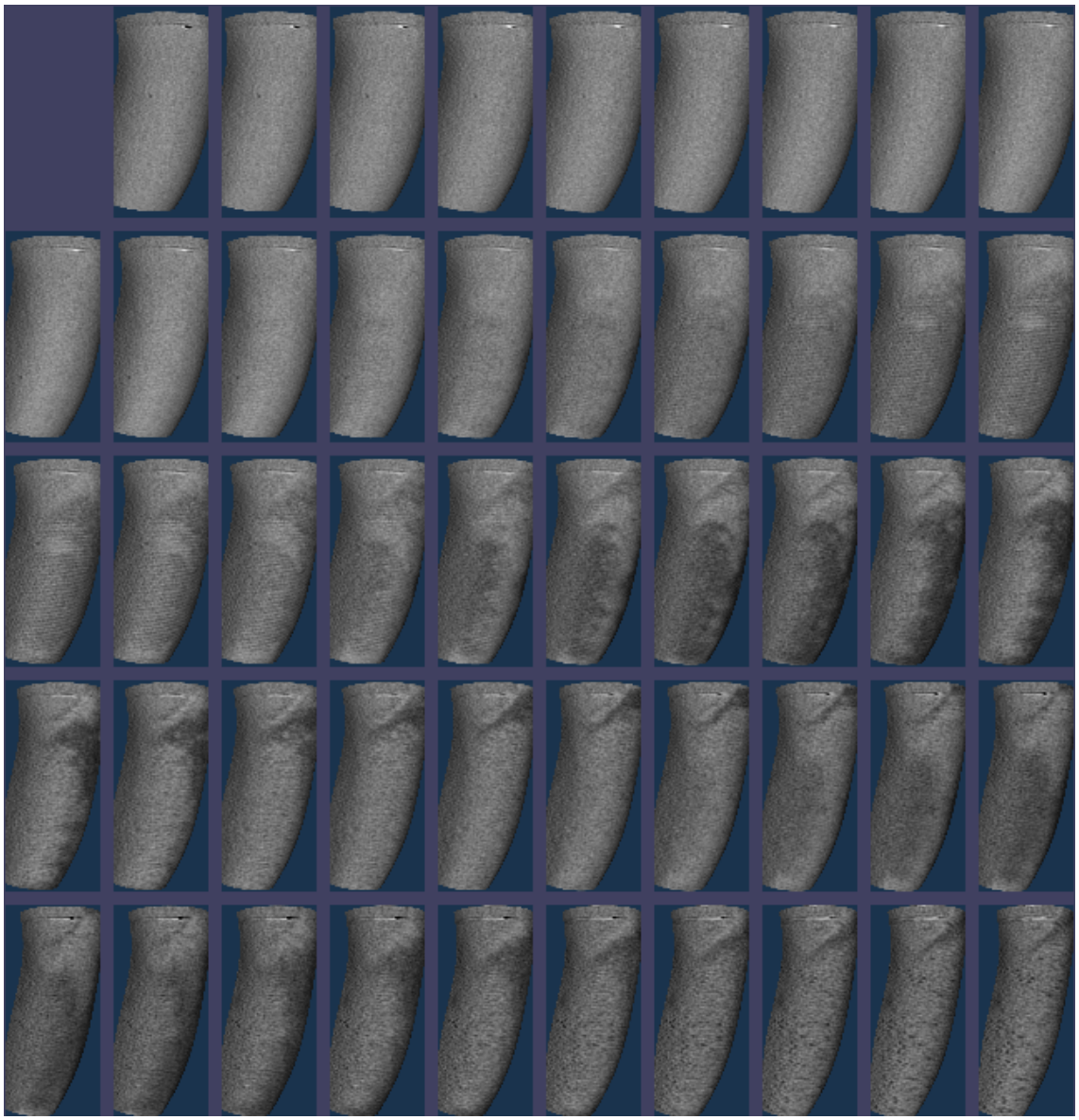

**Figure 5.6:** Rendering of extruded surfaces from the original detected S-layer. Top left is the outermost surface (ice) while bottom right is the inner most surface ($IM$). Sequence shows a visual test to check that the detection process has not missed S-layer components. Distance between rendered surfaces is approx. 2nm.

fit an exponential decay $Ae^{-\frac{x}{\tau}}$ to the envelope of the autocorrelation function (Fig. 5.7B-red line) we obtain $\tau = 0.015\text{nm}^{-1}$. This analysis indicates that, on average, every three periods we find a perturbation (phase shift) in the lattice that breaks the possibility

of long-range order. However, these are the averaged statistics: Fig. 5.7E and F show individual profiles following a perfect periodic pattern and individual profiles with very short range order respectively.

Fig. 5.7B and 5.7C show a comparison of the range order and the periodicity between wild type cells and RsaA mutant (using 177 profiles from three different datasets). Results from both cases are almost indistinguishable. The S-layer periodicity presents a very symmetric distribution with mean 22.8nm and a spread of ±2nm, which shows again a more amorphous structure than previously observed in different isolated S-layers.

Next we focus on the underlying $3D$ structure of each S-layer subunit. Looking at the global lattice structure in the previous section we were not able to find significant differences between wild type cells and RsaA mutant. Using subtomogram alignment techniques described in Chapter 4 we can obtain a better resolution estimation of the subunit structure and analyze possible structural differences. Fig. 5.8 shows the resulting average structures displaying the structural differences between "single" and "double" S-layer. Fig. 5.9 shows $FSC$ curves comparing the average for the "double" S-layer obtained with one group of 1106 sumtomograms and a second group of 983 subtomograms. We can see that the resolution of the average obtained using $TCCC$ metric described in Section 4.2.1 is very close to the first zero of the theoretical $CTF$, which suggests that adding more boxes to the averaging will not improve resolution. Moreover, we compare $FSC$ curves obtained using two existing metrics in the literature [FPSF08, BSL$^+$08] and using $TCCC$ to confirm the results shown in Chapter 4 with synthetic data: in scenarios with $SNR < 0.1$, such as tomograms of whole cells, $TCCC$ improves alignment accuracy. In this case, the improvement is approximately 5% in the resolution curves.

Layers of S-layers one on top of the other have been reported in the literature [SM90, SM83] for other bacteria. Our results show this phenomenon for the first time in cell wall of *Cc.*, which combines areas with a single S-layer and other areas with double S-layer, although they are spatially separated. Analyzing these results two observations stand out: First, for the wild type datasets, 47% of the subtomograms contained a single S-layer while

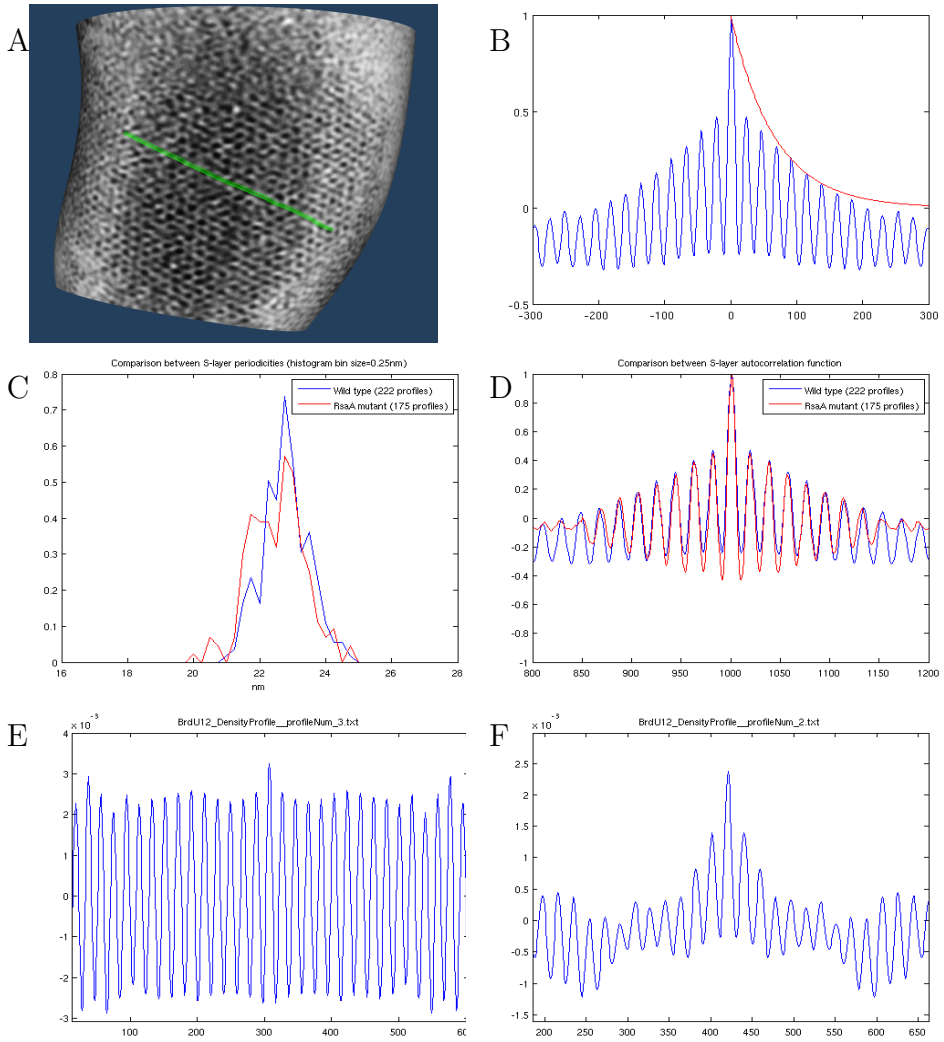


**Figure 5.7:** (A)Example of how line profiles (green line) were obtained from the rendered S-layer. (B) Averaged autocorrelation function across 222 profiles on wild type cells. The autocorrelation decays exponentially ($\tau = 60$nm) showing short-range order. (C) Histogram of S-layer periodicity extracted from line profiles for wild type (blue) and RsaA mutant (red). Both present a peak at 22.8nm and a spread of $\pm$2nm, showing the elasticity of the S-layer. (D) Comparison of averaged autocorrelation function between wild type (blue) and RsaA mutant (red). Both present the same short-range order. (E) Autocorrelation function for individual profile showing long range order. (F) Autocorrelation function for individual profile showing very short range order.

53% contained a double one. On the other hand, the RsaA mutant showed 36% of the subtomograms with single S-layer versus 64% with double S-layer. Second, the distance between the $OM$ and the outermost S-layer seems to be similar in both classes (Fig 5.8). However, a detailed quantitative analysis using the projected 1D profiles results on the distribution shown in Fig. 5.10. For the double S-layer, the mean relative distance between the $OM$ and the outermost S-layer is 25.3nm with standard deviation 1.3nm, while for the single S-layer the mean is 24.2nm and standard deviation 1.4nm. Moreover, the distance between each layer of S-layer has mean 8.7nm and standard deviation 1.7nm, measured from the distance between "dips" in Fig. 5.10C.

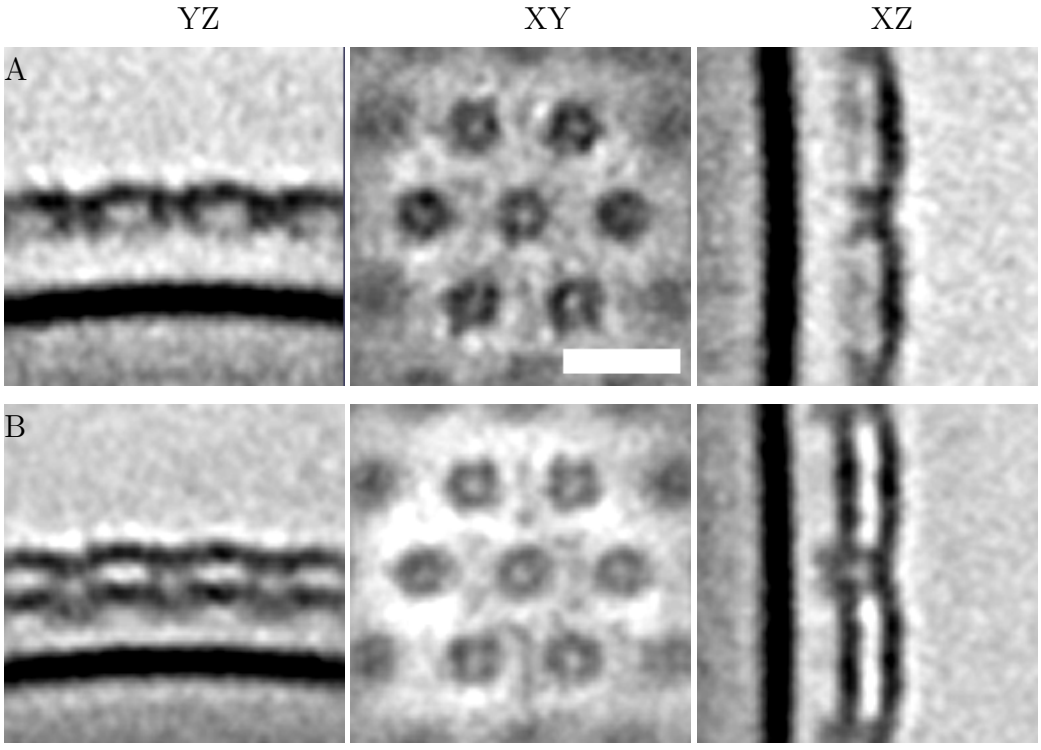


**Figure 5.8:** (A) Average structure of single S-layer using 1688 boxes. (B) Average structure of double S-layer using 1106 boxes. The structural difference between (A) and (B) is clear. YZ, XY and XZ represent perpendicular planes to display the results of the $3D$ density maps. White bar represents 20nm.

In the case of the double S-layer, we would like to answer if each layer is a mirrored

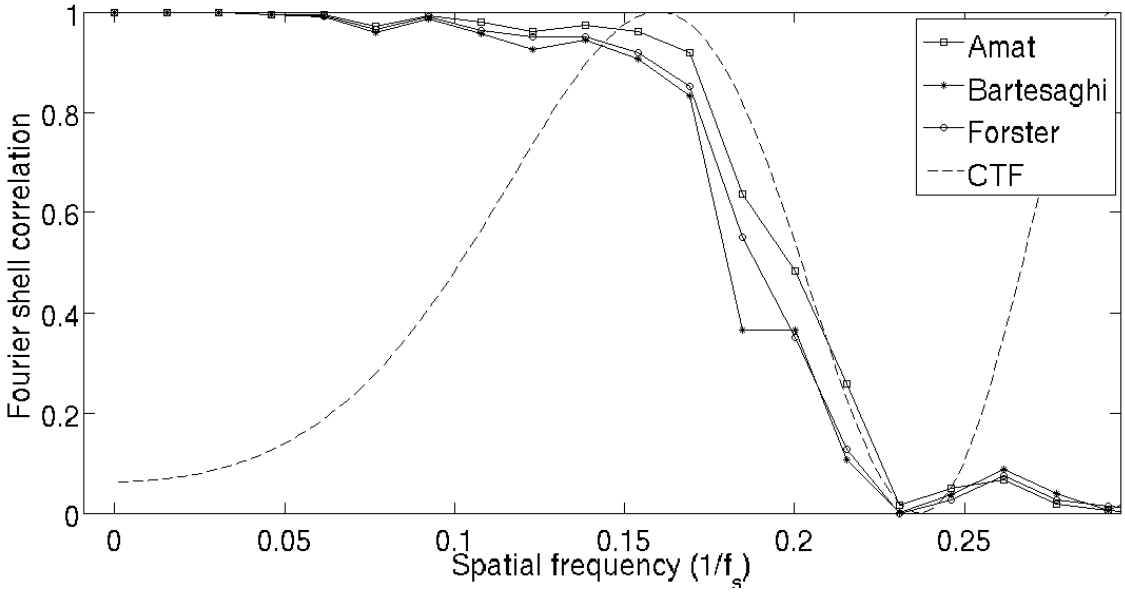


**Figure 5.9:** FSC curve for *Caulobacter crescentus* surface layer (continuous line) for three different dissimilarity scores: from [FPSF08] (circles), from Eq. (2) in [BSL$^+$08] (dot) and TCCC (square). Theoretical CTF for the imaging conditions assuming amplitude contrast of 0.2 (dashed line) and $Cs = 3.2$mm. Formulas for theoretical CTF from [Fra06]-Ch. 3).

image of each other or if the layers are built like scaffolds. Resolving the handedness in the arm links connecting each p6 symmetry center would answer this question. Unfortunately, the thickness of *Cc.* whole cells limits the resolution and the handedness can not be resolved in our tomograms. Another question we would like to answer is where is the beginning (N terminus) and the ending (C terminus) of the RsaA monomers forming S-layer subunits (Fig. 5.2A). Again, the resolution obtained using whole cells is not enough to provide primary or secondary structural information that will answer this question.

In order to resolve both questions, we change the imaging conditions to improve attainable resolution by obtaining tomograms at higher magnification (40,000X) of the *Cc.* stalks, which are the thinnest regions of the bacteria. Furthermore, we use immunogold labeling techniques to extract atomic resolution information about the S-layer structure. As described in the introduction, we grow strains of *Cc.* with cysteines in different positions (Cys277 and Cys944) of the RsaA chain with 1024 aminoacids, and nanogold should

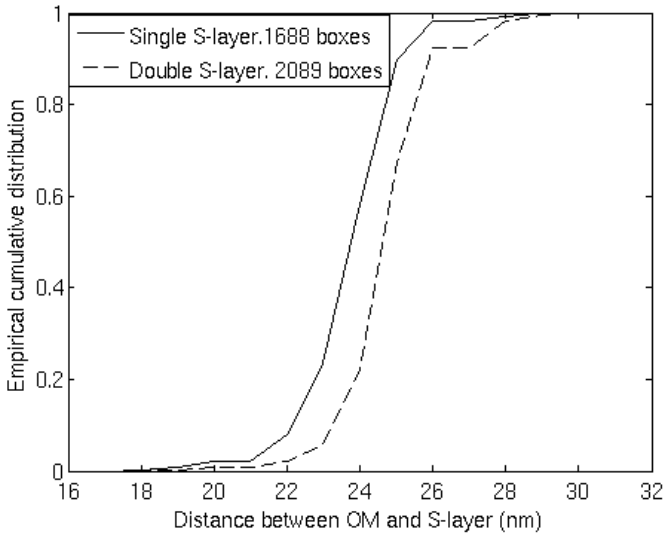


**Figure 5.10:** Empirical cumulative distribution of the distance between outermost S-layer and *OM* for single (continuous line) and double (dashed line) S-layer cases measured from 1D high $SNR$ profiles shown in Fig. 5.3. Single S-layer is closer than outermost double S-layer.

bind to these locations if it is added to the growing media. If we are able to consistently locate the nanogold with respect to the center of p6 symmetry in different strains, we will determine the location of the N and C terminus of an RsaA monomer in the S-layer. In general, labeling techniques for *EM* have been difficult to implement for general structural studies as explained in Section 2.3.6. However, nanogold labeling has been used in previous S-layer studies for *Cc.* [SA82] because the S-layer is the frontier between the bacterium and the external medium, and any nanogold diluted in the sample will bind to the targeted cysteine in the RsaA protein building the S-layer.

After screening the stalk tomograms, the nanogold was not apparent to the naked eye. However, if it is true that the location of the nanogold is the same in different S-layer subunits, we should be able to observe it after aligning and averaging multiple subunits. Fig. 5.11 shows how the subtomogram averaging of two different strains (Cys277 and Cys944) present high density spots in different locations of the S-layer structure, which indicates the location where the nanogold binds to the cysteine. In particular, using the

schematic in Fig. 5.2A as a guidance to interpret the average structure, the N terminus is located in the center of the hexagonal core (p6 symmetry) and the C terminus is located in the links between S-layer subunits (p3 symmetry). The average of wild type is needed as a control experiment to show that without labeling, the high density spots disappear. Moreover, the presence of the outer membrane as a high density element in all the averages is useful as a reference to compare relative density between datasets.

Unfortunately, we were not able to answer the handedness question with these new tomograms because we need to average thousands of boxes to resolve the links between S-layer subunits due to the very low $SNR$ of high magnification tomograms. From each stalk we can extract approximately between 200 and 600 S-layer subunits, and only a small percentage of those contain double S-layer, since the new strains do not overproduce RsaA. Thus, with the current pipeline we will need to acquire dozens of tomograms more to average enough boxes. Moreover, the handedness is even more subtle as a structural feature than the "double" versus "single" S-layer. In order to resolve it, we need to make sure that each subtomogram is oriented coherently with the rest, otherwise the handedness will be blurred by the averaging. This constraint is difficult to impose on each individual box separately. The handedness structure shows the limitations of the current $CET$ imaging and data analysis process.

## 5.3 Discussion

This chapter has presented one of the first studies of S-layer close to its native state using $CET$ of whole cells to obtain a global picture of the lattice-like structure with few nanometers resolution. To accomplish this task we used several image analysis techniques to process large amounts of data, visualize curved surfaces in $3D$ density volumes and average similar structures in low $SNR$ scenarios to obtain higher resolution images. This native state analysis presents a different picture from the structural analysis performed on isolated S-layers by shedding them from the membranes. In particular, the $CET$ datasets show an S-layer with a lattice structure closer to an amorphous solid rather than a crystal.

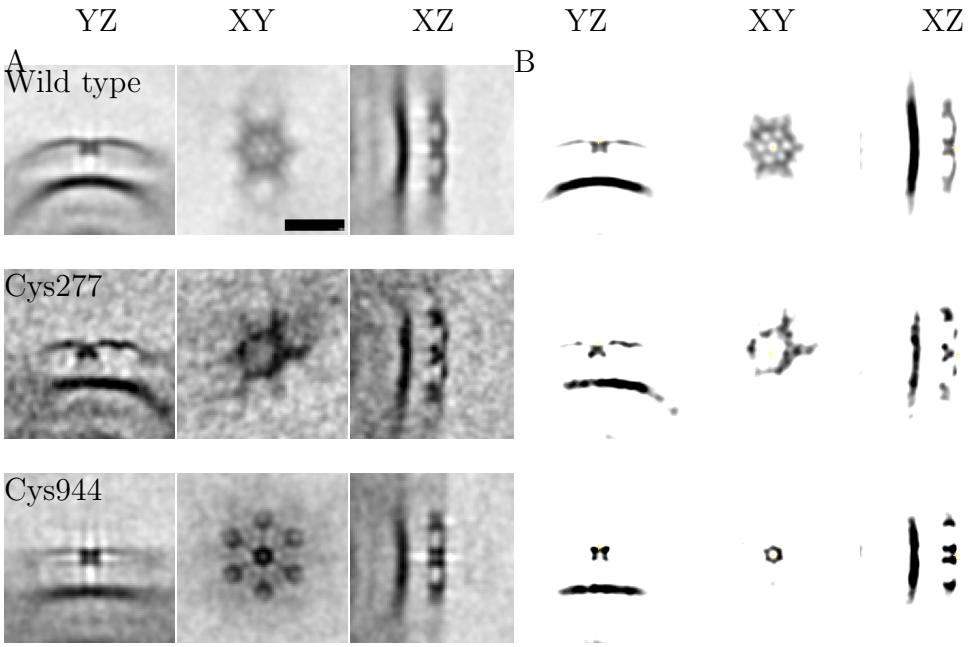


**Figure 5.11:** (A)Comparison between averages from different *Cc.* strains using high magnification (40,000X) tomograms of stalks. Wild type obtained using 879 boxes. Cys277 obtained using 196 boxes and Cys944 obtained using 619 boxes. (B) Same in images as (A) but enhancing contrast to visualize high density spots. *OM* present in all the image serves as a reference to compare relative densities between averages. Cys277 concentrates the high density spots between hexagonal cores while Cys944 concentrates them in the hexagonal core itself. YZ, XY and XZ indicate three different perpendicular planes to visualize the $3D$ average. Black bar represents 20nm.

Considering that the S-layer subunits are affected by two interactions, subunit to subunit and subunit to *OM*, and that isolated S-layers present self-assembling properties into $2D$ crystalline lattices, we deduce that the anchoring of the S-layer to the cell wall is not as regular. Moreover, the fact that hexagonal structures are not observed in all the surface even if RsaA is present reinforces the idea of a dynamic process to create new S-layers [PMS91]. Fig. 5.5 shows several "ridges" on the S-layer surface where different growth areas merge. Our quantitative analysis also agrees with Fig.6 by Smit et al. [SGGA81], which presents images from *Cc.* CB15 S-layer using freeze-etching techniques and briefly states: "In this typical view, there are indications of a periodic structure on the surface,

but the impression of precise, long-range order was not seen.".

Our main reference for the structural analysis in whole cell tomograms is [SEV$^+$92], which presents a detailed study of the *Cc.* CB15 strain using negative-stain electron tomography on isolated S-layer sheets. Their measured periodicity between subunit cores is 22nm, while [SGGA81] had previously measured 23.5nm. Fig. 5.7 not only shows a mean of 22.8nm, which is between the two previous measurements, but it also displays a distribution of possible periodicities, which represents the elasticity in the S-layer links between subunits. Our observations of average thickness 8.7nm disagrees with the estimates from Smit et al. [SEV$^+$92] of 7nm, although the measured standard deviation of 1.7nm makes that number feasible given the elasticity of the S-layer. Overall, the structure between *in vitro* and *in vivo* are similar.

One of the main applications in nanotechnology for S-layers is to generate high density arrays to display useful proteins. Therefore, it is important to research if certain mutant strains can display even a denser amount of them. One first obvious choice is to produce more RsaA. However, our results show that *Cc.* stores the extra RsaA in a second layer of S-layer instead of packing the hexagon more regularly or reducing the distance between subunit cores. It is still an open question how this second layer is formed, since it seems that the initial single layer is the one closer to the external medium. In the same direction, the localization of the N and C terminus in each RsaA protein forming the S-layer of *Cc.*, which was unknown until now, it is also useful to design the arrays since each terminus shows different symmetries (p3 or p6), creating different lattices.

All these results have been possible thanks to the combination of hundreds of subtomograms, which allows the discovery of signals such as nanogold labeling that are not detectable in a single tomogram. However, in order to obtain and process all the necessary data we needed to automate every step of the pipeline as much as possible. Even if the methods presented throughout this thesis are targeted to be robust against the low $SNR$, missing wedge and heterogeneity in $CET$ dataset, this chapter has presented examples where there is still room for improvement. Based on these experiences, the next

section presents a discussion of possible ideas for future directions in order to improve some of the methodology discussed in previous chapters.

# Chapter 6

# Conclusion

We have seen how $CET$ delivers unique $3D$ information that can help understand the function and complex spatial relationships of microtubules and other large structures within cells. However, there are several specific challenges associated with $CET$ images, such as low $SNR$, missing data and heterogeneity, that make the data analysis challenging. In this thesis, we have shown a probabilistic framework based on $MRF$ that helps aligning pairs of $2D$ and $3D$ $CET$ images. The graphical model adds local spatial correlations between adjacent features on top of standard template matching techniques to guide the alignment and make it more robust against noise. In particular, Chapter 5 showed a case study of the surface layer of *Caulobacter crescentus* bacterium where the alignment techniques were used to visualize the S-layer in its native state without sample preparation artifacts.

Looking at the pipeline to acquire and process tomographic data, it was clear that one of the bottlenecks for real high-throughput tomography was the alignment of tilt series using fiducial markers. Because of the low $SNR$ in the datasets and the lack of ground truth, high-throughput is crucial in $CET$ to obtain more images and validate the structural findings. Our software RAPTOR has improved the state-of-the-art alignment methods for $CET$ samples, shifting the computational burden from data acquisition and reconstruction to data analysis. This method shows the power of considering contextual

information among multiple pairs of close markers to make a final decision, instead of looking at the best correspondence for each marker separately.

Among the possible methodologies that high-throughput tomography opens for data analysis, one of the most successful ones in the recent years has been subtomogram averaging. However, good selection of initial locations and orientations to avoid reference bias and to obtain high resolution structures is still a challenge mostly done by hand. In Chapter 4 we show how we can extend the same $MRF$ framework used in $2D$ images to $3D$ using minor modifications. Basically, we need to incorporate more degrees of freedom for possible deformations and modified the dissimilarity score to include the missing wedge effect and noise statistics. One of the critical changes was to choose a sparse representation of the volumes to make the comparison more robust against noise.

Even with all this progress, we are still far from reaching the full potential of cryo-electron tomography. Arguably, one of the main dreams for electron tomography is to be able to reconstruct mammalian cells *in situ* at molecular resolution in $3D$ by stitching together hundreds of tomograms [MVMH04]). Comparing the structural differences between healthy and disease conditions in combination with genomic, proteomic, and biochemical data will bring key insights to develop new drugs. However, many steps of the $CET$ pipeline shown in this thesis need to be improved further to reach that goal. For example, we have seen in Chapter 5 how RAPTOR achieves 70% success rate in a real case study with few dozens tomograms. If we want to process thousands of tomograms, 30% of alignment by hand is still a bottleneck.

Some ideas in how to improve RAPTOR were hinted at the end of Chapter 3 and are compiled here as future directions. First, approximate inference tends to break down when the number of markers in the image is more than a hundred: marker locations become too close to each other to establish good correspondence and the $MRF$ does not return clear assignments (Fig. 3.15). Thus, global correspondence returns short overlapping partial trajectories instead of global ones. In the future, the probabilistic model and the projection map estimation should be combined to mitigate these effects. One possible direction to

do this is to modify the greedy iterative refinement described in Section 3.4.3. Instead of picking the assignment with higher singleton score that is closer to the reprojection point indicated by the projection model, we should do a second pass of pairwise correspondence incorporating the information from the projection model. Since the projection model provides accurate information of marker locations, the roll-off constraint $\sigma_1$ can be made much tighter, which will help in the inference procedure to distinguish between close by markers.

Second, marker detection success depends on the type of images we are trying to align. $NCC$ works very well for $CET$ images because they have low contrast and the only high contrast features are the gold beads. However, in plastic embedded samples or in soft X-ray tomography, the images have much higher contrast, especially the edges of stained membranes, and the $NCC$ with the gold bead template returns too many false positives along these edges. Different template matching techniques need to be tested in the future for different types of samples.

Subtomogram alignment and classification will also play a key role to discover structural differences between different cell populations. We have seen two examples in Chapter 5 to resolve "double" versus "single" S-layer and the location of nanogold labeling. However, we have also shown how many subtle differences are elusive to current classification methods based on cross-correlation. Strong edges, such as the $OM$ in the S-layer or the protofilaments in the microtubule structure, dominate the dissimilarity score and mask other subtler differences. In Chapter 5 we have seen already small *ad hoc* modifications that suggest possible future directions for subtomogram alignment, averaging and classification where a combination of multiple local features could mitigate these effects. However, one of the main problems is that noise masks these local features in each individual subtomogram (they only become apparent after alignment and averaging). Techniques like the projection in XY to obtain enhanced 1D profiles across the cell wall presented in Chapter 5 solve the two problems at once: we have higher $SNR$ in each subtomogram because we are averaging along sides within each box and we obtain local features of

the cell wall. Finally, approaches that integrate alignment and classification in a single framework should also improve performance. Scheres et al. [SMVC09] presented a first step in this direction by describing a maximum likelihood approach solved by expectation-maximization. However, they had to simplify the generative model for subtomograms to make the computation of the likelihood computationally tractable. More research in this direction is needed to achieve better resolutions.

The case study in Chapter 5 also showed how modifying the imaging parameters can improve subtomogram alignment for certain goals. Currently, microscopists are used to set parameters in order to extract as much "visual" information from each individual tomogram as possible. However, if we know the project at hand requires averaging, maximizing "visual" information per single tomogram might not be optimal. For example, in single particle protocols it is common to obtain micrographs at different defocus for two reasons: first, the high defocus images lack any high frequency information but present high contrast, which makes particle detection easy. This step helps generate a first low resolution template that can be used for detection and alignment of other particles. Second, the simulation shown in Fig. 4.9 shows that even with perfect alignment, the first zero of the CTF always affects resolution. Therefore, combining data with nulls at different locations allows to recover information in all frequency bands. A similar approach could be applied to subtomogram averaging by acquiring tomograms with different defocus. The main assumption here is that we are able to estimate and correct the CTF in each tomogram, which is an active research topic.

Finally, one of the holy grails to exploit all the information generated by high-throughput tomography would be the ability to detect similar structures in new tomograms once we have labeled several examples of the structure we are looking for by hand. This problem is known as object detection in the computer vision literature. However, challenges in $CET$ images presented here make clear that this is a hard problem to solve: managing all the degrees of freedom in $3D$ efficiently, accounting for the missing wedge effect and using robust features against noise set strong requirements in the methodology design. However, if we

aspire to reach the full potential of electron tomography, we need to overcome these problems. It is not reasonable to expect that users will select thousands of particles from $3D$ volumes to compare structural differences. Approaches based on machine learning techniques are currently being investigated in single particle projects [AHGM09, SRA$^+$09]). Because in $2D$ micrographs there is no missing wedge and the degrees of freedom are reduced, making the extension to tomograms is not straight forward.

In summary, $CET$ is still a growing field and will be able to deliver many new biological insights in the future since neither the hardware nor the software has reached its full potential. As the amount of acquired data increases, the possibilities for developing and applying new statistical image processing techniques grows too. We hope this thesis has shown several examples that can help in this direction.

# Appendix A

# Materials and methods

Here we describe all the materials and methods for the different datasets presented in the results section used of chapter 5..

## A.1 Bacterial strains and growth conditions

*Caulobacter crescentus* strain JS 1023 is a CB15 NA1000 derivative, modified by inactivation of rsaA, sapA and introduction of repBAC from the broad host range plasmid RSF1010. The native sapA and rsaA genes were inactivated by gene replacement with an internal deletion (sapA) or an amber mutation (rsaA353∅B). The gene replacements and chromosomal insertion of repBAC was accomplished using pK18mobsacB ([STJ$^+$94]) and as previously described ([FNS07]). Escherichia coli strain DH5 alpha (Invitrogen, Carlsbad, CA) was grown at 37°C in LB medium (1% tryptone, 0.5% NaCl, 0.5% yeast extract). C. crescentus JS1023 grown in PYE medium (0.2% peptone, 0.1% yeast extract, 0.01% $CaCl_2$, 0.02% $MgSO_4$), at 30°C. For growth on solid medium, agar was added at 1.3% (wt/vol). When needed, media contained chloramphenicol at 20 $\mu$g/ml (*E. coli*) or 2 $\mu$g/ml (*C. crescentus*). Electroporation of *C. crescentus* was performed as previously described ([GS91]). Plasmid DNA was isolated using a Gene Jet miniprep plasmid isolation kit (Fermentas).

## A.2 Introducing unique cysteines into rsaA

The oligonucleotides JN CYS N-1 (5' GATCGTCATGCGCTAGCC 3') and JN CYS N-2 (5' GATCGGCTAGCGCATGAC 3') when annealed and ligated into a BamH1 site, destroyed the original BamH1 site and, in the correct orientation, coded for the amino acids SSCASR. This segment was cloned into the unique BamH1 sites that were previously installed in rsaA at positions corresponding to amino acids 277 and 353 ([BNS97a]). Similarly, the oligonucleotides CYS10F (5' TCATGTACAG 3') and CYS10R (5' TCCTGTACAT 3'), encoding the amino acids SCTG were annealed and ligated into a BamH1 site corresponding to amino acid 944. All introductions were confirmed by DNA sequencing. A BglII site was introduced into rsaA at a position corresponding to amino acid 2 by utilizing an EcoN1 site 45 bp 3' to the ATG start codon. An oligonucleotide that included the EcoN1 site and the 5 sequence of rsaA except that it altered the 2nd and 3rd codons to create a BglII site and changed the codons from Ala-Tyr to Arg-Ser. Using this oligonucleotide and an oligonucleotide matching sequence 5' to an upstream EcoR1 site for PCR amplification resulted in a product that could be substituted for native sequence as an EcoR1/EcoN1 segment. Annealed JN CYS N-1 and JN CYS N-2 oligonucleotides were then ligated into the BglII site. The modified versions of rsaA were transferred to p4A as EcoR1/HindIII segments ([NDDS07]). The resulting plasmids were introduced into JS1023 by electroporation.

## A.3 Protein isolation and separation

Native and modified RsaA protein was recovered from the cell surface by low-pH extraction, as described previously ([WSS92]). Proteins were visualized using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 7.5% gels and stained with Coomassie brilliant blue R.

## A.4 Bacterial growth media

Different *Caulobacter crescentus* strains were grown to log phase before cryoplunging or reacting with nanogold in peptone yeast extract (PYE) media. Media were supplemented with kanamycin (5 $\mu$g/ml) or chloramphenicol (1 $\mu$g/ml) as necessary. Transcription from the xylose promoter was induced by adding 0.3% xylose to the growth media. Cells were grown overnight in PYE. They were then washed and diluted in M2G and incubated until the $A_{660}$ reached  0.2. Cells were then diluted to an $A_{660}$ of  0.05 in M2GX (M2G plus 0.3% xylose) to induce transcription from the xylose promoter. During induction, cultures were diluted as necessary to keep the $A_{660}$ at <0.2.

## A.5 Nanogold labeling

Cells were grown in PYE media to an optical density of approximately 0.4 to 0.6 at 620nm ($ODU_{620} \sim 0.4$  0.6 ). The cell number is estimated to be at most $10^8$ or less. A rough calculation of the number of S-layer subunits in a *Cc.* cell results in the estimate of approximately 14,000 subunits[1]. Using these calculations, monomaleimido nanogold® labeling reagent (Nanoprobes) was dissolved in 1ml of double distilled water. 100$\mu$l of the nanogold suspension was mixed with 900$\mu$l of the whole cell culture, then thoroughly mixed with a pipetted several times for approximately 30 min. The resulting sample was cryo-plunged as described below.

## A.6 Cryo-electron microscopy specimen preparation

Aliquots of 5$\mu$l were taken directly from the cultures and placed onto lacey carbon grids (Ted Pella 01881) that were pre-treated by glow-discharge. The Formvar support was not removed from the lacey carbon. The grids were manually blotted and plunged into liquid

[1]We approximate a *Cc.* cell as a cylinder of diameter 500nm and length of 1000nm to calculate the total S-layer area. Assuming distance of 22.8nm between S-layer subunits, we can calculate the area of each subunit. Dividing both areas results in approximately 14,000 subunits per cell.

ethane by a compressed air piston, then stored in liquid nitrogen. For CET, fresh sample aliquots were deposited onto support grids pre-loaded with 10nm colloidal gold particles.

## A.7 Cryo-electron tomography

Images were acquired on a JEOL3100 electron microscope equipped a FEG electron source operating at 300 kV, an Omega energy filter, a Gatan 795 2Kx2K CCD camera, and cryo-transfer stage. The stage was cooled using liquid nitrogen to 80K. Approximately 10% of the data was acquired using a Gatan 795 4Kx4K CCD camera mounted at the exit of an Electron Decelerator [DM08]. The decelerator was operated at 248 kV, resulting in images formed by a 52 kV electron beam at the CCD. A total of 26 tomographic tilt series were acquired under low dose conditions using the program Serial-EM [Mas05], typically over an angular range between +62 deg and -62 deg, ±2 deg with increments of 1° or 2°. Tilt series of whole cells images were acquired using a magnification of 25K at the 2Kx2K CCD giving a pixel size of 1.2nm at the specimen. Tilt series of stalks images were acquired using a magnification of 40K giving a pixel size of 0.373nm at the 4Kx4K Decelerator-coupled CCD and of 0.685nm at the 2Kx2K CCD. The data acquired on the 4Kx4K CCD was subsequently binned by a factor of two. Underfocus values ranged between 3.6$\mu$m ± 0.5$\mu$m to 12$\mu$m ± 0.5$\mu$m, depending on the goal of the data set, and energy filter widths ranged between 22 eV to 28 eV, also depending on the data set. For all data sets the maximum dose used per complete tilt series was approximately 140 e-/Å$^2$, with typical values of approximately 100 e-/Å$^2$.

# Appendix B

# Noise model for CET volumes

In this appendix we derive a noise model for CET volumes. The noise model is useful for any sort of statistical image processing applied to CET volumes. In particular, define $p_i^j$ as the i-th pixel in the j-th projection of a tilt series. We assume as a first approximation that the noise on each pixel is independent from its neighbors and it is distributed as $p_i^j \sim Poisson(\lambda_i^j)$. Assuming we use a basic weighted back projection (WBP) algorithm as described in ([Fra06]-pp. 254) we have that each voxel in the tomogram is the sum of pixels from different projections convoluted with a weighting function $w$ to filter high noise and equalize Fourier components. Formally:

$$f_i = (\sum_{(i,j)\in\chi} p_i^j) * w \tag{B.1}$$

Where $\chi$ is the set of pixels in different projections that contribute to voxel $f_i$ in the reconstruction. The summation of independent Poisson random variables is a Poisson random variable such that $\sum_{(i,j)\in\chi} p_i^j \sim Poisson(\lambda_i)$ with $\lambda_i = \sum_{(i,j)\in\chi} \lambda_i^j$. Typically in ET we have 100+ projections, so $\lambda_i > 100$. In this regime, the Poisson distribution can be approximated very well by a Gaussian distribution with mean and variance equal to $\lambda_i$. The weighting filter $w$ adds correlation between neighboring voxels. Therefore, for each

tomogram $f$, we have the following noise model:

$$f = x + \varepsilon \tag{B.2}$$

Where $x$ is the true underlying signal and $\varepsilon$ is additive noise with Gaussian distribution $\varepsilon \sim N(\mathbf{0}, \Sigma)$. If a tomogram has $N$ voxels, then $\Sigma^{-1}$ is a $NxN$ sparse matrix with non-zero elements only in the diagonal and in the entries corresponding to neighboring pixels because the weighting function $w$ is a convolution in a local neighborhood. Moreover, the elements in the diagonal are all different, since the variance $\lambda_i$ is different for each voxel.

This model of additive noise defines a Gaussian Markov Random Field (GMRF) between $N$ voxels. Estimating the parameters of a GMRF with such a large $N$ as in our tomograms is not an easy computational task and it is an active research topic ([DGK08, FHT07, MJW06]). Thus, we approximate the noise model by considering only a diagonal covariance matrix with different variance for each voxel. In other words, we assume $\Sigma = diag(\sigma_1^2, \ldots, \sigma_N^2)$ considering only marginal noise statistics for each voxel[1]. This makes the computation much more efficient and yields good estimation results (see results in section 4.4 of this thesis). Notice that any affine scaling of the intensities in the tomogram does not change this noise model.

Fig. B.1 shows an example for the real data described in materials and methods for the *Cc.* S-layer. We obtain $\mu_i$ and $\sigma_i$ for one of the voxels using 500 aligned subtomograms and plot the theoretical cumulative distribution function (cdf) for the estimated values versus the empirical cdf obtained from the data. The agreement with the Gaussian assumption is very good even in the tails.

[1]Since $\varepsilon$ has a multidimensional Normal distribution, each of the marginals follows a univariate Normal distribution.

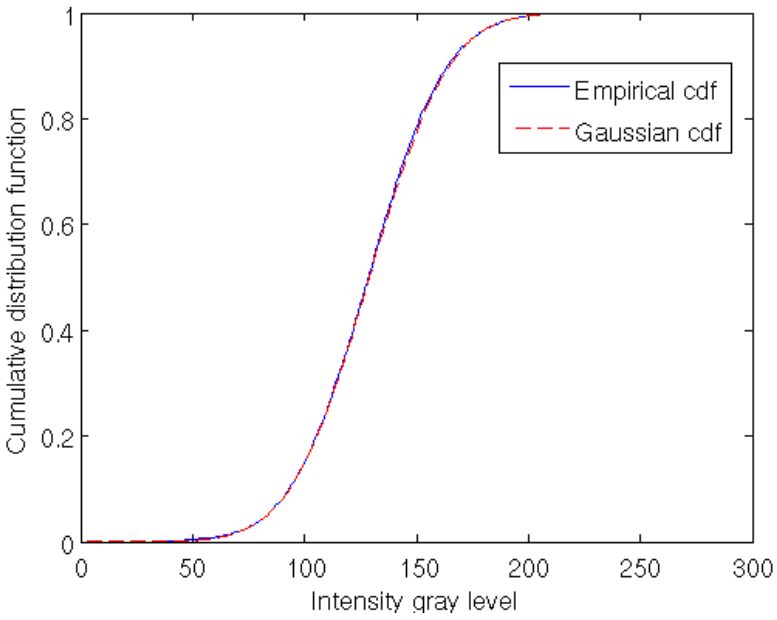


**Figure B.1:** Comparison between empirical cumulative distribution functions (cdf): first, obtained from 500 aligned subtomograms of *C.c.* whole cells for one voxel (blue continuous line); second, the analytical cdf of a Gaussian distribution (red dashed line) with $\mu$ and $\sigma$ estimated from the empirical data. The agreement is almost perfect making hard to distinguish both curves.

# Appendix C

# Cylindrical coordinates

In this appendix we described the coordinates system used to parametrize the surface of *Caulobacter crescentus* bacteria. Due to its banana-like shape the best representation is obtained using cylindrical coordinates with a polynomial axis that follows the mid-axis of the cell instead of a straight axis. Fig. C.1 shows a schematic of this set of coordinates.

Without loss of generality, we can assume the bacterium lies parallel to the XY plane. Thus, we can parametrize the mid-axis of the bacterium as follows

$$
\begin{aligned}
x(t) &= f(t) \\
y(t) &= t \\
z(t) &= z_0 \\
f(t) &= a + bt + ct^2 + dt^3
\end{aligned}
$$

where $t$ is the parameter that indicates movement along the mid-axis. Given the banana-like shape of *Cc.*., defining $f(t)$ as a cubic polynomial fits any sample. Fig. C.1 displays the mid-axis in green. Once we have the mid-axis, for each value of $t$ along the axis we can obtain a plane perpendicular to the tangent vector of the axis $(f'(t), 1, 0)$. Fig. C.1 displays this cross-section in red. In each cross-section we can define polar coordinates $r$ and $\theta$ as is done in standard cylindrical coordinates.

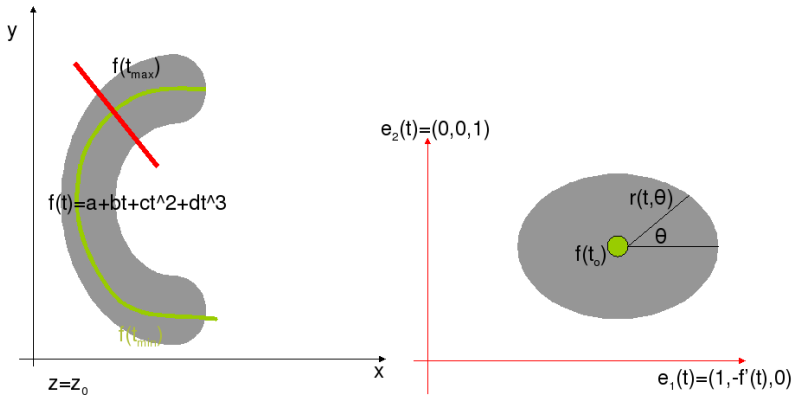


**Figure C.1:** Schematic representing cylindrical coordinates with cubic polynomial axis (green line) to parametrize the surface layer of *Caulobacter crescentus*. Without loss of generality, we assume the bacterium is parallel to the XY plane.

Using the ideas above we can parametrize the surface of *Caulobacter crescentus* S-layer as follows:

$$
\begin{aligned}
x(t,\theta) &= f(t) + r(t,\theta)\cdot\cos(\theta) \\
y(t,\theta) &= t - r(t,\theta)\cdot\cos(\theta)\cdot f'(t) \\
z(t,\theta) &= z_0 + r(t,\theta)\cdot\sin(\theta)
\end{aligned}
$$

The advantage is of this representation is that $r(t,\theta)$ is very close to a plane if the mid-axis has been selected appropriately. Therefore, $r(t,\theta)$ can be represented with bicubic splines with a small number of knots. The only constrain we have to impose is periodcity in $\theta$ to ensure a close surface. However, with this representation it is not possible to impose $\mathcal{C}^2$ continuity at the poles where $r(t,\theta)$ tends to zero. $\mathcal{C}^0$ continuity at the poles is sufficient for all the applications described in this thesis.

# Bibliography

[AHGM09] Pablo Arbelaez, Bong-Gyoon Han, Robert M. Glaeser, and Jitendra Malik. Protein complex analysis project (pcap): Automated particle picking in electron microscopy images. In *Genomics:GTL Awardee Workshop VII*, 2009.

[AMC+08] F. Amat, F. Moussavi, L. R. Comolli, G. Elidan, K. H. Downing, and M. Horowitz. Markov random field based automatic image alignment for electron tomography. *Journal of Structural biology*, 161(3):260–275, July 2008.

[AScPK04] Dragomir Anguelov, Praveen Srinivasan, Hoi cheung Pang, and Daphne Koller. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *Advances in neural information processing systems*, pages 33–40, 2004.

[BAN+06] P. K. Bhatnagar, A. Awasthi, J. F. Nomellini, J. Smit, and M. R. Suresh. Anti-tumor effects of the bacterium caulobacter crescentus in murine tumor models. *Cancer biology & therapy*, 5(5):485–91, 2006.

[Bau02] W. Baumeister. Electron tomography: towards visualizing the molecular organization of the cytoplasm. *Current Opinion in Structural Biology*, 12:679–684, 2002.

[Bev79] T. J. Beveridge. Surface arrays on the wall of sporosarcina ureae. *Journal of Bacteriology*, 139:1039–1048, 1979.

[BG91] Terry J. Beveridge and Lori L. Graham. Surface layers of bacteria. *Microbiological Reviews*, 55(4):684–705, December 1991.

[BHE01] Sami Brandt, Jukka Heikkonen, and Peter Engelhardt. Multiphase method for automatic alignment of transmission electron microscope images using markers. *Journal of Structural Biology*, 133(1):10 – 22, 2001.

[Bis07] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2nd edition, 2007.

[BK04] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.

[BM92] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *Pat. Anal. and Mach. Intel*, 14(2):239–256, 1992.

[BNS97a] W.H. Bingle, J.F. Nomellini, and J. Smit. Cell-surface display of a pseudomonas aeruginosa strain k pilin peptide within the paracrystalline s-layer of caulobacter crescentus. *Molecular Microbiology*, 26(2):277–288, 1997.

[BNS97b] W.H. Bingle, J.F. Nomellini, and J. Smit. Linker mutagenesis of the caulobacter crescentus s-layer protein: toward a definition of an n-terminal anchoring region and a c-terminal secretion signal and the potential for heterologous protein secretion. *Journal of Bacteriology*, 179(3):601–611, February 1997.

[BPS+97] T.J. Beveridge, P.H. Pouwels, M. Sra, A. Kotiranta, K. Lounatmaa, K. Kari, E. Kerosuo, M. Haapasalo, E.M. Egelseer, I. Schocher, U.B. Sleytr,

L. Morelli, M.L. Callegari, J.F. Nomellini, W.H. Bingle, J. Smit, E. Leibovitz, M. Lemaire, I. Miras, S. Salamitou, P. Bguin, H. Ohayon, P. Gounon, M. Matuschek, and S.F. Koval. Functions of s-layers. *FEMS Microbiol Rev.*, 20(1-2):99–149, June 1997.

[BPS+06] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, 313(5793):1642–1645, 2006.

[Bra86] R. N. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 2nd edition, 1986.

[BRG+09] J. A. G. Briggs, J. D. Riches, B. Glass, V. Bartonova, G. Zanetti, and H.-G. Krusslich. Structure and assembly of immature hiv. *Proceedings of the National Academy of Sciences*, 106(27):11090–11095, 2009.

[BS09] A. Bartesaghi and S. Subramaniam. Membrane protein structure determination using cryo-electron tomography and 3d image averaging. *Journal of Structural Biolology*, 19:1–6, 2009.

[BSL+08] A. Bartesaghi, P. Sprechmann, J. Liu, G. Randall, G. Sapiro, and S. Subramaniam. Classification and 3d averaging with missing wedge correction in biological electron tomography. *Journal of Structural Biolology*, 162:436–450, 2008.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[BVZ01] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[BZ06] S. S. Brandt and U. Ziese. Automatic tem image alignment by trifocal geometry. *Journal of Microscopy*, 222(1):1–14, 2006.

[CD05] Luis R. Comolli and Kenneth H. Downing. Dose tolerance at helium and nitrogen temperatures for whole cell electron tomography. *Journal of Structural Biology*, 152(3):149 – 156, 2005.

[CDE+05] H. Claus, T. Debaerdemaeker, C. Evrard, J. Declercq, J. R. Harris, B. Schlott, and H. Knig. Molecular organization of selected prokaryotic s-layer proteins. *Can. J. Microbiol*, 51:731743, 2005.

[CDT+07] Rossana Cambie, Kenneth H. Downing, Dieter Typke, Robert M. Glaeser, and Jian Jin. Design of a microfabricated, two-electrode phase-contrast element suitable for electron microscopy. *Ultramicroscopy*, 107(4-5):329 – 339, 2007.

[CF02] James M. Coughlan and Sabino J. Ferreira. Finding deformable shapes using loopy belief propagation. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III*, pages 453–468, London, UK, 2002. Springer-Verlag.

[CGS05] G. Cardone, K. Grnewald, and A.C. Steven. A resolution criterion for electron tomography based on cross-validation. *Journal of Structural Biology*, 151:117–129, 2005.

[CHH04] W. R. Crum, T. Hartkens, and D. L. G. Hill. Non-rigid image registration: theory and practice. *Br J Radiol*, 77:140–153, 2004.

[CRH+94] S. Chen, A.M. Roseman, A. Hunter, S.P. Wood, S.G. Burston, Ranson, A.R. N., Clarke, and H.R. Saibil. Location of a folding protein and shape changes in groel-groes complexes imaged by cryo-electron microscopy. *Nature*, 371:261–264, 1994.

[DGK08] John Duchi, Stephen Gould, and Daphne Koller. Projected subgradient methods for learning sparse gaussians. In David A. McAllester and Petri Myllymki, editors, *UAI*, pages 145–152. AUAI Press, 2008.

[DJJ93] D. Donoho, I. Johnstone, and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455, 1993.

[DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.

[DM08] Kenneth H. Downing and Paul E. Mooney. A charge coupled device camera with electron decelerator for intermediate voltage electron microscopy. *Rev Sci Instrum.*, 79, 2008.

[DN01] R. Danev and K. Nagayama. Transmission electron microscopy with Zernike phase plate. *Ultramicroscopy*, 88(4):243–252, 2001.

[Don95] D. Donoho. De-noising by soft-thresholding. *IEEE Transaction on Information Theory*, 41:613–627, May 1995.

[EG79] Murray Eisenberg and Robert Guy. A proof of the hairy ball theorem. *The American Mathematical Monthly*, 86:571–574, 1979.

[EHSH09] C. J. Ehrhardt, R. M. Haymon, S. M. Sievert, and P. A. Holden. An improved method for nanogold in situ hybridization visualized with environmental scanning electron microscopy. *Journal of Microscopy*, 236:5–10, 2009.

[EMK06] G. Elidan, I. McGraw, and D. Koller. Residual belief propagation: Informed scheduling for asynchronous message passing. In *Proceedings of the Twenty-second Conference on Uncertainty in AI (UAI)*, Boston, Massachussetts, July 2006.

[Eng07] Harald Engelhardt. Are s-layers exoskeletons? the basic function of protein surface layers revisited. *Journal of Structural Biology*, 160:115–124, 2007.

[EP98] Harald Engelhardt and Jurgen Peters. Structural research on surface layers: A focus on stability, surface layer homology domains, and surface layercell wall interactions. *Journal of Structural Biology*, 124:276302, 1998.

[FB81] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. Of the ACM*, 24:381395, 1981.

[FHT07] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostat*, page 432, December 2007.

[FJ05] Matteo Frigo and Steven G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[FK03] I. Fodor and C. Kamath. Denoising through wavelet shrinkage: an empirical study. 12:151–160, 2003.

[FLC06] J.J. Fernandez, S. Lia, and R.A. Crowther. Ctf determination and correction in electron cryotomography. *Ultramicroscopy*, 106:587596, 2006.

[Fle08] Patrick Van Fleet. *Discrete Wavelet Transformations: An Elementary Approach with Applications*. Wiley-Interscience, January 2008.

[FMZ+05] F. Forster, O. Medalia, Nathan Z., W. Baumeister, and D. Fass. 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):4729–4734, 2005.

[FNS07] M.J. Ford, J.F. Nomellini, and J. Smit. S-layer anchoring and localization of an s-layer-associated protease in caulobacter crescentus. *Journal of bacteriology*, 189(6):2226, 2007.

[FP88] Janez Funda and Richard P. Paul. A comparison of transforms and quaternions in robotics. *IEEE International Conference on Robotics and Automation*, 2:886 – 891, 1988.

[FPSF08] F. Forster, S. Pruggnaller, A. Seybert, and A. Frangakis. Classification of cryo-electron sub-tomograms using constrained correlation. *Journal of Structural Biolology*, 161:276–286, 2008.

[Fra96] J. Frank. *Three Dimensional Electron Microscopy of Macromolecular Assemblies*. Academic Press Inc., 1996.

[Fra06] J. Frank. *Electron Tomography: Methods for Three-dimensional Visualization of Structures in the Cell. Second Edition.* Springer, 2006.

[GS91] A. Gilchrist and J. Smit. Transformation of freshwater and marine caulobacters by electroporation. *Journal of bacteriology*, 173(2):921, 1991.

[GSR+98] R Grimm, H Singh, R Rachel, D Typke, W Zillig, and W Baumeister. Electron tomography of ice-embedded prokaryotic cells. *Byophysics journal*, 74(2):10311042, February 1998.

[Guc82] R. Guckenberger. Determination of a common origin in the micrographs of tilt series in three-dimensional electron microscopy. *Ultramicroscopy*, 9(1-2):167 – 173, 1982.

[HCWS08] J. Bernard Heymann, Giovanni Cardone, Dennis C. Winkler, and Alasdair C. Steven. Computational resources for cryo-electron tomography in bsoft. *Journal of Structural Biology*, 161(3):232 – 242, 2008. The 4th International Conference on Electron Tomography, The 4th International Conference on Electron Tomography.

[Hen04] Amy Henderson. *The ParaView Guide: A Parallel Visualization Application*. Kitware, 2004.

[Her09] Gabor T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer, 2nd edition, 2009.

[HJB09] Y. He, Grant J. J., and P. J. Bjorkman. Nanogold as a specific marker for electron cryotomography. *Microscopy and Microanalysis*, 15(03):183–188, 2009.

[Hoa63] David Hoag. Apollo guidance and navigation considerations of apollo imu gimbal lock. Technical report, MIT Instrumentation Laboratory, 1963.

[HS88] C.G. Harris and M.J. Stephens. A combined corner and edge detector. In *Proceedings Fourth Alvey Vision Conference*, pages 147–151, Manchester, 1988.

[Hub81] Peter J. Huber. *Robust Statistics*. John Wiley & Sons, New York, NY, 1981.

[HZ04] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

[IMF+05] Kenji Iwasaki, Kaoru Mitsuoka, Yoshinori Fujiyoshi, Yukio Fujisawa, Masakazu Kikuchi, Kiyotoshi Sekiguchi, and Takao Yamada. Electron tomography reveals diverse conformations of integrin alpha-iib-beta3 in the active state. *Journal of Structural Biology*, 150(3):259 – 267, 2005.

[JH99] Andrew E. Johnson and Martial Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, 1999.

[Jor98] Michael Jordan. *Learning in Graphical Models*. The MIT Press, 1st edition, 1998.

[Kar07] Charles F. F. Karney. Quaternions in molecular modeling. *J. Mol. Graph. Mod.*, 25(5):595–604, 2007.

[KF01] Anatoly B. Kolomeisky and Michael E. Fisher. Force-velocity relation for growing microtubules. *Biophysical Journal*, 80(1):149 – 154, 2001.

[KF09] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 1st edition, 2009.

[KK05] Qifa Ke and Takeo Kanade. Robust l1 norm factorization in the presence of outliers and missing data by alternative convex programming. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 592–599, 2005.

[KKT09] M. P. Kumar, V. Kolmogorov, and P. H. S. Torr. An analysis of convex relaxations for map estimation of discrete mrfs. *Journal of Machine Learning Research*, 10:71–106, January 2009.

[KMM96] J. R. Kremer, D. N. Mastronarde, and J. R. McIntosh. Computer visualization of three-dimensional image data using imod. *Journal of structural biology*, 116(1):71–76, January 1996.

[Kui98] Jack B. Kuipers. *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press, 1998.

[KZ04] V. Kolmogorov and R. Zabih. What Energy Functions Can Be Minimized via Graph Cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147, 2004.

[LAS+08] R. D. Leapman, M. A. Aronova, A. A. Sousa, G. Zhang, and M. F. Hohmann-Marriott. Quantitative electron tomography of biological structures using elastic and inelastic scattering. In *EMC 2008 14th European Microscopy Congress 15 September 2008, Aachen, Germany*, 2008.

[Law92] M. C. Lawrence. Least-squares method of alignment. *Chapter8: Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope*, 1992.

[LBB+08] J. Liu, A. Bartesaghi, M.J. Borgnia, G. Sapiro, and S. Subramaniam. Molecular architecture of native hiv-1 gp120 trimers. *Nature*, 455(7209):109–113, 2008.

[LBC99] S. Ludtke, P. Baldwin, and W. Chiu. Eman: Semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology*, 128(1):82–97, 1999.

[LBPE06] Albert Lawrence, James C. Bouwer, Guy Perkins, and Mark H. Ellisman. Transform-based backprojection for volume reconstruction of large format electron microscope tilt series. *Journal of Structural Biology*, 154(2):144 – 167, 2006.

[LCB+05] S. Lanzavecchia, F. Cantele, P.L. Bellon, L. Zampighi, M. Kreman, E. Wright, and G.A. Zampighi. Conical tomography of freeze-fracture replicas: a method for the study of integral membrane proteins inserted in phospholipid bilayers. *Journal of Structural Biology*, 149(1):87 – 98, 2005.

[LFB05] V. Lucic, F. Forster, and W. Baumeister. Structural studies by electron tomography: From cells to molecules. *Annual Review of Biochemistry*, 74(1):833–865, 2005.

[LJ09] Zhuo Li and Grant J Jensen. Electron cryotomography: a new view into microbial ultrastructure. *Current Opinion in Microbiology*, 12(3):333 – 340, 2009. Ecology and Industrial Microbiology - Techniques.

[Low04] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[LWR+06] Jun Liu, Shenping Wu, Mary C. Reedy, Hanspeter Winkler, Carmen Lucaveche, Yifan Cheng, Michael K. Reedy, and Kenneth A. Taylor. Electron tomography of swollen rigor fibers of insect flight muscle reveals a short and variably angled s2 domain. *Journal of Molecular Biology*, 362(4):844 – 860, 2006.

[Mal99] Stéphane Mallat. *A Wavelet Tour of Signal Processing: Wavelet Analysis & Its Applications*. Academic Press, 2nd edition, September 1999.

[Mas97] David N. Mastronarde. Dual-axis tomography: An approach with alignment methods that preserve resolution. *Journal of Structural Biology*, 120(3):343 – 352, 1997.

[Mas05] David N. Mastronarde. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, 152(1):36 – 51, 2005.

[MBS+07] E. Majorovits, B. Barton, K. Schulthei, F. Prez-Willard, D. Gerthsen, and R.R. Schrder. Optimizing phase contrast in transmission electron microscopy with an electrostatic (boersch) phase plate. *Ultramicroscopy*, 107(2-3):213 – 226, 2007.

[MFM04] David R. Martin, Charless C. Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *PAMI*, 26:530–549, 2004.

[MHA+10] F. Moussavi, G. Heitz, F. Amat, L. R. Comolli, D. Koller, and M. Horowitz. 3d segmentation of cell boundaries from whole cell cryogenic electron tomography volumes. *Journal of Structural Biology*, 170:134 – 145, 2010.

[MJW06] Dmitry M. Malioutov, Jason K. Johnson, and Alan S. Willsky. Low-rank variance estimation in large-scale gmrf models. In *ICASSP*, 2006.

[MLJ06] G.E. Murphy, J.R. Leadbetter, , and G.J. Jensen. In-situ structure of the complete treponema primitia flagellar motor. *Nature*, 442:1062–1064, 2006.

[MNM05] Richard McIntosh, Daniela Nicastro, and David Mastronarde. New views of cells in 3d: an introduction to electron tomography. *Trends in Cell Biology*, 15(1):43 – 51, 2005.

[MP87] R. H. J. Gmeligg Meyling and P. R. Pfluger. B-spline approximation of a closed surface. *IMA Journal of Numerical Analysis*, 7:73–96, 1987.

[MS05] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(10):1615–1630, 2005.

[MSKS03] Yi Ma, Stefano Soatto, Jana Kosecka, and S. Shankar Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. SpringerVerlag, 2003.

[MVMH04] Brad J. Marsh, Niels Volkmann, J. Richard McIntosh, and Kathryn E. Howell. Direct continuities between cisternae at different levels of the golgi complex in glucose-stimulated mouse islet beta cells. *PNAS*, 101:5565 – 5570, 2004.

[MWJ99] Kevin P. Murphy, Yair Weiss, and Michael I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In *In Proceedings of Uncertainty in AI*, pages 467–475, 1999.

[NDDS07] J.F. Nomellini, G. Duncan, I.R. Dorocicz, and J. Smit. S-layer-mediated display of the immunoglobulin G-binding domain of streptococcal protein G on the surface of Caulobacter crescentus: development of an immunoactive reagent. *Applied and Environmental Microbiology*, 73(10):3245, 2007.

[NFL+05] Stephan Nickell, Friedrich Frster, Alexandros Linaroudis, William Del Net, Florian Beck, Reiner Hegerl, Wolfgang Baumeister, and Jrgen M. Plitzko.

Tom software toolbox: acquisition and analysis for electron tomography. *Journal of Structural Biology*, 149(3):227 – 234, 2005.

[NKK+07] JE Norville, DF Kelly, TF Jr Knight, AM Belcher, and T. Walz. 7a projection map of the s-layer protein sbpa obtained with trehalose-embedded monolayer crystals. *Journal of Structural Biology*, 160(3):313–323, June 2007.

[NLL+10] John F. Nomellini, Carmen Li, Danielle Lavallee, Iryna Shanina, Lisa A. Cavacini, Marc S. Horwitz, and John Smit. Development of an hiv-1 specific microbicide using ¡italic¿caulobacter crescentus¡/italic¿ s-layer mediated display of cd4 and mip1. *PLoS ONE*, 5(4):e10366, 04 2010.

[NSP+06] Daniela Nicastro, Cindi Schwartz, Jason Pierson, Richard Gaudette, Mary E. Porter, and J. Richard McIntosh. The Molecular Architecture of Axonemes Revealed by Cryoelectron Tomography. *Science*, 313(5789):944–948, 2006.

[OFK+06] Julio O. Ortiz, Friedrich Frster, Julia Krner, Alexandros A. Linaroudis, and Wolfgang Baumeister. Mapping 70s ribosomes in intact cells by cryoelectron tomography and pattern recognition. *Journal of Structural Biology*, 156(2):334 – 341, 2006.

[PEA07] A. Philippsen, H. Engel, and Engel A. The contrast-imaging function for tilted specimens. *Ultramicroscopy*, 107:202–212, 2007.

[PET+08] T. Pavkov, E. M. Egelseer, M. Tesarz, D. I. Svergun, U. B. Sleytr, and W. Keller. The structure and binding behavior of the bacterial cell surface layer protein sbsc. *Structure*, 16(8):1226 – 1237, 2008.

[Phi01] G. Phillips. Green fluorescent protein–a bright idea for the study of bacterial protein localization. *FEMS Microbiol Lett*, 204(1):9–18, 2001.

[PMBF95] P. Penczek, M. Marko, K. Buttle, and J. Frank. Double-tilt electron tomography. *Ultramicroscopy*, 60(3):393–410, 1995.

[PMS91] D. Pum, P. Messner, and U. B. Sleytr. Role of the s layer in morphogenesis and cell division of the archaebacterium methanocorpusculum sinense. *Journal of Bacteriology*, 173(21):6865–6873, 1991.

[PS00] Mary E. Porter and Winfield S. Sale. The 9 + 2 axoneme anchors multiple inner arm dyneins and a network of kinases and phosphatases that control motility. *Journal of Cell Biology*, 151:37–42, 2000.

[PYFS06] Pawel A. Penczek, Chao Yang, Joachim Frank, and Christian M. T. Spahn. Estimation of variance in single-particle reconstruction using the bootstrap technique. *Journal of Structural Biology*, 154(2):168–183, May 2006.

[RBZ06] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3:793–795, 2006.

[RG07] David DeRosier Wah Chiu Joachim Frank Robert Glaeser, Kenneth Downing. *Electron Crystallography of Biological Macromolecules*. Oxford University Press, 2007.

[RK08] L. Reimer and H. Kohl. *Transmission Electron Microscopy: Physics of Image Formation*. Springer Series in Optical Sciences. Springer, 5th edition, 2008.

[SA82] J. Smit and N. Agabian. Cell surface patterning and morphogenesis: Biogenesis of a periodic surface array during caulobacter development. *The Journal of Cell Biology*, 95(1):41–49, 1982.

[SAK+07] A. Sousa, M.A. Aronova, Y. Kim, L. Dorward, G. Zhang, and R.D. Leapman. Limits of detection of ultrasmall gold labels in biological specimens by stem tomography. *Microscopy and Microanalysis*, 13:1328–1329, 2007.

[SD06] Haixin Sui and Kenneth H. Downing. Molecular architecture of axonemal microtubule doublets revealed by cryo-electron tomography. *Nature*, 442:475–478, 2006.

[SEV+92] J. Smit, H. Engelhardt, S. Volker, S. H. Smith, and W. Baumeister. The s-layer of caulobacter crescentus: three-dimensional image reconstruction and structure analysis by electron microscopy. *Journal of Bacteriology*, 174(20):6527 – 6538, 1992.

[SGGA81] J. Smit, D. A. Grano, R. M. Glaeser, and N. Agabian. Periodic surface array in caulobacter crescentus: fine structure and chemical analysis. *Journal of Bacteriology*, 146(3):1135 – 1150, June 1981.

[SH02] Guntram Seltmann and Otto Holst. *The Bacterial Cell Wall*. Springer, 1st edition, 2002.

[SHI+07] U. B. Sleytr, C. Huber, N. Ilk, D. Pum, B. Schuster, and E. M. Egelseer. S-layers as a tool kit for nanobiotechnological applications. *FEMS Microbiology Letters*, 267(2):131–144, 2007.

[Sho85] Ken Shoemake. Animating rotation with quaternion curves. *ACM SIGGRAPH Computer Graphics*, 19:245–254, 1985.

[Sle78] U.B. Sleytr. Regular arrays of macromolecules on bacterial cell walls: structure, chemistry, assembly, and function. *Int. Rev. Cytol.*, 53:1–62, 1978.

[SLR97] Lucy Shapiro, Losick, and Richard. Protein Localization and Cell Fate in Bacteria. *Science*, 276(5313):712–718, 1997.

[SM83] U. B. Sleytr and P. Messner. Crystalline surface layers on bacteria. *Annual Review of Microbiology*, 37:311–339, 1983.

[SM90] SH Smith and RG. Murray. The structure and associations of the double s layer on the cell wall of aquaspirillum sinuosum. *Canadian journal of microbiology*, 36(5):327–335, May 1990.

[SMG+08] David Sontag, Talya Meltzer, Amir Globerson, Yair Weiss, and Tommi Jaakkola. Tightening LP relaxations for MAP using message-passing. In *24th*

*Conference in Uncertainty in Artificial Intelligence*, pages 503–510. AUAI Press, 2008.

[Smi08] John Smit. Heads up s-layer display: The power of many. *Structure*, 16(8):1151 – 1153, 2008.

[SML97] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit: An Object-Oriented Approach to 3-D Graphics*. Prentice Hall, 2nd edition, November 1997.

[SMVC09] Sjors H.W. Scheres, Roberto Melero, Mikel Valle, and Jose-Maria Carazo. Averaging of electron subtomograms and random conical tilt reconstructions through likelihood optimization. *Structure*, 17(12):1563 – 1572, 2009.

[SMVM+04] C.O.S. Sorzano, R. Marabini, J. Velzquez-Muriel, J.R. Bilbao-Castro, S.H.W. Scheres, J.M. Carazo, and A. Pascual-Montano. Xmipp: a new generation of an open-source image processing package for electron microscopy. *Journal of Structural Biology*, 148(2):194 – 204, 2004.

[SOLR06] C.O.S. Sorzano, E. Ortiz, M. Lpez, and J. Rodrigo. Improved bayesian image denoising based on wavelets with applications to electron microscopy. *Pattern Recognition*, 39(6):1205 – 1213, 2006.

[SPK+06] Michael F. Schmid, Angel M. Paredes, Htet A. Khant, Ferda Soyer, Henry C. Aldrich, Wah Chiu, and Jessup M. Shively. Structure of halothiobacillus neapolitanus carboxysomes by cryo-electron tomography. *Journal of Molecular Biology*, 364(3):526 – 535, 2006.

[SRA+09] C.O.S. Sorzano, E. Recarte, M. Alcorlo, J.R. Bilbao-Castro, C. San-Martn, R. Marabini, and J.M. Carazo. Automatic particle selection from electron micrographs using machine learning techniques. *Journal of Structural Biology*, 167(3):252 – 260, 2009.

[SSC+09] Christian Suloway, Jian Shi, Anchi Cheng, James Pulokas, Bridget Carragher, Clinton S. Potter, Shawn Q. Zheng, David A. Agard, and Grant J. Jensen. Fully automated, sequential tilt-series acquisition with leginon. *Journal of Structural Biology*, 167(1):11 – 18, 2009.

[SSMP94] U. B. Sleytr, M. Sra, P. Messner, and D. Pum. Two-dimensional protein crystals (s-layers): Fundamentals and applications. *Journal of Cellular Biochemistry*, 56(2):171–176, 1994.

[STJ+94] A. Sch
"afer, A. Tauch, W. J
"ager, J. Kalinowski, G. Thierbach, and A. P
"uhler. Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. *Gene*, 145(1):69–73, 1994.

[Sub06] Sriram Subramaniam. Bridging the imaging gap: visualizing subcellular architecture with electron tomography. *Curr Opin Microbiol*, 8(3):316–322, November 2006.

[TKR08] Lorenzo Torresani, Vladimir Kolmogorov, and Carsten Rother. Feature correspondence via graph matching: Models and global optimization. *Computer Vision ECCV 2008*, pages 596–609, 2008.

[TM08] Tinne Tuytelaars and Krystian Mikolajczyk. *Local Invariant Feature Detectors: A Survey*. Now Publishers, 2008.

[TNC+02] Mohamed Tagari, Richard Newman, Monica Chagoyen, Jose-Maria Carazo, and Kim Henrick. New electron microscopy database and deposition system. *Trends in Biochemical Sciences*, 27(11):589 – 589, 2002.

[TV98] Emanuele Trucco and Alessandro Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.

[vHS05] Marin van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250 – 262, 2005.

[VV95] Paula Viola and Paul A. Viola. Alignment by maximization of mutual information. In *International Journal of Computer Vision*, pages 16–23, 1995.

[Win07] H. Winkler. 3d reconstruction and processing of volumetric data in cryo-electron tomography. *Journal of Structural Biolology*, 157:126–137, 2007.

[WKRS94] S. G. Walker, D. N. Karunaratne, N. Ravenscroft, and J. Smit. Characterization of mutants of caulobacter crescentus defective in surface attachment of the paracrystalline surface layer. *Journal of Bacteriology*, 176(20):6312 – 6323, October 1994.

[WLR+09] Shenping Wu, Jun Liu, Mary C. Reedy, Hanspeter Winkler, Michael K. Reedy, and Kenneth A. Taylor. Methods for identifying and averaging variable molecular conformations in tomograms of actively contracting insect flight muscle. *Journal of Structural Biology*, 168(3):485 – 502, 2009.

[WSS92] S.G. Walker, S.H. Smith, and J. Smit. Isolation and comparison of the paracrystalline surface layer proteins of freshwater caulobacters. *Journal of bacteriology*, 174(6):1783, 1992.

[XMS+09] Quanren Xiong, Mary K. Morphew, Cindi L. Schwartz, Andreas H. Hoenger, and David N. Mastronarde. Ctf determination and correction for low dose tomographic tilt series. *Journal of Structural Biology*, 168(3):378 – 387, 2009.

[YFW03] Jonathan S. Yedidia, William T. Freeman, and Yair Weiss. Understanding belief propagation and its generalizations. *Exploring artificial intelligence in the new millennium*, pages 239–269, 2003.

[YFW05] Jonathan S. Yedidia, William T. Freeman, and Yair Weiss. Constructing free energy approximations and generalized belief propagation algorithms. *IEEE Transactions on Information Theory*, 51:2282–2313, 2005.

[ZMB+09] Shawn Q. Zheng, Atsushi Matsuda, Michael B. Braunfeld, John W. Sedat, and David A. Agard. Dual-axis target mapping and automated sequential acquisition of dual-axis em tomographic data. *Journal of Structural Biology*, 168(2):323 – 331, 2009.

[ZPSF97] J. Zhu, P. Penczek, R. Schrder, and J. Frank. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70sescherichia coliribosome. *Journal of Structural Biology*, 118(3):197–219, 1997.