# Universidad Autónoma de Madrid

## PhD Thesis

---

# New computational methods toward atomic resolution in single particle cryo-electron microscopy

---

*Supervisor:*

Dr. José Maria Carazo

*Author:*

Vahid Abrishami

*Co-Supervisors:*

Dr. Roberto Marabini Ruiz,

Dr. Carlos Óscar Sánchez

Sorzano

*Submitted in partial fulfillment of the requirements for the degree of*

*Doctor of Philosophy in Investigation and Innovation in Information and*

*Communication Technologies*

*in the*

Computer Science Department,

Escuela Politécnica Superior,

Universidad Autónoma de Madrid,

C/ Francisco Tomás y Valiente, 11,

Madrid 28049 Spain.

April 2016

# Declaration of Authorship

I, Vahid Abrishami, declare that this thesis titled, 'New computational methods toward atomic resolution in single particle cryo-electron microscopy' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

*"Wisdom is not a product of schooling but of the lifelong attempt to acquire it."*

Albert Einstein

# *Abstract*

**New computational methods toward atomic resolution in single particle cryo-electron microscopy**

by Vahid Abrishami

Structural information of macromolecular complexes provides key insights into the way they carry out their biological functions. In turn, Electron microscopy (EM) is an essential tool to study the structure and function of biological macromolecules at a medium-high resolution. In this context, Single-Particle Analysis (SPA), as an EM modality, is able to yield Three-Dimensional (3-D) structural information for large biological complexes at near atomic resolution by combining many thousands of projection images. However, these views suffer from low Signal-to-Noise Ratios (SNRs), since an extremely low total electron dose is used during exposure to reduce radiation damage and preserve the functional structure of macromolecules. In recent years, the emergence of Direct Detection Devices (DDDs) has opened up the possibility of obtaining images with higher SNRs. These detectors provide a set of frames instead of just one micrograph, which makes it possible to study the behavior of frozen hydrated specimens as a function of electron dose and rate. In this way, it has become apparent that biological specimens embedded in a solid matrix of amorphous ice move during imaging, resulting in Beam-Induced Motion (BIM). Therefore, alignment of frames should be added to the classical standard data processing workflow of single-particle reconstruction, which includes: particle selection, particle alignment, particle classification, 3-D reconstruction, and model refinement. In this thesis, we propose new algorithms and improvements for three important steps of this workflow: movie alignment, particles selection, and 3-D reconstruction. For movie alignment, a methodology based on a robust to noise optical flow approach is proposed that can efficiently correct for local movements and provide quantitative analysis of the BIM pattern. We then introduce a method for automatic particle selection in micrographs that uses some new image features to train two classifiers to learn from the user the kind of particles he is interested in. Finally, for 3-D reconstruction, we introduce a gridding-based direct Fourier method that uses a weighting technique to compute a uniform sampled Fourier transform. The algorithms are fully implemented in the open-source Xmipp package (http://xmipp.cnb.csic.es).

# Acknowledgements

Although this thesis has one name on it, its existence is possible only because of the help and support of many people. I would like to thank all those individuals who made this thesis possible:

First of all, I would like to express my deepest sense of gratitude to my advisor and my co-advisors Drs. Jose María Carazó, Roberto Marabini, and Carlos Oscar Sánchez Sorzano for their valuable advice and support. I thank Roberto for always been there to answer my questions with patience and sincerity. I really thank you for always being willing to help and share your knowledge and experiences. I thank José María for giving me this opportunity to work in one of the best laboratories in this field, and providing me whatever I needed to perform my research. I never forget your efforts to solve my problems before and after moving to Spain. I thank Carlos Oscar, with a great personality and one of the smartest person I know, for all his help. He helped me without hesitation even when he was very busy. Without him, it would not be possible for me to finish this thesis. I will miss you all very much!!!

I want to thank Javier for helping me with his brilliant ideas and discussions. He didn't just give me a hand in science, but helped me a lot when I moved to Madrid. I would also like to thank Kino, Jose Miguel and Airen for their friendship, helpful aid and support. They helped me a lot when I arrived in the laboratory. I am also thankful to Jesus, who maintained all the machines in the lab so efficiently that we never had to worry about losing files, creating backups or installing software. I am also indebted to Blanca for her kindness and the help she gave me before and during my Ph.D.

I am also grateful to the former or current staff of the laboratory for their various forms of support during my Ph.D. study–Nacho, Jordi, Josue, Daniel Tabas, Ana Lucia, Laura, Vilas, Mohsen, Joan, Pablo, Johan, Roberto Melero, Monica. You are the best colleagues.

Last but not the least, I would like to thank my wife Maryam with whom I began this long journey full of love thirteen years ago. Her support and encouragement were in the end what made this dissertation possible. I have to thank my parents for giving me the strengths to reach for the stars and chase my dreams.

# Contents

# List of Figures

# Abbreviations

| | |
|---|---|
| **2-D** | **T**wo-**D**imensional |
| **3-D** | **T**hree-**D**imensional |
| **APP** | **A**utomatic **P**article **P**icker |
| **ART** | **A**lgebraic **R**econstruction **T**echnique |
| **BIM** | **B**eam-**I**nduced **M**otion |
| **CCD** | **C**harge **C**oupled **D**evice |
| **CMOS** | **C**omplementary **M**etal **O**xide **S**emiconductor |
| **CTF** | **C**ontrast **T**ransfer **F**unction |
| **DDD** | **D**irect **D**etector **D**evice |
| **DED** | **D**irect **E**lectron **D**etector |
| **DFR** | **D**irect **F**ourier **R**econstruction |
| **ECT** | **E**lectron **C**ryo-**T**omography |
| **EM** | **E**lectron **M**icroscopy |
| **FFT** | **F**ast **F**ourier **T**ransform |
| **FP** | **F**alse **P**ositive |
| **FSC** | **F**ourier **S**hell **C**orrelation |
| **HSV** | **H**ue, **S**aturation and **V**alue |
| **LK** | **L**ucas-**K**anade |
| **MKB** | **M**odified **K**aiser-**B**essel |
| **MPI** | **M**essage **P**assing **I**nterface |
| **MSA** | **M**ultivariate **S**tatistical **A**nalysis |
| **NB** | **N**aive **B**aysian |
| **NMR** | **N**uclear **M**agnetic **R**esonance |
| **NN** | **N**earest **N**eighbor |
| **OF** | **O**ptical **F**low |

| | |
|---|---|
| **PCA** | **P**rinciple **C**omponent **A**nalysis |
| **PDB** | **P**rotein **D**ata **B**ank |
| **PET** | **P**ositron **E**lectron **T**omography |
| **POSIX** | **P**ortable **O**perating **S**ystem Interface |
| **PSD** | **P**ower **S**pectral **D**ensity |
| **RANSAC** | **RAN**dom **SA**mple Consensus |
| **RBF** | **R**adial **B**asis **F**unction |
| **RCT** | **R**andom **C**anonical **T**ilt |
| **SIRT** | **S**imultaneous **I**terative **R**econstruction **T** |
| **SNR** | **S**ignal-to-**N**oise **R**atio |
| **SPA** | **Single** **P**article **A**nalysis |
| **SSNR** | **S**pectral **S**ignal-to-**N**oise **R**atio |
| **SVM** | **S**upport **V**ector **M**achine |
| **TEM** | **T**ransmission **E**lectron **M**icroscope |
| **TP** | **T**rue **P**ositive |
| **PSF** | **P**oint **S**pread **F**unction |
| **WBP** | **W**eighted **B**ack **P**rojection |

*Dedicated to my wife, Maryam*

# Chapter 1

# Introduction

Structural biology is about the determination of the structure of biological macromolecules, desirably down to atomic resolution, in order to understand the way they carry out their biological functions, and to discover how changes in their structures influence their function (Banaszak, 2000). Various techniques of structural biology are employed to yield the Three-Dimensional (3-D) structure of a macromolecule at the atomic level. In all cases, experimental information is combined with stereo chemical and biophysical information to build consistent atomic models. These models can provide vital information about the underlying principles for human diseases, opening new ways to treat them.

Proteins are one of the main topics of structure analysis since they are responsible for cell shape and a large variety of crucial tasks within a cell. Protein is made from a sequence of amino acids which, in turn, results from a genomic sequence. Proteins can bind with other molecules, and this interaction can translate into either subtle or dramatic conformational changes. This possible heterogeneity makes the structure determination of proteins a very challenging task (Henderson *et al.*, 2011).

There are several techniques, including X-ray crystallography, Nuclear Magnetic Resonance (NMR) spectroscopy, and cryo-electron microscopy (cryo-EM) in order to determine the structure of macromolecules. Normally, the size of a macromolecule is in the tens of nanometers or even less, which makes the determination of its structure impossible using a light microscope where the minimal separation between two resolvable objects is around $0.2 - 0.4 \, \mu m$. However, new super-resolution techniques of light microscopy can image beyond the optical diffraction limit but still far away from atomic resolution (Moerner, 2015).
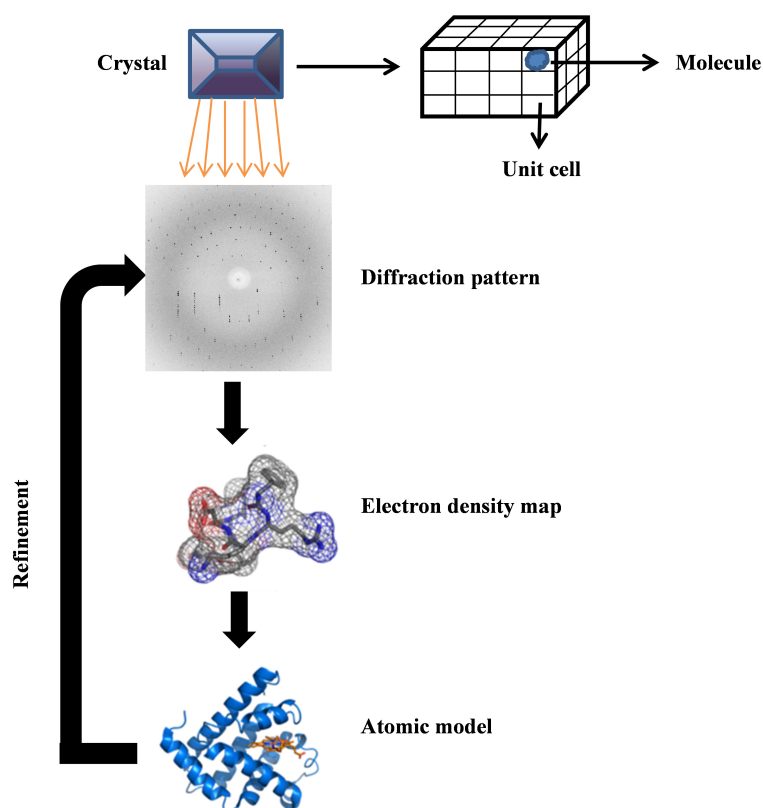
FIGURE 1.1: The workflow of X-ray crystallography for the determination of the structure of a molecule.

## 1.1 Structure determination techniques

There are several well-established methods to determine the structures of biological structures in different ranges from millimeter to atomic resolution, each with their advantages and disadvantages. Practically, a single method is not able to reveal all the structural information of macromolecules, so these methods should work in a complementary manner to understand how macromolecules interact dynamically with their functional environment.

### 1.1.1 X-ray crystallography

X-ray crystallography is a powerful tool with well-established methods to determine the arrangement of atoms in a large macromolecule. Over 80% of all 3-D atomic models in the Protein Data Bank (PDB) (Rose *et al.*, 2015) have been obtained using X-ray crystallography. For this technique, first a crystal containing the molecules that are ordered in a specified spatial arrangement must be grown, and then this crystal is bombarded with an intense beam of X-ray. The electrons of atoms diffract X-rays and produce
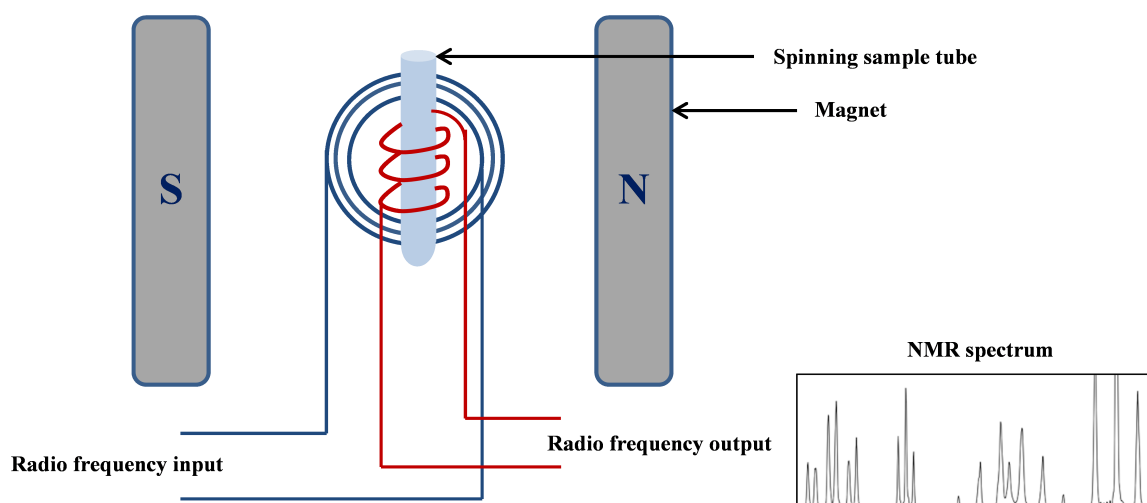
FIGURE 1.2: The basic arrangement of an NMR spectrometer.

images that consist of thousands of points (diffraction spots) in reciprocal space (Bragg, 1913). These images have information about the content of the crystal's unit cell. For each point, the amplitude can be simply computed from the intensity of the spot, but computing the phase is tricky (this is the well known "phase problem"). Figure 1.1 shows how X-ray crystallography determines the structure of a molecule. For more information about X-ray crystallography, the reader is referred to Ladd and Palmer (1985).

Although X-ray crystallography is a powerful technique to determine the atomic structure of a molecule, there are some limitations that make its usage hard or impossible in some conditions (Acharya and Lloyd, 2005). First, the quality of the determined structure tightly depends on the quality of the crystal, so that a crystal of bad quality can reflect the structure of the molecule poorly. Second, some complexes do not form nice and ordered crystals since they are either flexible or undergo biochemical or conformational changes that contradict the assumption of having many molecules aligned in the same orientation within a crystal. These flexible parts of the molecules cannot be captured using X-ray crystallography.

### 1.1.2   NMR spectroscopy

NMR spectroscopy is another technique of structural biology to recover the organization of atoms in small macromolecules. In this method, first a purified protein is placed in an external magnetic field, and then a radio frequency pulse is applied. The magnetic fields around atoms of the protein change their resonance frequencies, thus producing resonance lines which can be analyzed to build a model of the protein that shows the location of each atom (Keeler, 2011). Figure 1.2 shows a simple schema of NMR spectroscopy.

NMR spectroscopy has the advantage of allowing to study the dynamics of macromolecular systems. However, this method is expensive to use (Chatham and Blackband, 2001). Additionally, for big macromolecules, it is often hard to analyze NMR spectra.

### 1.1.3 Cryo-EM

X-ray crystallography is limited to complexes that can form a crystal and NMR spectroscopy is limited to small macromolecules. These limitations prevent from determining the structure of many macromolecules. Cryo-EM does not have the constraints of these two methods, and with the new advances in electron microscopy it can provide near-atomic resolution of biological structures (Bai *et al.*, 2015, Carazo *et al.*, 2015, Cheng, 2015, Doerr, 2016, Eisenstein, 2016, Glaeser, 2016, Nogales, 2016). In this method, the electron beam generated by a Transmission Electron Microscope (TEM) passes through the frozen specimen. The scattered electrons are focused by the electromagnetic lenses of the TEM to form an image. Freezing the sample in cryo-EM protects the sample from radiation damage, that is high enough to change the shape of the specimen. Negative staining is another method to mitigate electron-induced sample damage where the surface of the sample is coated with reagents containing heavy atoms. However, in the latter case, the internal information of the macromolecule is lost, and artifacts such as sample flattening are possible. Indeed, negative staining is better suited for an early phase study, to verify the initial quality of the sample and have a general idea about the macromolecule shape.

Single-Particle Analysis (SPA) and Electron Cryo-Tomography (ECT) are two subdisciplines of cryo-EM to determine the 3-D structure of macromolecules.

In SPA, multiple copies of the same macromolecule, ideally with random orientations, are simultaneously imaged into a single micrograph (see Figure 1.3). Due to the low electron dose used for the observation, the particle images suffer from a low signal-to-noise ratio (SNR). Thus, similar particles are aligned and averaged to obtain particle images with higher SNR. Thus, particles from several hundreds or even thousands of micrographs are commonly employed in a structural study. Before retrieving the original 3-D map from its projections, we need to estimate the orientation of the object in each image. Euler angles can describe such an orientation in 3-D Euclidean space. Once the Euler angles are assigned, average images are back projected using a reconstruction algorithm (Penczek, 2010) to build the final 3-D model of the specimen. The standard data processing workflow of single-particle reconstruction includes: particle selection, particle alignment, particle classification, 3-D reconstruction, and model refinement (Bai *et al.*, 2015, Cheng, 2015, Frank, 1996).
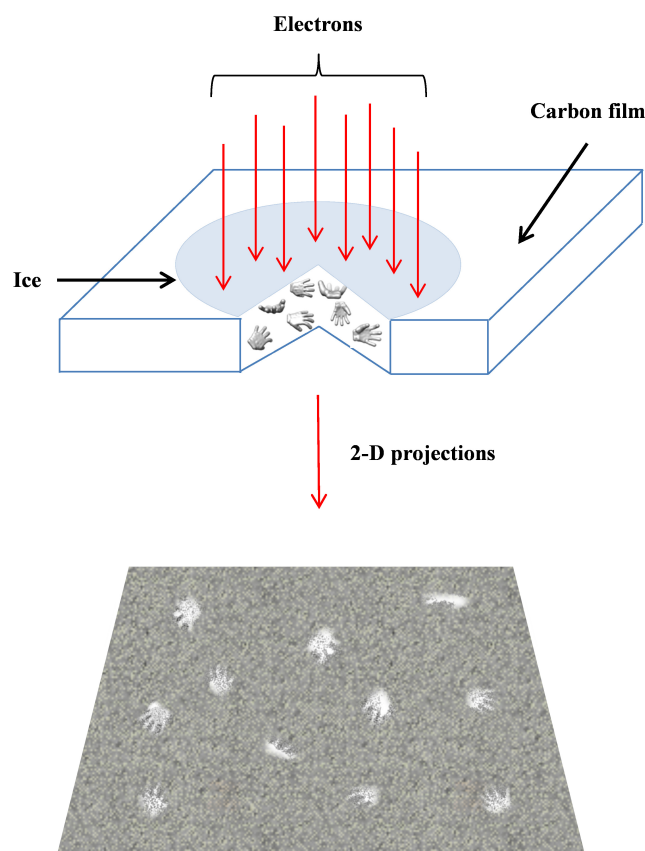
FIGURE 1.3: 2-D projections of particles of the same 3-D object with random orientations in SPA (adapted from Schmeisser (2009)).

ECT is a method that allows the study of cells and large macromolecular machines. In contrast to SPA, this approach does not require purification of macromolecular complexes, thus, it can provide detailed 3-D information on the structure of macromolecular complexes in *situ*. Additionally, there is no need to use several copies of the same macromolecule, assumed to be conformationally homogeneous and placed in random orientations (Schmid, 2011). In this approach, a series of images (tilt series) is acquired by tilting the sample around a fixed axis (see Figure 1.4A). Physical restrictions limit the tilt range to $\pm 60 - 70°$. Tilt series images are then computationally combined by a 3-D reconstruction process to form a 3-D map (see Figure 1.4C), commonly referred to as tomogram. In subsequent analysis steps, similar sub-volumes (sub-tomograms) within the tomogram can be identified and selected, so that they can be aligned to a reference and then averaged, providing a more detailed structural information than is present in the original tomogram (Briggs (2013)).
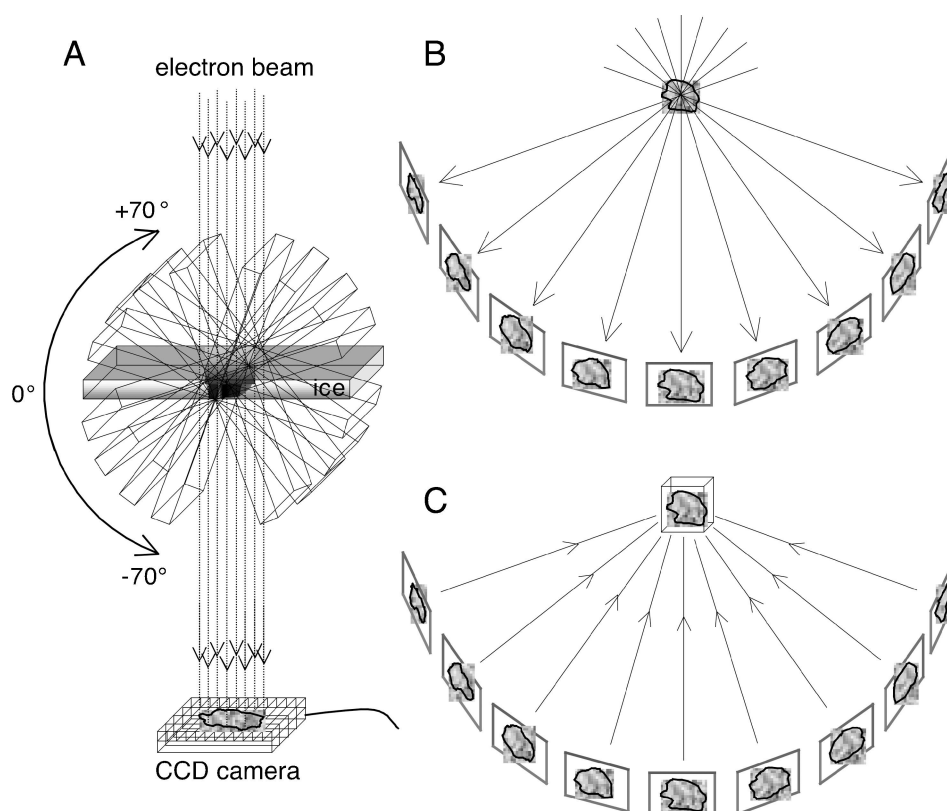
FIGURE 1.4: Principle of electron tomography. A series of images (tilt series) is acquired by tilting the sample around a fixed axis (A,B), and then computationally combined by a 3-D reconstruction process to form a 3-D map (C) (from Steven and Belnap (2001)).

## 1.2 Fundamentals of 3-D electron microscopy

### 1.2.1 Transmission Electron Microscope (TEM)

TEM transmits a beam of electrons through a very thin object. The scattered electrons by matter are then focused using electromagnetic lenses onto an imaging device such as a fluorescent screen, photographic film, or a charge-coupled device (CCD) camera. Since the wavelength of high-energy electrons ($2 - 5pm$) is much shorter than the wavelength of the visible light, TEM gives a superior resolving power that is enough for atomic resolution imaging of biological objects. However, in practice, the expected atomic resolution is not commonly obtained, particularly in Life Science applications, mainly due to lens aberrations associated with the microscope, the very small electron doses used in the imaging, Beam-Induced Motion (BIM), and the sample heterogeneity (Frank, 1996). Figure 1.5 shows different components of a TEM. In a classical microscope, the electron gun generates the electron beam (other sources are used for high resolution studies). For this, the cathode filament emits the electrons which are made converge toward the optical axis using a Wehnelt cylinder. The anode accelerates the emitted
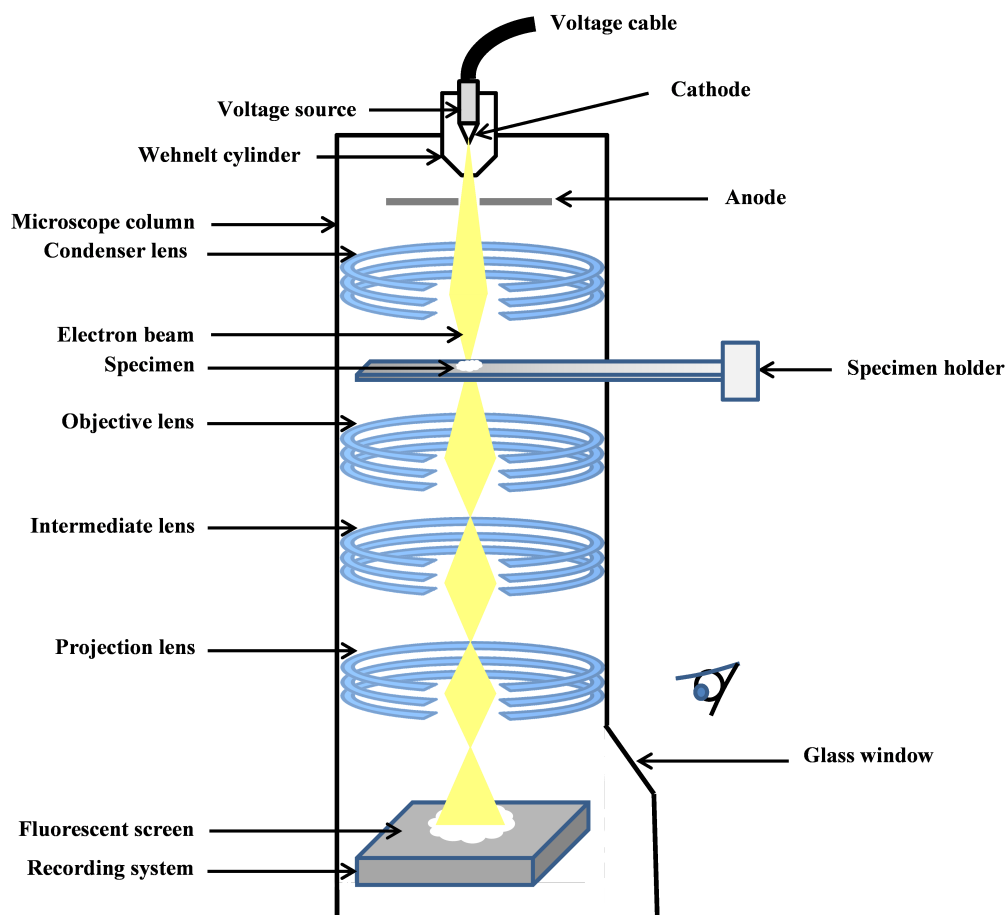
FIGURE 1.5: Principle components of a TEM.

electrons, based on the selected voltage, as a fine beam. Condenser lenses form an image of the source on the sample by adjusting the coherency and the physical size of the beam. After electrons pass through the sample, the objective lens focuses them and initially magnifies the image. Finally, projection lenses project the magnified image on the fluorescent screen.

TEM images are approximately projections (X-ray transform) of the electron density in the sample. However, microscopy images are not exactly the projections of the object under study, and the image is affected by the distortion added by the device. In fact, the final image is the convolution result of the original projection and a mathematical function that depends on geometrical and physical properties of the microscope. This mathematical function is called Point Spread Function (PSF) which affects each point of the original projection. In Fourier space, this convolution change to multiplication, and PSF is called Contrast Transfer Function (CTF). A typical contrast transfer function observed from an electron micrograph, and its 1-D profile is shown in Figure 1.6. As seen in this figure, CTF is oscillatory. When it is negative, positive phase contrast occurs, meaning that atoms will appear dark on a bright background; while when it
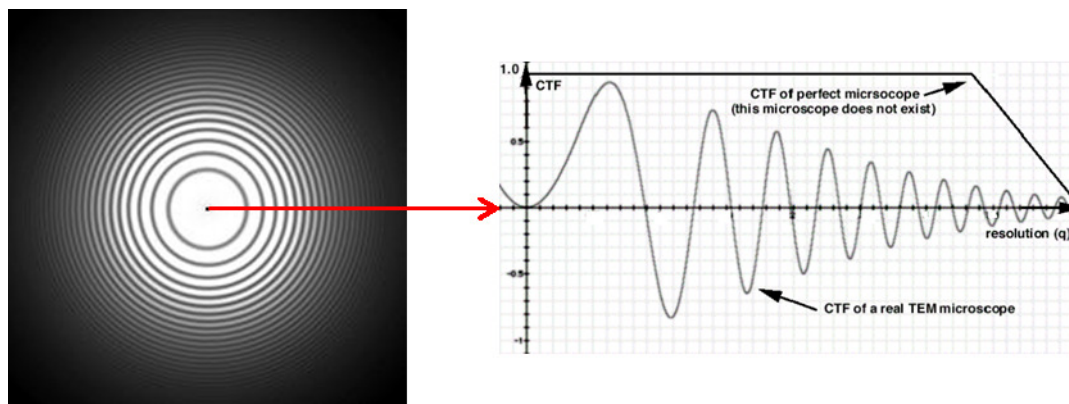
FIGURE 1.6: The CTF of a micrograph and its 1-D profile. CTF is oscillatory: there are "passbands" where it is not equal to zero (good "transmittance") and there are "gaps" where it is equal (or very close to) zero (no "transmittance").

is positive, negative phase contrast occurs, meaning that atoms will appear bright on a dark background. Zero CTF means there is no information transfer for this spatial frequency. To have an accurate final reconstruction of the real macromolecule, we should estimate CTF properly and correct for it.

In theory, an electron microscope should provide resolutions around 0.037 Å when it uses a 100 kV acceleration voltage (expected from the Broglie wavelength) (Frank, 1996). However, in the practice for macromolecular machines, the best electron microscopes can reach a resolution in the range $2 - 3$ Å because of some limiting factors such as noise introduced by the device, sample preparation, CTF, and radiation damage to the sample. Researchers have proposed new solutions to minimize the effect of these factors and recover the maximum resolving power of TEMs. The emergence of new direct detectors, new image processing techniques, and new preservation techniques are some efforts aiming to this goal Nogales (2016).

### 1.2.2  Direct Detection Device (DDD)

Traditional TEM digital cameras use indirect electron detection to form an image. In these cameras, a scintillator is coupled to a sensor (CCD or CMOS) through a lens or fiber optic coupling. The scintillator converts incoming electrons to photons to be detectable by the sensor (Faruqi and Subramaniam, 2000, Mooney, 2007). In contrast to indirect electron detectors, DDDs can detect the image-forming electrons directly. These cameras typically use Complementary Metal Oxide Semiconductor (CMOS) pixel sensors instead of CCD's. DDDs use two distinct modes to translate the striking electrons into pixels: integration mode and counting mode. In integration mode (Figure 1.7c), the generated charge by striking electrons is accumulated by the sensor over a fixed internal frame rate and then read out to produce the final image. By contrast, in counting mode
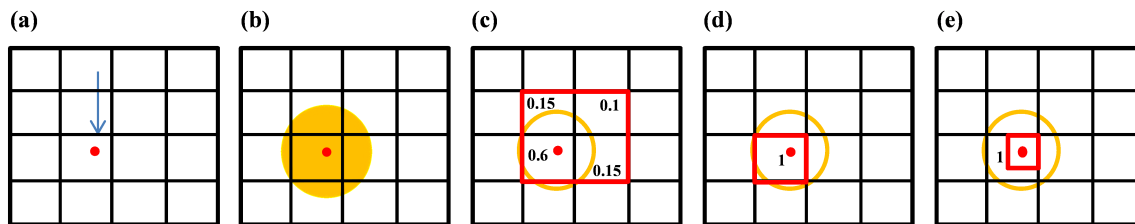
FIGURE 1.7: Methods to convert the electron interacting with a DDD into pixels. Striking electron (a), scattered signal (b), charge integration mode (c), counting mode (d) and super-resolution mode (e) (adapted from Booth and Mooney (2013)).

(Figure 1.7d), each electron event is detected, thus making it possible to do electron counting for each pixel (sub-pixel). The use of this mode is not feasible when the dose rate is high where electron events are too close to each other (Booth and Mooney, 2013). The Gatan K2 Summit detector has a super-resolution mode where sensor pixel size is slightly smaller than the area that the electron interacts with; as a result, each incoming electron deposits signal in a small cluster of pixels (Figure 1.7e).

DDDs have several advantages over traditional sensors (Deptuch *et al.*, 2007, Faruqi *et al.*, 2003, Milazzo *et al.*, 2005, Xuong *et al.*, 2004). First, the thin layer sensor allows electrons to pass through with a much-reduced backscattering, which results in a smaller PSF, and, therefore, higher resolution. Moreover, image artefacts related to scintillators and fiber optics or lenses do not limit the achievable resolution. Another advantage of DDDs is its high frame rate which can be exploited for motion correction, damage compensation, and other image processing techniques (Wu *et al.*, 2015).

Currently, there are three DDDs brands available in the market: the Gatan K2 Summit, which has the highest detective quantum efficiency (DQE) at low spatial frequency, the FEI Falcon II (Kuijper *et al.*, 2015) having the highest DQE beyond one-half the Nyquist frequency, and finally the Direct Electron DE-12, which has the fastest data acquisition rate (McMullan *et al.*, 2014, Ruskin *et al.*, 2013).

### 1.2.3   Sample preparation for TEM

Special sample preparation techniques are necessary for the observation of hydrated specimens using TEM. The sample should be first placed in the electron microscope column that is under high vacuum to enable the electron beam to travel in straight lines. Since biological samples contain significant amounts of water, they would evaporate in vacuum. Additionally, the radiation of TEM is another potential source of specimen damage. Embedding in negative stain and embedding in vitreous ice are two well-known sample preparation methods to overcome the problems mentioned above, at least, to some extent (Frank, 1996, Ruiz and Radermacher, 2006).
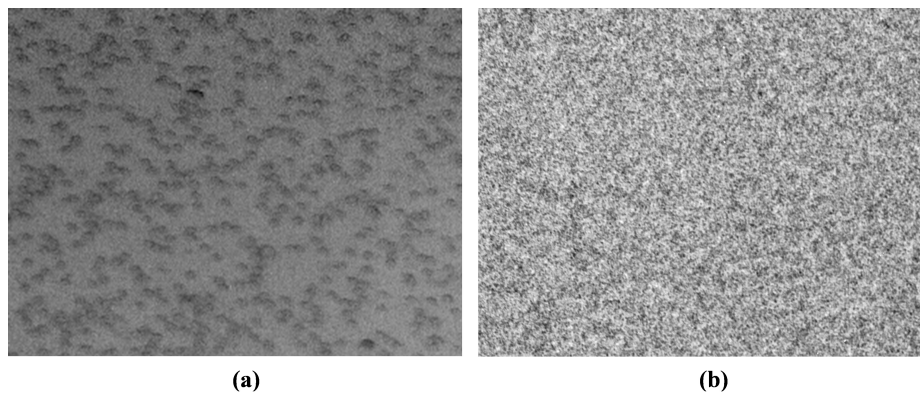
FIGURE 1.8: Sample preparation techniques in single-particle cryo-EM. Negative stained (a) and frozen hydrated (b) biological samples.

For negative staining, the sample is covered with heavy metal salts such as uranyl acetate. Negative stain protects the sample from the evaporation and provides high contrast particle images (Figure 1.8a). However, negative stain EM just gives a cast of the exterior surface of the molecule, and cannot resolve interior information from the specimen. Additionally, negative staining may flatten the samples, which may result in considerable structural distortions.

In contrast, embedding the samples in vitreous ice allows the imaging of the sample in its hydrated state without interfering with its shape. Indeed, this method protects the sample much better than negative stain, including reducing radiation damage due to the low-temperature environment where the sample is kept and provides vital information about the macromolecule itself. However, cryo-EM generates projection images of low contrast since the scattering density of ice is just a little smaller than the scattering density of proteins (Figure 1.8b). Large numbers of particles are required to compensate for the low SNR and obtain high-resolution maps.

## 1.3 Single-particle cryo-EM

Single-particle cryo-EM provides an accurate estimation of the 3-D structure of a macro-molecule by computationally combining images of many individual macromolecules in identical or similar conformations. Using an electron beam, these individual macro-molecules at random orientations in a layer of vitreous ice generate some Two-Dimensional (2-D) projection images (particles) in a micrograph (image). Since macromolecules are sensitive to radiation, low-dose radiation exposures are commonly employed. As a consequence, the particle images are extremely noisy. Averaging techniques can combine thousands to hundreds of thousands of similar particle images to calculate a high SNR
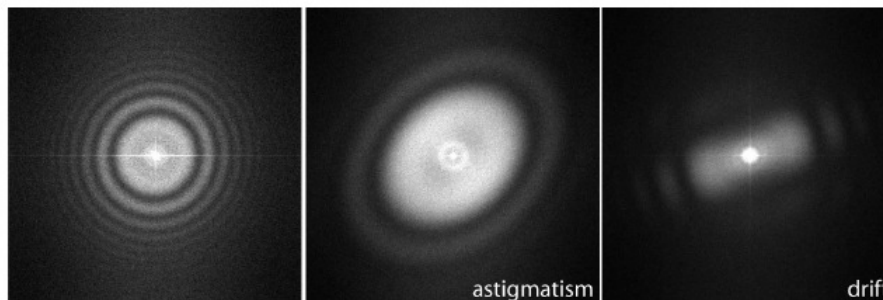
FIGURE 1.9: CTFs of good (left), astigmatic (center) and drift (right) micrographs, respectively (from Scheres *et al.* (2008)).

structure from noisy images. The standard data processing of single-particle reconstruction includes movie alignment (for DDD movies), estimation of CTF parameters, particle selection, 2-D particle alignment and classification, initial model calculation of initial model, structure refinement and 3-D classification. We will go into details of each step in the rest of this section.

## 1.3.1 Movie alignment

In contrast to the classical workflow for SPA, the first input of the workflow may not be a set of micrographs but a set of movies. An easy way to process this input is to compute the average of each movie and reduce it to classical micrograph images. However, and as we mentioned before, aligning the individual frames of a movie can substantially reduce the image blurring caused by BIM. This alignment can be at the micrograph level (Abrishami *et al.*, 2015a, Grant and Grigorieff, 2015, Li *et al.*, 2013, Wu *et al.*, 2015) or at the particle level (Bai *et al.*, 2013, Rubinstein and Brubaker, 2015). At the micrograph level, the algorithm aligns the frames of a movie and generates a corrected average that is easy to integrate with the classical workflow. At the particle level, however, a change in the traditional workflow is necessary, since the algorithm tries to align the particle instances in different frames. In Section 2.1 of this thesis, we go into more details of methods for movie alignment.

## 1.3.2 Estimation of CTF and screening of micrographs

The estimation of the CTF parameters is a classical research topic with an abundant bibliography (Rohou and Grigorieff, 2015, Sorzano *et al.*, 2007, Vargas *et al.*, 2013a, Zhang, 2016). In order to achieve an accurate structural determination, it is important first to screen the micrographs, then to estimate the CTF, and then correct for it. To estimate CTF, we need some parameters like acceleration voltage, spherical aberration,
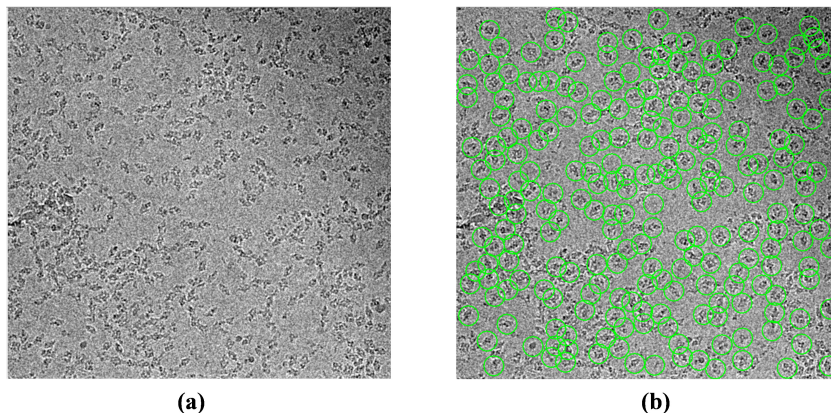
FIGURE 1.10: An electron micrograph of $\beta$-galactosidase (Bartesaghi *et al.*, 2014) before (a) and after picking (b) using Xmipp automatic picking (Abrishami *et al.*, 2013).

and amplitude contrast to describe the frequency region to be analyzed. We can estimate defocus and astigmatism values by fitting CTF pattern to the 2-D power spectrum of the micrograph. The CTF pattern of a good micrograph consists of several concentric rings, extending from the image center towards its edges (see Figure 1.9 left). However, for micrographs of low quality, this pattern consists of asymmetric rings (astigmatism) or rings that fade in a particular direction (drift) (see Figure 1.9). Micrographs with CTF patterns of the latter two types should not be further processed and should be removed from the dataset.

### 1.3.3 Particle selection

Particle selection from electron micrographs is the first step of the single-particle analysis workflow. As it is discussed before, tens or even hundreds of thousands of particles are required to achieve high-resolution maps. Manually identifying this number of particles from many low-dose micrographs is a tedious task. Since this step is the first and one of the crucial measures towards a high-resolution reconstruction, researchers have proposed many automatic and semi-automatic approaches to facilitate this task. These algorithms are indeed targeted to select particles, but a number of false positive is unavoidable (Figure 1.10). Therefore, the output particles should be additionally processed using particle quality assessment algorithms (Vargas *et al.*, 2013b) and classification to generate a dataset with high-quality particles. To further improve the quality of particles, one can filter the images to remove the noise and correct for the CTF phase.

In Section 2.2 of this thesis, I go into the details of a number of different methods for the automatic selection of particles, describing my proposed approach (Abrishami *et al.*, 2013) for this task.
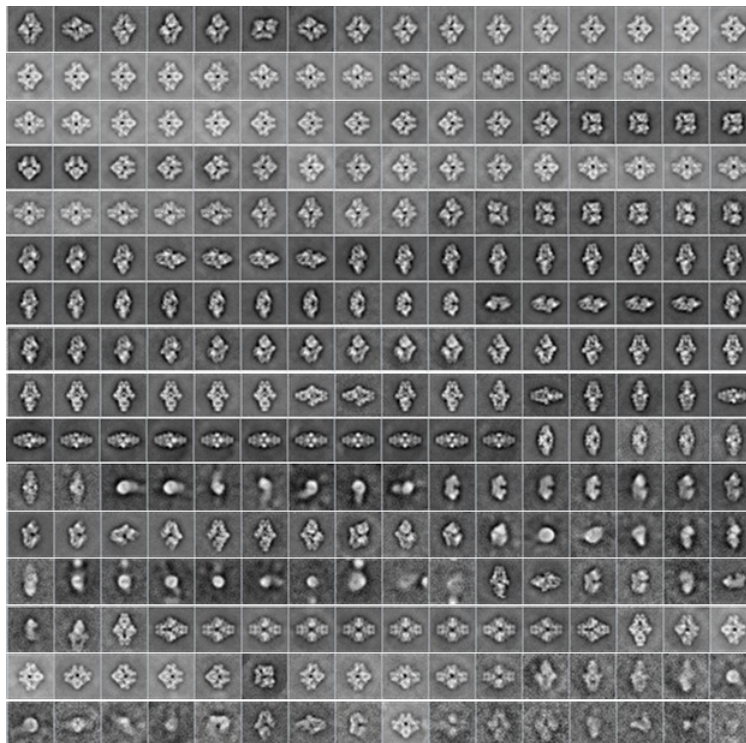
FIGURE 1.11: 256 class avearges after classification of 90,503 $\beta$-galactosidase particles (Bartesaghi *et al.*, 2014) using the proposed method by Sorzano *et al.* (2010), sorted by number of class members.

### 1.3.4 Particle alignment and classification

Extracted particles from electron micrographs are very noisy and many times are not directly used in the 3-D reconstruction workflow. However, one can increase the SNR by averaging multiple particles (Frank, 1975, van Heel *et al.*, 2000). This averaging is only possible if the similar macromolecules have the same orientations inside the vitreous ice, except for rotation around an axis perpendicular to the plane of the detector. Then, by knowing the rotations and translations of each particle, we can align them and calculate an average with higher SNR. However, the existence of projections from various directions, projections from different conformations, and even projections from various macromolecules make this dataset heterogeneous. Therefore, we should divide the dataset into several homogeneous classes, and do the alignment inside each class. However, classification requires that these classes are aligned first, unless they are classified using shift invariants, rotational invariants, or both (Chong *et al.*, 2003, Flusser *et al.*, 1999, Lowe, 1999, 2004, Pun, 2003, Schmid and Mohr, 1997, Tuytelaars and Van Gool, 1999). However, in high noise levels, these invariants do not have enough information for classification. Multivariate statistical analysis (MSA) and classification is one solution to this problem. MSA can show each image of a dataset as a linear combination of its main eigenvectors, and thus, facilitates the classification of particles into groups
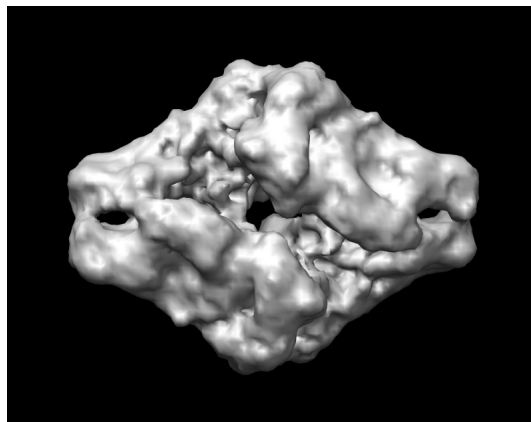
FIGURE 1.12: One of the initial volumes for $\beta$-galactosidase (Bartesaghi *et al.*, 2014)
generated by RANSAC method (Vargas *et al.*, 2014).

with similar orientations in an iterative manner (Frank, 1990, Frank *et al.*, 1981, van Heel, 1984, 1986, van Heel and Frank, 1981). Multi-reference classification alignment is another solution to this problem which carries out 2-D alignment and classification iteratively until convergence (Scheres *et al.*, 2005, Sorzano *et al.*, 2010, van Heel and Stoffler-Meilicke, 1985). In this method, class representatives are considered as references for alignment. First, by assuming we have some initial references, the algorithm aligns particles to each reference and calculates a similarity measure. Then, each particle is assigned to the class with the highest similarity. Now we can calculate the new class representatives by averaging over the members of each class. This process can be repeated until stable class averages are obtained, i.e. no improvement with respect to the previous iteration can be observed. Figure 1.11 shows the class averages of 256 classes (sorted by the number of members) after six iterations. As seen in this figure, some of the classes, normally with few particles, mostly contain noise and should be excluded from the dataset.

### 1.3.5 Initial 3-D model generation

Refinement methods require an approximate model of the final 3-D map to assign each particle to a direction initially. A model of a similar macromolecule, already available from other studies is a good initial estimate, after being low-pass filtered (to prevent from reference-model induced bias). However, one such a map is not always available. Therefore, we need a method to determine this first volume automatically. Researchers have proposed several approaches to this challenging issue. In random canonical tilt (RCT) (Leschziner and Nogales, 2006, Radermacher *et al.*, 1987, Sorzano *et al.*, 2015b), a set of tilt pairs is obtained by recording one untilted and one tilted orientation (e.g. 0 and $\pm 45$). Alignment and classification of untilted and tilted images provide a constraint
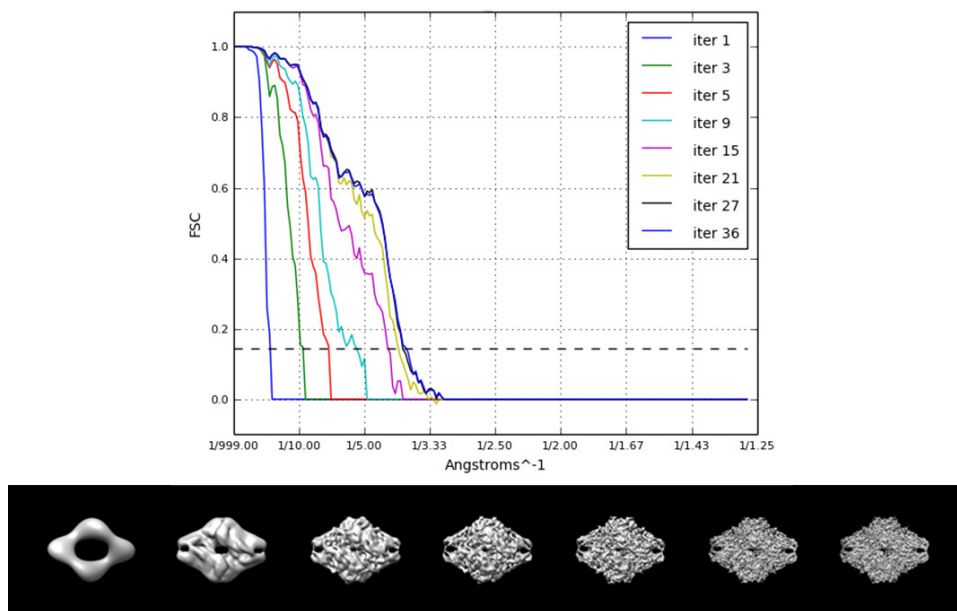
FIGURE 1.13: Evolution of the reconstruction of $\beta$-galactosidase (Bartesaghi *et al.*, 2014) along iterations (iteration 1, 3, 5, 9, 15, 21, 27 and 36) using the proposed method by Scheres (2012b).

on the Euler angles which can be exploited to make an angular assignment without the need of a reference volume. Knowing the Euler angles of each projection, we can generate a 3-D model. There are also many methods based on common lines (Castón *et al.*, 1999, Crowther *et al.*, 1970, Elmlund and Elmlund, 2012, Elmlund *et al.*, 2010, 2008, Liu *et al.*, 2007, Ogura and Sato, 2006, Penczek *et al.*, 1996, Thuman-Commike and Chiu, 1997, van Heel, 1987) to generate an *ab initio* reconstruction without tilting the specimen. In Xmipp (de la Rosa-Trevín *et al.*, 2013), we have a range of different approaches to this task, including an RANSAC-based approach to generate this initial 3-D map (Vargas *et al.*, 2014) and "Significant" (Sorzano *et al.*, 2015a). In the former, some volumes are created by assigning random orientations to different subsets of class averages. Then, for each volume, we calculate a score by comparing it with the rest of the class averages. The volumes with the highest scores are refined using projection matching, presenting to the user several possible solutions. Figure 1.12 shows an initial volume generated using this method. The latter addresses the initial volume problem in SPA by setting it in a Weighted Least Squares framework and calculating the weights through a statistical approach based on the cumulative density function of different image similarity measures.

In Section 2.3 of this thesis, we go over the details of different methods for 3-D reconstruction (Abrishami *et al.*, 2015b) after the determination of projection directions.

### 1.3.6 Refinement

In this step, the low-resolution volume previously obtained is iteratively refined to achieve a 3-D map with high resolution. In each iteration, the current volume is projected in different directions (according to a defined angular sampling) to generate a set of reference images for 3-D alignment and classification. Each projection image is then compared to reprojections of this volume (for example, by cross-correlation), selecting the orientation parameters of the reprojection image with the highest correlation. The new direction parameters of each single-particle images are now used to generate a new 3-D model that normally has a better resolution than the one from the previous step. The algorithm stops when there is no further improvement in resolution. Typically the so called Fourier Shell Correlation (FSC) (Harauz and van Heel, 1986, van Heel and Schatz, 2005) curve is used to monitor the resolution during the refinement process. The FSC curve is obtained by computing correlation coefficients along shells of two volumes, each calculated from one-half of the dataset. Most of the algorithms for refinement (such as Scheres (2012b), Sorzano *et al.* (2004, 2015a)) use the same basic comparison principle as projection matching (Penczek *et al.*, 1994), but they are different in the way that they re-estimate the orientation parameters and the degree to which the user can control the process.

### 1.3.7 3-D classification

Living cells typically require various binding interactions to perform their task. These interactions may be accompanied by conformational changes. In SPA, macromolecular complexes exist in multiple conformations in solution, and it imposes no restrictions on the range of existing conformations (Leschziner and Nogales, 2007). Therefore, SPA allows the study of macromolecular assemblies in distinct functional states. However, SPA assumes that all the projections come from identical three-dimensional objects to determine the structure of a biological object. So, it is important to identify different states present in the experimental data and quantify the percentage of each state within the biological sample.

Several methods have been proposed to detect heterogeneity within cryo-EM data and classify heterogeneous projection data (Spahn and Penczek, 2009). Methods for 3-D classification can be divided into two categories, supervised and unsupervised methods (Frank, 2006). Supervised classification methods use guessed or known structures of conformations as the reference to separate particles according to the best correlation with re-projections of each 3D structure (e.g. Gao *et al.* (2004), Heymann *et al.* (2004), Valle *et al.* (2002)). The shortcomings of these methods are the danger of reference bias and
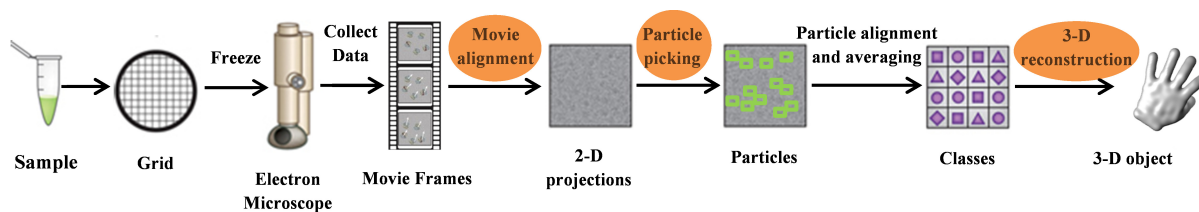
FIGURE 1.14: Different steps of SPA for resolving the structure of biological objects (adapted from Doerr (2016)).

difficulty to discover the unexpected behavior of the specimen. In contrast, unsupervised classification methods classify particles according to their mutual relationships without any such guidance (e.g. Elmlund *et al.* (2013), Fu *et al.* (2007), Herman and Kalinowski (2008), Liao *et al.* (2015), Scheres *et al.* (2007), Shatsky *et al.* (2010), Tang *et al.* (2007), Wang *et al.* (2013)).

## 1.4 Thesis challenges

Recent technical advances in single-particle cryo-EM allow researchers to determine the structure of macromolecules near-atomic-resolution. The emergence of new detectors that can provide images with higher contrast together with new computational tools are the keys that make SPA today as an alternative for X-ray crystallography (Doerr, 2016, Eisenstein, 2016, Nogales, 2016). The introduction of this method in 2015 as the "Method of the Year" by the *Nature Methods* conveys the significant role of SPA in this field.

The high-contrast images provided by direct detectors and sophisticated image-processing software tools together are necessary to perform a successful cryo-EM study. Today many research groups provide software packages with required image-analysis methodologies for SPA. Some of the packages that are broadly used by EM community are: EMAN (Ludtke *et al.*, 1999), Spider (Frank *et al.*, 1996), SPARX (Hohn *et al.*, 2007), Imagic (van Heel *et al.*, 1996), Frealign (Grigorieff, 2007), Xmipp (de la Rosa-Trevín *et al.*, 2013), Relion (Scheres, 2012a).

The aim of the research in this thesis is to improve three significant computational steps of SPA inside Xmipp package to enhance the throughput and to obtain a higher resolution of biological specimens. Figure 1.14 shows the SPA workflow with specified steps that this thesis will focus on. As is seen in this figure, movie alignment, particle picking and 3-D reconstruction are the main challenges of this thesis. Integration of these new algorithms in Xmipp package should lead to better resolutions.

### 1.4.1 Movie alignment

For several years, BIM was a barrier for extracting high-resolution information of biological objects using SPA. Nowadays it is possible to correct for BIM using the frames provided by DDDs. Several methods have been proposed by different researchers for BIM correction, and they showed that how this correction can improve the final resolution. However, still it is a challenging problem since there is no general agreement about the BIM pattern. Some researchers believe that displacements between frames are global while others think the movements in between frames is local. A better understanding of BIM pattern to propose a method for its correction is challenging.

### 1.4.2 Particle picking

Picking particles from electron micrographs is the first and one of the important steps of SPA that can affect all the steps significantly. A wrong picking causes resolution loss or even a bad 3-D structure due to biasing (Henderson, 2013). Picking hundred thousands or millions of particles from electron micrographs is a tedious task and constitutes a hurdle towards high-throughput. Several methods have been proposed for automatic selection of particles from electron micrographs by researchers. However, there is no balance between speed and accuracy, and methods with high accuracy usually are slow. Therefore, development of an automatic particle picking method which can detect particles fast and accurate is a challenging problem.

### 1.4.3 3-D reconstruction

After the determination of Euler angles for particles, 3-D reconstruction algorithms re-project back each particle to generate the final density map. In refinement step, we require calling a 3-D reconstruction algorithm in each iteration to refine our reconstruction. Consequently, A slow algorithm for reconstruction works as a bottleneck for refinement algorithms. A refinement algorithm can take long if the reconstruction algorithm works slowly. A fast reconstruction algorithm that can re-project back particles accurately is another challenge.

## 1.5 Thesis objectives

In conclusion, the collection of novel works within this Ph.D. will contribute to the final objective of reaching atomic resolution in 3D-EM. With this goal in mind, the thesis objectives are:

- Proposing a new approach for alignment of DDD micrographs to correct for BIM by characterizing the BIM pattern.

To design such a method, we need to develop some tools to describe the movement of particles inside ice as a function of electron dose and rate correctly, and then according to this decide which methodology to use to correct for that individual pattern.

- A fast and accurate method for automatic selection of particles from electron micrographs with low SNR.

Develop a method to learn features of desired particles from the user and to use that to detect particles in new micrographs. It is important to choose features that are not only robust to noise but also discriminative to distinguish between noise and particle. Additionally, we should select a classifier which is fast to train and recognize new particles.

- A fast direct Fourier 3-D reconstruction algorithm with correction for the CTF

To propose a fast method for direct Fourier three-dimensional reconstruction which can correct for the CTF. Having a proper weighting technique to compute a uniform sampled Fourier transform is of great importance in this method. Additionally, since this approach is used by refinement algorithms frequently, a parallel implementation of this algorithm is necessary to reduce the required time for final reconstruction.

## 1.6 Publications: specific contribution of the author

This thesis consists of three peer-reviewed journals (Abrishami *et al.*, 2013, 2015a,b). It is organized as follows: in the next section, a short description of each publication and the author's contribution is given. A summary of each article and its results is described in Chapter 2. Chapter 3 includes a discussion for each method and a final discussion. Conclusions are drawn in Chapter 4. After the list of references, the below-mentioned publications are attached in Chapter A.

### 1.6.1 Summaries of the original publications

**Publication I:**

*V. Abrishami, J. Vargas, X. Li, Y. Cheng, R. Marabini, C.O.S. Sorzano, J.M. Carazo, Alignment of direct detection device micrographs using a robust Optical Flow approach", J. Struct. Biol. 189 (3), 163-176, 2015.*

In this publication, we proposed an algorithm for BIM correction by aligning frames of DDDs. This method uses a modified version of optical flow as a tool to calculate how pixels move from one frame to the next. This new method can correct for local sample movements in a rapid and efficient manner. Additionally, it provides a quantitative analysis of the BIM pattern which helps the researchers to decide about the quality of the data. A comparison is made with other well-known methods for BIM correction methods.

**Publication II:**

*V. Abrishami, A. Zaldívar-Peraza, J.M. de la Rosa-Trevín, J. Vargas, J. Otón, R. Marabini, Y. Shkolnisky, J.M. Carazo, C.O.S Sorzano, A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs", Bioinformatics, btt429, 2013*

In this paper, a method for automatic selection of particles was proposed. This method extracts some discriminative features from the manually selected particles and uses them to train an SVM classifier. Then it suggests some new particles to the user and enables him to correct the classifier by removing the noise and adding the missed particles. After the user is satisfied with the classifier performance, he allows the algorithm to pick the particles from the rest of micrographs automatically.

**Publication III:**

*V. Abrishami, J.R. Bilbao-Castro, J. Vargas, R. Marabini, J.M. Carazo, C.O.S. Sorzano, A fast iterative convolution weighting approach for gridding-based direct Fourier three-dimensional reconstruction with correction for the contrast transfer function", Ultramicroscopy 157, 79-87, 2015*

This paper presents a new method for gridding-based direct Fourier three-dimensional reconstruction that uses an iterative convolution weighting approach to computing a uniform sampled Fourier transform. This method can correct for the CTF, which is a limiting factor in pursuing a high-resolution reconstruction. The parallel implementation of this algorithm makes the process of three-dimensional reconstruction even faster. The comparison of the proposed method with similar existing methods shows that our method provides slightly better results.

# Chapter 2

# Summary of results

In this chapter, we give a summary of all articles that form this thesis. Each review provides an overview of the proposed method and its results.

## 2.1 New method for characterization of BIM and its correction by alignment of direct detection device micrographs

As was mentioned in the introduction, BIM was a barrier to obtaining high-resolution 3-D maps. By the emergence of DDDs, it is possible to acquire several frames instead of just one micrograph. This property allows the researchers to see how particles move in vitreous ice layer and correct for BIM to restore high-resolution information. Figure 2.1 shows how a whole-frame motion-correction algorithm (Li *et al.*, 2013) restores Thon rings to close to 3-Å resolution.

Still there is no agreement among researchers about the pattern of the BIM. Some scientists believe this movement is local while others consider that as a global movement. Based on this assumption, two methods are distinguishable for BIM correction:

- Global motion-correction methods: These methods use a whole-frame motion-correction algorithm to trace motion of all movie frames (Grant and Grigorieff, 2015, Li *et al.*, 2013, Wu *et al.*, 2015). The input of these algorithms is a movie and the output is a micrograph corrected for BIM.

- Local motion-correction methods: These methods correct for BIM by aligning each particle (Bai *et al.*, 2013, Rubinstein and Brubaker, 2015) or small sections of the
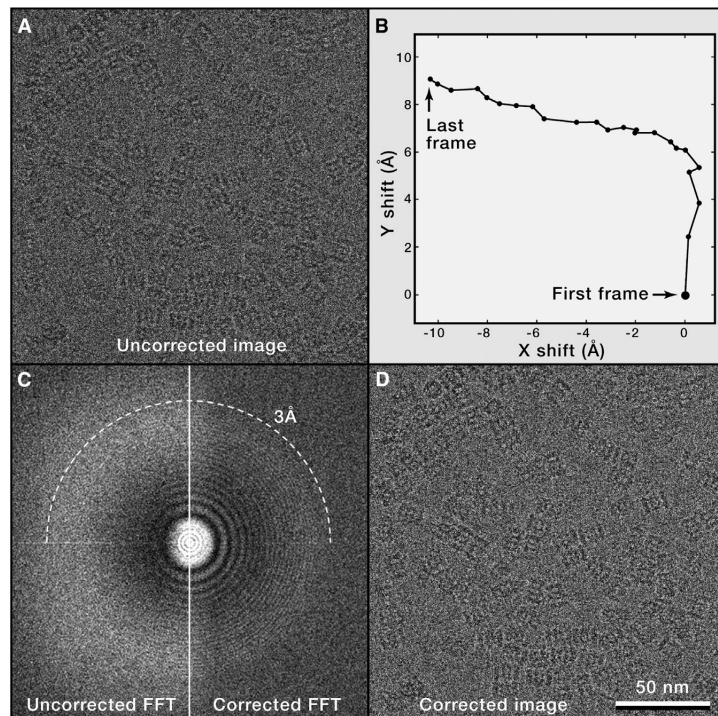
FIGURE 2.1: A micrograph obtained from the average of all frames (A), trace of the movement of all movie frames determined by Li *et al.* (2013) (B), power spectrum of the average micrograph before (left) and after correction (right) (C), the average micrograph after correction for BIM (D) (from Cheng *et al.* (2015)).

image individually (Abrishami *et al.*, 2015a, Scheres, 2014). The input for these method is a movie while the output can be a set of corrected particles or a set of corrected micrographs.

To correct for local motion without analyzing each particle, which requires considerable time and effort, a method based on optical flow has been proposed (Abrishami *et al.*, 2015a). This method can efficiently correct for BIM pattern presenting a substantial degree of local movements. Additionally, spatial analysis of the optical flow allows for quantitative analysis of the BIM pattern.

## Method

This method uses a modified version of optical flow method to align each frame against a reference.

### Optical flow estimation

Optical flow method is widely used in machine vision to estimate pixel motion from one image to another. There are two key assumptions for this approach:
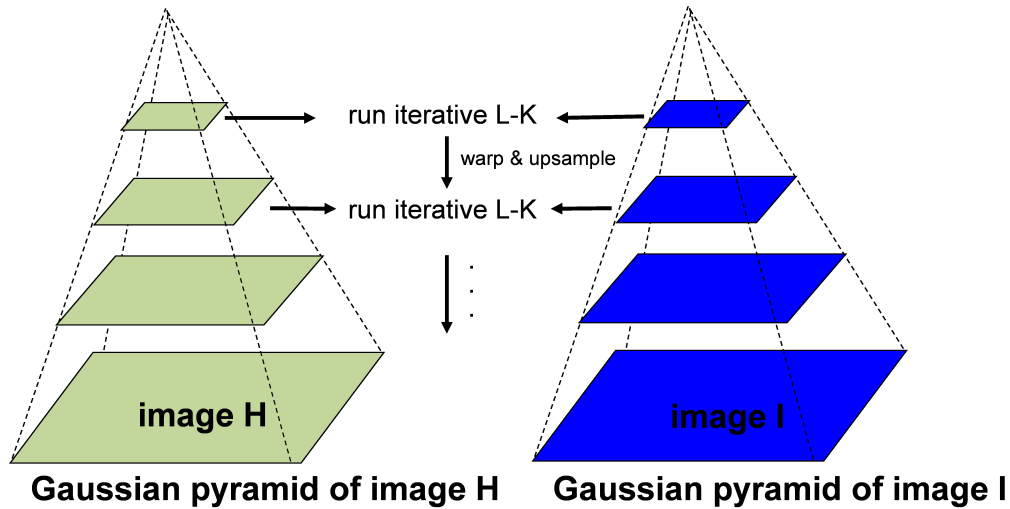
FIGURE 2.2: A coarse to fine Gaussian iterative method to recover displacements larger than one pixel (adapted from Bradski and Kaehler (2008)).

- Brightness constancy $H(x, y) = I(x + u, y + v)$.

- Small motion: $u$ and $v$ are less than 1 pixel.

where $H(x, y)$ is the first image, $I(x, y)$ is the second image after movement, and $u$ and $v$ are displacements along x and y, respectively. Small motion assumption allows taking Taylor series expansion of $I(x, y)$

$$I(x + u, y + v) = I(x, y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \tag{2.1}$$

By substitution of 2.1 in brightness constancy assumption we have

$$I_t + I_x u + I_y v = 0 \tag{2.2}$$

where $I_t = I(x, y) - H(x, y)$, $I_x$ and $I_y$ are shorthand for $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$, respectively. Equation 2.2 is called the optical flow equation. Since there is one equation, it is not possible to calculate two unknowns.

Several methods have been proposed to solve this equation. Lukas and Kanade (Lucas and Kanade, 1981) solved this equation by pretending the pixel's neighbors have the same $(u, v)$. For instance, if we use a 5x5 window, there are 25 equations per pixel as shown below

$$\begin{pmatrix} I_x(p1) & I_y(p1) \\ ... & ... \\ I_x(p25) & I_y(p25) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} I_t(p1) \\ ... \\ I_t(p25) \end{pmatrix}$$

Least squares is a standard approach to solving this over-determined system. Therefore, we have

$$\begin{pmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix} \tag{2.3}$$

Note that the summations are over all pixels in the K x K window.

However, in real scenarios, pixel displacement is more than one pixel which is in contradiction to the small motion assumption. To solve this problem, the algorithm is run in a coarse to fine manner using Gaussian pyramids (see Figure 2.2). This means that, at first, the two frames in consideration are smoothed and down-sampled, effectively reducing their resolution, so that a general estimate for the displacement vectors can be calculated. The images are then up-sampled, and the process is repeated at different levels of detail until the images have returned to their original resolutions and the most precise estimate for the feature's displacement is derived.

**Robust to noise Optical flow**

The methodology is based on a robust Optical Flow (OF) approach that can efficiently correct for local movements in a rapid manner. To describe the method, let $I_i$ be the $i$th frame in the stack, $\sum_j^k$ be the unaligned average of the $j$th subgroup of the frames in the $k$th iteration (how these subsets are formed will be described later), $\tilde{\sum}^k$ be the aligned average in the $k$th iteration obtained using all images and, finally, let $\tilde{\sum}_j^k$ be the aligned average of the images belonging to the subset $j$ in iteration $k$. Next, the optical flow algorithm is described as follows:

1. For k=1, the algorithm computes the average of all $N$ unaligned frames $\sum_1^1 = (\sum_{i=1}^N I_i)$ and initializes the refined average estimation as $\tilde{\sum}_1^1 = \sum_1^1$. Next, $k$ is updated as $k = k + 1$, and for $k > 1$, the following iterative process is executed.

2. Images are split into $2^{(}k - 1$ subgroups, and the unaligned average of each group is computed. Therefore, there are $2^{k-1}$ averages with $\frac{N}{2^{k-1}}$ images in each average such that $\sum_j^k = (\sum_{i=(j-1)s^k+1}^{js^k} I_i)$, where $s^k = \frac{N}{2^{k-1}}$. Note that for $k = 2$, the set

**Level 1**

$$\Sigma_1^1 = \left(\sum_{i=1}^{16} I_i\right) \longrightarrow \quad \widetilde{\Sigma}^1 = \Sigma_1^1 \quad OF_1^1 = 0$$

**Level 2**

$$OF_1^2 = OptFlow_{OF_1^1}(\Sigma_1^2, \widetilde{\Sigma}^1) \quad \widetilde{\Sigma}_1^2 = (\Sigma_1^2)_{OF_1^2}$$

$$OF_2^2 = OptFlow_{OF_1^1}(\Sigma_2^2, \widetilde{\Sigma}^1) \quad \widetilde{\Sigma}_2^2 = (\Sigma_2^2)_{OF_2^2}$$

$$\Sigma_1^2 = \left(\sum_{i=1}^{8} I_i\right) \Sigma_2^2 = \left(\sum_{i=9}^{16} I_i\right) \longrightarrow \quad \widetilde{\Sigma}^2 = \frac{1}{2}(\widetilde{\Sigma}_1^2 + \widetilde{\Sigma}_2^2)$$

**Level 3**

$$OF_1^3 = OptFlow_{OF_1^2}(\Sigma_1^3, \widetilde{\Sigma}^2) \quad \widetilde{\Sigma}_1^3 = (\Sigma_1^3)_{OF_1^3}$$

$$OF_2^3 = OptFlow_{OF_1^2}(\Sigma_2^3, \widetilde{\Sigma}^2) \quad \widetilde{\Sigma}_2^3 = (\Sigma_2^3)_{OF_2^3}$$

$$\Sigma_1^3 = \left(\sum_{i=1}^{4} I_i\right) \Sigma_2^3 = \left(\sum_{i=5}^{8} I_i\right) \Sigma_3^3 = \left(\sum_{i=9}^{12} I_i\right) \Sigma_4^3 = \left(\sum_{i=13}^{16} I_i\right) \quad OF_3^3 = OptFlow_{OF_2^2}(\Sigma_3^3, \widetilde{\Sigma}^2) \quad \widetilde{\Sigma}_3^3 = (\Sigma_3^3)_{OF_3^3}$$

$$OF_4^3 = OptFlow_{OF_2^2}(\Sigma_4^3, \widetilde{\Sigma}^2) \quad \widetilde{\Sigma}_4^3 = (\Sigma_4^3)_{OF_4^3}$$

$$\widetilde{\Sigma}^3 = \frac{1}{4}(\widetilde{\Sigma}_1^3 + \widetilde{\Sigma}_2^3 + \widetilde{\Sigma}_3^3 + \widetilde{\Sigma}_4^3)$$

**Level 4**
$\vdots$

**Level 5**

$$\longrightarrow \quad \widetilde{\Sigma}^5 = \frac{1}{16}\sum_{i=1}^{16} \widetilde{\Sigma}_i^5$$

FIGURE 2.3: Diagram of the proposed Optical Flow method: The required steps for alignment of a video with 16 frames using the proposed Optical Flow method (from Abrishami *et al.* (2015a)).

is divided into two groups, and for each group, we obtain the unaligned averages $\sum_1^2$ and $\sum_2^2$.

3. Next, we compute the LKPDI optical flows between the new set of unaligned averages $\sum_j^k$ and the corresponding set of aligned images computed in the previous iteration $\widetilde{\sum}^{k-1}$. The result, which is a map of shift vectors centered at each pixel that aligns $\sum_j^k$ with $\widetilde{\sum}^{k-1}$, is known as $OF_j^k$. The estimation of $OF_j^k$ requires an initial guess; our algorithm uses $OF_{j/2+j mod 2}^{k-1}$ for $k > 2$ and a zero-valued matrix of displacement or shift vectors for $k = 2$.

4. Using the obtained shift vector matrix $OF_j^k$, the averages $\sum_j^k$ are aligned; we refer to these aligned averages as $\widetilde{\sum}_j^k$. A new corrected average is computed as $\widetilde{\sum}^k = \sum_{j=1}^{2^k}(\widetilde{\sum}_j^k)/2^k$. If the groups contain more than one frame, Steps 2 to 4 are repeated.

Figure 2.3 shows a diagram of the proposed approach for a movie composed of 16 frames. This procedure is applicable to movies with a number of images that is not a power of 2 by simply readjusting the appropriate indices. Note that this alignment strategy improves the robustness of the approach against noise, and at the same time, it reduces

other potential issues, such as the validity of the brightness constancy equation due to radiation damage or changes in brightness resulting from mass loss during exposure.

**Implementation**

The software that incorporates the new approach is available in Xmipp (`http://xmipp.cnb.csic.es`). To compute optical flow, the OpenCV (Open Source Computer Vision Library) (Bradski and Kaehler, 2008) implementation of the Lucas-Kanade dense optical flow (Lucas and Kanade, 1981) was used, which provides both CPU and GPU execution modes.

## Results

The proposed method was compared to the global motion correction algorithm of Li *et al.* (2013) which uses cross-correlation to align frames against a reference. Results were shown for both the Falcon II (FEI) and the K2 Summit (in counting mode). In this research, the envelope of the CTF and SSNR were introduced as efficient criteria to compare the result of different alignment algorithms.

**Results for the Falcon II (FEI)**

The results for the Falcon II showed that the BIM pattern presents a high degree of local movements. Since OF works particularly well if the BIM pattern displays a substantial degree of local changes, for this case OF worked perfectly and provided better results than Li *et al.* (2013). Dividing the image into nine overlapping regions and applying the Li *et al.* (2013) algorithm to each part improved the results of the Li *et al.* (2013) algorithm, but still OF could align frames more efficiently than Li *et al.* (2013). Figure 2.4 shows the alignment result for a movie of the Falcon II. Note that in Figure 2.4c, the area of each circle, indicates the degree of locality of the movement between two frames.

**Results for the K2 Summit**

The results showed that, in contrast to the results for the Falcon II (FEI), the BIM pattern for the K2 Summit corresponds to global movements. Therefore, for K2 data the additional improvement of OF method is reduced. However, applying the OF after correction for global movements using the Li *et al.* (2013) algorithm improved the results slightly. The results also showed that for movies where there is a pure in-plane drift,

the OF method can not recover the movements since it has a first step of averaging sets of frames without correction. Figure 2.5 shows the alignment result for a movie of the K2 Summit.

**(a)**



**(b)**



**(c)**
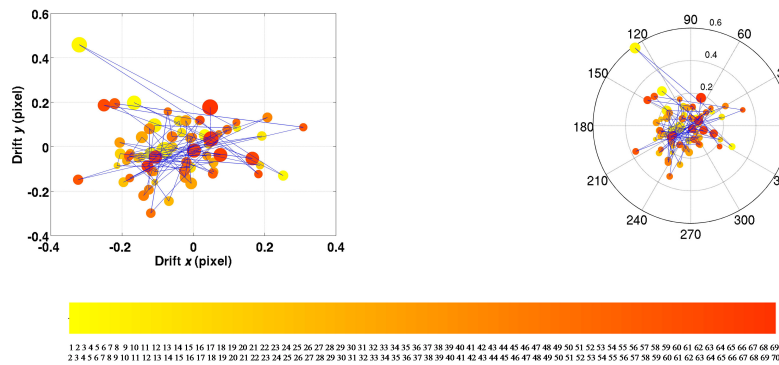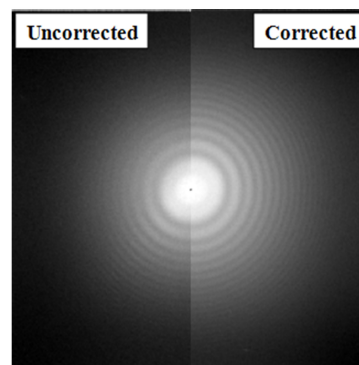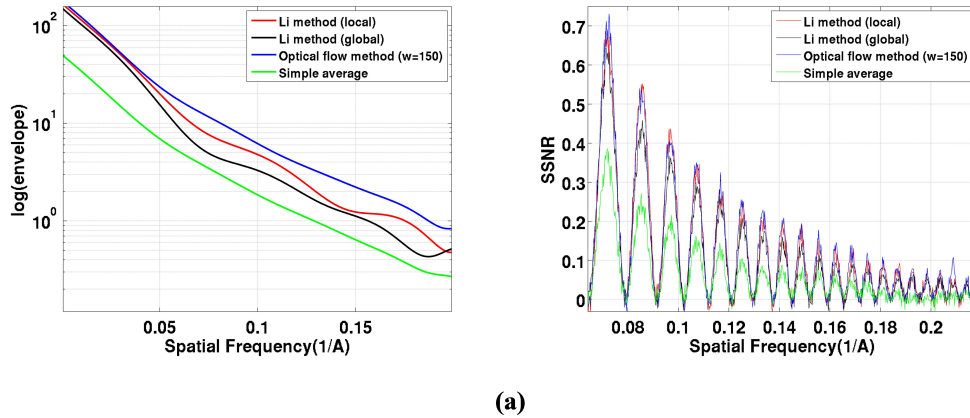
FIGURE 2.4: Comparison and analysis of the alignment result for the influenza virus RNP dataset: comparison of the envelopes of the CTF (in logarithmic scale) as well as the SSNR functions corresponding to the average image obtained from the summation of unaligned frames and aligned frames using Li Local, Li Global and Optical Flow with a window size of 150 pixels (a). The periodogram for the uncorrected (left) and corrected averages (right) (b), and the analysis of the movements between consecutive frames (c) (from Abrishami *et al.* (2015a)).
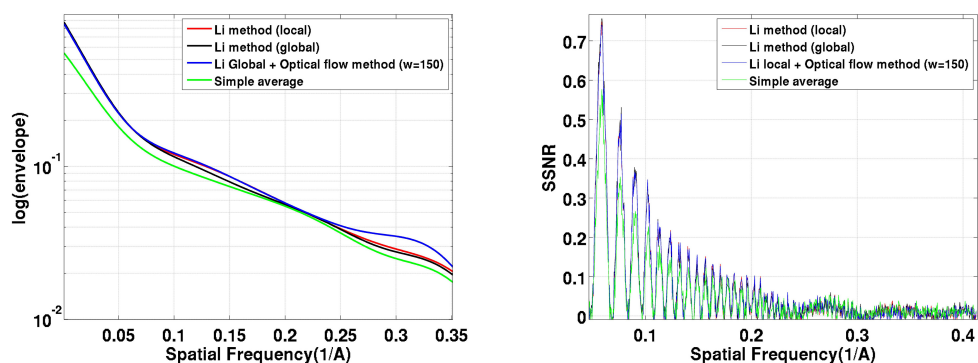
**(a)**
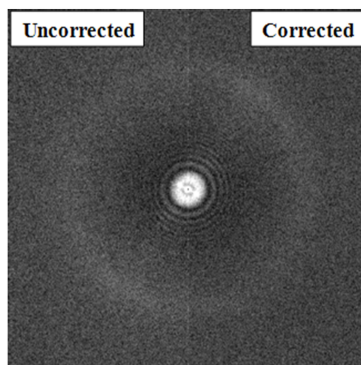


**(b)**
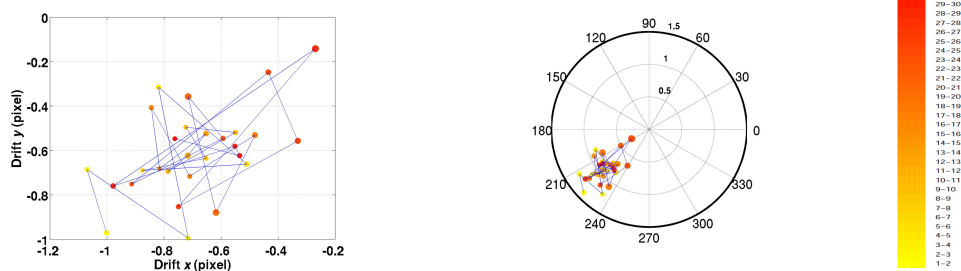


**(c)**

FIGURE 2.5: Analysis of 20S proteasome images obtained on a K2: 1D profiles corresponding to the envelope of the CTF function (in logarithmic scale) as well as the SSNR functions for the average image of a representative stack obtained using simple average (green), Li Local (red), Li Global (black) and, finally, Li Global + Optical Flow (blue) with a window size of 150 (a). Comparison of the periodograms of the uncorrected (left) and Li Global + Optical Flow (right) aligned averages (b), and analysis of the movements in between consecutive frames of the stack (c) (from Abrishami *et al.* (2015a)).

## 2.2 A fast and accurate method for automatic selection of particles from electron micrographs

After alignment of movies and obtaining a set of good micrographs, the project continues with the labor-intensive process of particle picking. This step is significant since the selection of many poor quality particles may preclude successful structure determination. Methods for assessment the quality of particles normally are not that robust, and allow many types of noise to contribute to the subsequent data processing which can affect the 3-D final map adversely.

Particles can be selected from micrographs using one of the following methods:

- Manually: where a professional can pick particles. Manually identifying that number of particles from noisy micrographs is not just time-consuming and laborious, but also an error-prone process.

- Semi-automated: where the algorithm suggests particle candidates to the user and he can remove poor quality particles.

- Fully automated: where the algorithm suggests particles to the user.

Fully automated method can be classified into two categories:

- Template matching-based methods (Chen and Grigorieff, 2007, Huang and Penczek, 2004, Ludtke *et al.*, 1999, Plaisier *et al.*, 2004, Roseman, 2003, Sigworth, 2004, Wong *et al.*, 2004) calculate the cross-correlation (or any other measure of similarity) between a set of templates and a micrograph image to seek for particle candidates.

- Feature-based approaches where particles are sought through the calculation of some prominent geometric and/or statistical features of the particle images (Arbeláez *et al.*, 2011, Hall and Patwardhan, 2004, Langlois *et al.*, 2011, Mallick *et al.*, 2004, Ogura and Sato, 2004, Sorzano *et al.*, 2009, Volkmann, 2004, Yu and Bajaj, 2004, Zhao *et al.*, 2013, Zhu *et al.*, 2003). Feature-based methods can be reference-free or learning-based. In the former, features corresponding to particles are known to fall within a particular region of the feature space and, therefore, no training is necessary, and the algorithm can start picking particles straightaway. In the latter, however, a set of particles and non-particles are required to train a classifier which is then able to distinguish between particles and non-particles based on the training features.
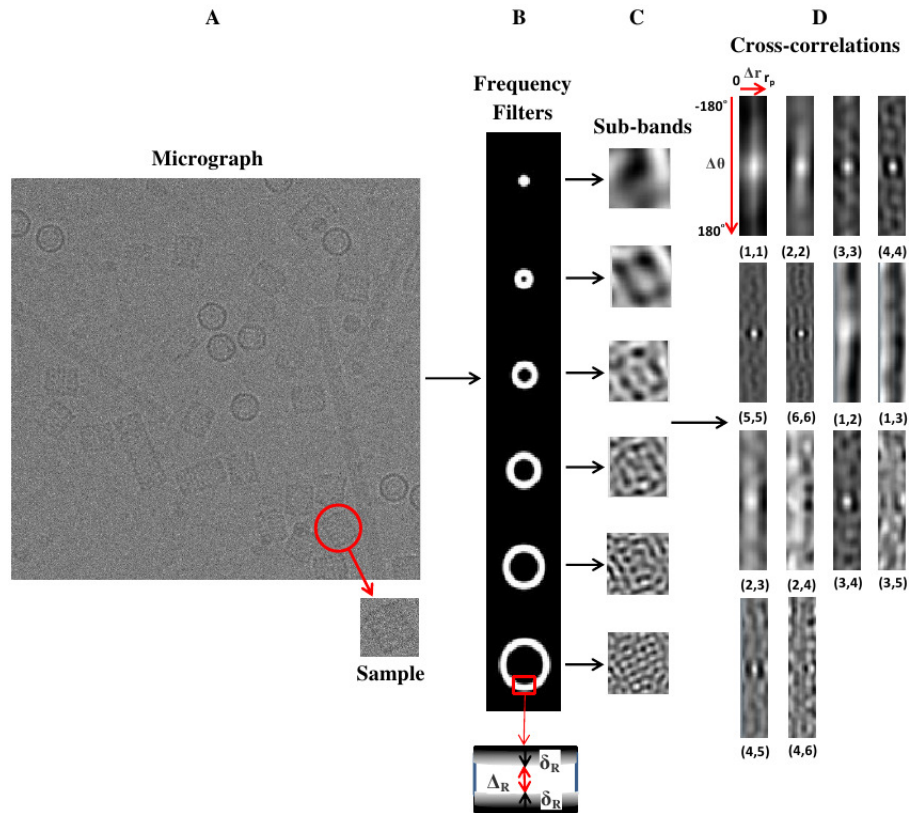
FIGURE 2.6: Particle shape description at different frequencies. First, 6 band-pass filters (column B) are applied to the input micrograph (column A), and boxed images are extracted at the location of particles or particle candidates (column C). Sub-band images for the particle (column C) are converted to polar form and cross-correlations between different sub-bands are calculated (column D) (from Abrishami *et al.* (2013)).

In Abrishami *et al.* (2013), a feature-based approach for automatic picking of particles was proposed. This method uses a classifier to learn some extracted features from particles, and then to detect particles in new micrographs.

## Method

### Identification of possible particle centers

After choosing a few particles by the user in the training phase, the algorithm clusters manually picked up particles into a few number of classes. Then it cross-correlates the input micrograph with the rotational average of each class representative to identify the local maxima (particle candidates).

**(a)**



**(b)**
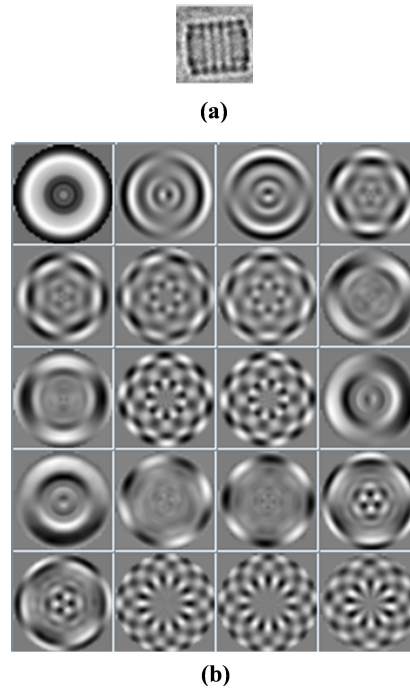
FIGURE 2.7: Eigenvectors (Ponce and Singer, 2011) corresponding to a given template. Given particle template (a) and first 20 eigenvectors of the template that generate a rotational invariant sub-space (b) (from Abrishami *et al.* (2013)).

### Feature extraction

Three types of features were used to distinguish between particle and noise. These features are robust to noise (due to the low SNR in the micrographs) and rotational invariant (to save computational time and avoid having to look for the particles in all possible orientations). These features are:

- *Particle shape description at different frequencies*: Micrographs were submitted to a filter bank with some raised cosine band-pass filters to decompose them into several sub-band images to being able later on to extract features associated with particular frequencies. Then sub-band images of each particle were extracted from filtered micrographs and were converted to polar form. The auto-correlation of a given band, as well as the cross-correlation between sub-bands, were computed to reveal the particle shape information and the linear relationships between the shapes at two different frequency bands, respectively (see Figure 2.6). The cross-correlation functions of the training set of particles for a particular combination (two sub-band images) were compressed using principal component analysis (PCA).

- *Particle shape description in a particle-adapted rotational invariant subspace*: Ponce and Singer (2011) proposed a method to calculate an image basis that is adapted

to the kind of images being studied and their in-plane rotations. Given the training particles provided by the user, the proposed method aligns them into a few templates using a clustering algorithm and computes the PCA basis associated with these templates (see Figure 2.7). Projection coefficients onto some of this basis are kept as part of the feature vector.

- *Particle intensity*: The mean, the standard deviation as well as a number of equidistributed deciles of the intensity histogram are used in feature vector.

A feature vector of size 91 were obtained (using default parameters of the program).

### Classification

This method uses non-linear, binary support vector machines (SVMs) as classifiers due to their good performance in other classification problems. Two SVM classifiers are used. The first classifier discriminates between particle objects and any other kind of objects (non-particles and errors). Because of the similarity between errors and particles, the output of the first classifier is not so accurate, and some errors are labeled as particles. Therefore, to reduce the false positives to a feasible extent, the second classifier is dedicated to just focus on distinguishing particles from errors (see Figure 2.8).

### Implementation

The software that incorporates the new approach is available in Xmipp (`http://xmipp.cnb.csic.es`). LIBSVM (Chang and Lin, 2011) is an efficient and widely used implementation of the SVM. This package suggests a variety of kernels to perform the non-linear classification. This package was used with a Radial Basis Function (RBF) kernel to gain a high accuracy classification.

### Proposed automatic particle picking in action

This method first allows the user to pick some particles manually (at least 15 particles) from the first micrograph. Then it automatically selects some non-particles and computes feature vectors for particles and non-particles. These feature vectors are used to train the first classifier. After training the classifier, it proposes new particles in the next micrograph and allows the user to correct the classifier by adding those missed particles (false negatives) and removing wrongly picked particles (false positives). At this point, the first classifier is trained with both newly added particles and removed
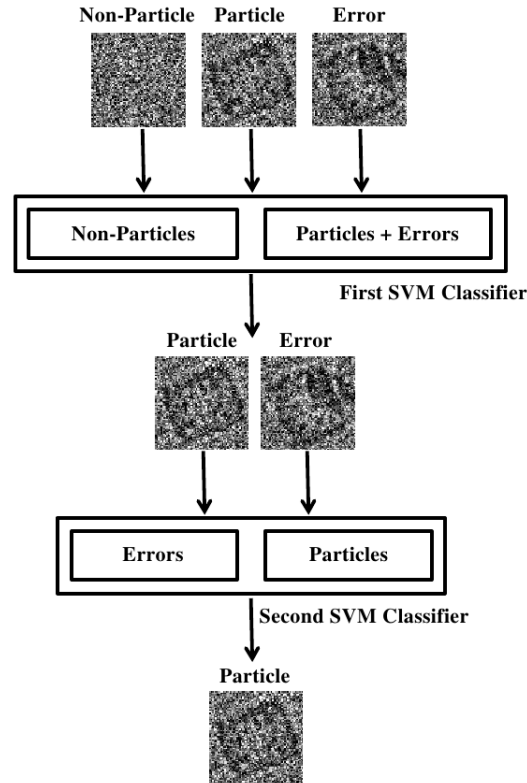
FIGURE 2.8: The behavior of the classifiers for three types of objects. The first classifier may classify errors as particles, but the second classifier is designed to remove errors from the final result (from Abrishami *et al.* (2013)).

particles while the second classifier is trained with removed particles (errors) and particles. Then, the user can continue with other micrographs to improve the accuracy of the classifiers. After the user is satisfied with the result of the classifiers, the algorithm selects the particles from the rest of micrographs automatically.

## Results

Three datasets were used to assess the speed and accuracy of the proposed algorithm for micrographs with different contrast, density and particle shape and size: KLH dataset, Adenovirus dataset, and Helicase dataset. In the 3DEM Benchmark site (http://i2pc.cnb.csic.es/3dembenchmark), there are datasets for KLH (with 30 micrographs for training and another one with 50 micrographs for testing) and Adenovirus (30 micrographs for training and 275 for testing).

To show the results in a quantitative way, we have used the performance metrics introduced by Langlois and Frank (2011). If TP is the number of true positives, FP the number of false positives, and FN the number of false negatives, then these metrics are defined as
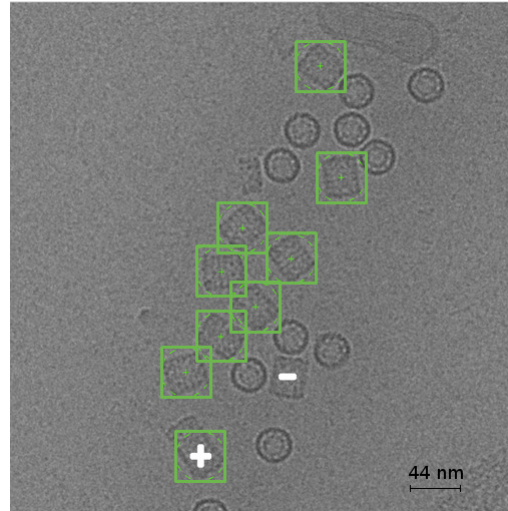
FIGURE 2.9: The result of the algorithm for one micrograph of the KLH dataset with 9 particles. 9 objects are boxed in green as particles. There is one false positive and one false negative, which are identified by + and - respectively (from Abrishami *et al.* (2013)).

- Precision $= \dfrac{\text{TP}}{\text{FP+TP}}$

- Recall $= \dfrac{\text{TP}}{\text{FN+TP}}$

- F-measure $= 2 \times \dfrac{\text{precision} \times \text{recall}}{\text{precision+recall}}$

For KLH dataset, Training with more than 173 particles (13 micrographs), a precision range of [80.0, 85.2] and a recall range of [73.1, 77.8] were achievable. F-measure ranges as a summary of precision and recall in the interval [77.53, 80.06]. Training the algorithm with 39 manually picked particles took 3.5 s. After training, the algorithm needs 1 s to suggest new particles on each new micrograph and 2.5 s to retrain after being corrected by the user. Picking particles from 50 micrographs of the test dataset took 51 s (see Figure 2.9).

For Adenovirus dataset, 334 particles and 690 non-particles from the first 14 micrographs were used to train fully the classifier. The classifier was used to automatically pick the particles of the test dataset. This experiment resulted in picking 9216 particles with precision rate 92.17% and recall rate 90.24% (see Figure 2.10). The average processing time per micrograph was 34 s.

For Helicase dataset, no benchmark was available to check the accuracy quantitatively. Therefore, the result was given qualitatively (see Figure 2.11).
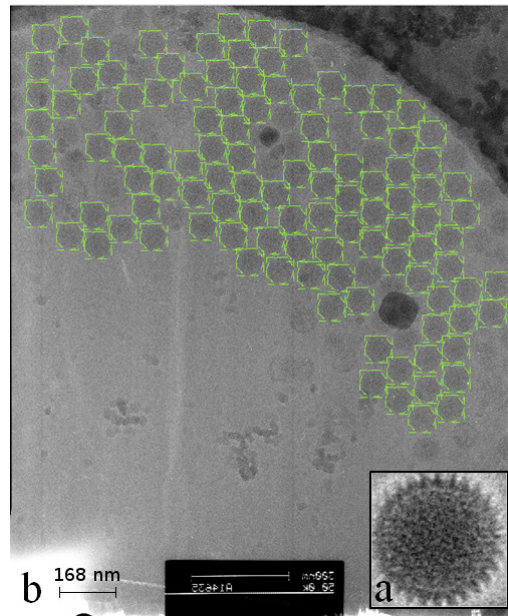
FIGURE 2.10: Automatic selection for one micrograph of the Adenovirus dataset. The template for cross-correlation (a), the micrograph with 115 automatically selected particles (b) (from Abrishami *et al.* (2013)).
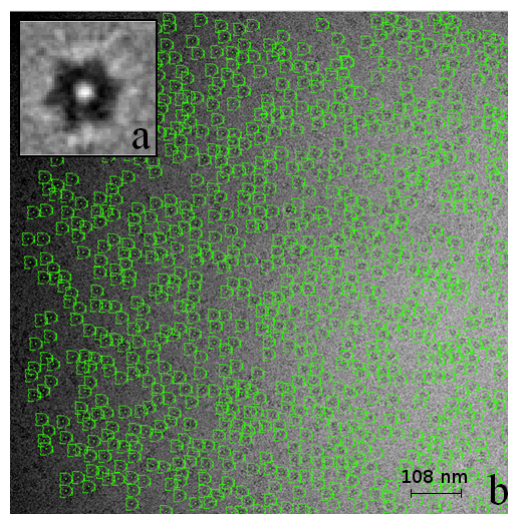


FIGURE 2.11: Automatic selection for the third micrograph of the Helicase dataset. The template for cross-correlation (a), 874 automatically picked particles from the micrograph (b) (from Abrishami *et al.* (2013)).

## 2.3   A fast and accurate direct Fourier three-dimensional reconstruction with correction for the CTF

Based on the central slice theorem, Direct Fourier Reconstruction (DFR) methods try to obtain the 3-D Fourier transform of an object directly from the 2-D Fourier transform of its projections, so that an estimation of the original 3-D object can be quickly obtained through an inverse 3-D Fast Fourier Transform (FFT). In practice, the irregularity of the spatial distribution of the frequency of samples in the set of projections in experimental SPA studies makes the direct use of the inverse FFT unfeasible. Thus, an additional interpolation step is required to obtain the 3-D Fourier transform of the object on a regular grid. In Abrishami *et al.* (2015b) a gridding-based direct Fourier method for the 3-D reconstruction approach that uses a weighting technique to compute a uniform sampled Fourier transform was proposed. Moreover, the CTF of the microscope, which is a limiting factor in pursuing a high-resolution reconstruction, is corrected by the algorithm.

### Method

The goal of the gridding-based direct Fourier method is to approximate frequency samples on a regular 3-D Cartesian lattice $F_{3D}(\bar{R})$ from the measured samples of the 3-D frequency domain $\hat{F}_{3D}(\bar{Q})$ as

$$F_{3D}(\bar{R}) = \int CTF^{-1}(\bar{Q})\hat{F}_{3D}(\bar{Q})K(\bar{R} - \bar{Q})d\bar{Q} \qquad (2.4)$$

where $\bar{R}$ is the frequency coordinate within the regular 3-D grid and $K$ is the kernel function by which the integration is accomplished.

Under experimental conditions, a limited number of projections from the specimen are available. Therefore, the discrete form of the integral in Eq. 2.4 should be considered because measurements are only available for a finite set of frequencies $\bar{Q} \in \{\bar{R}_i\}$ . To obtain a discrete form of Eq. 2.4, the integral is substituted by a summation as shown

$$F_{3D} = \sum_i CTF^{-1}(\bar{R}_i)\hat{F}_{3D}(\bar{R}_i)K(\bar{R} - \bar{R}_i)w(\bar{R}_i) \qquad (2.5)$$

where $w(\bar{R}_i)$ is the weighting factor for the i-th irregular sample. It is important to note that the weighting function is a substitution for $d\bar{Q}$ within Eq. 2.4.

**Calculation of the weighting function**

The proposed method is an iterative numerical approach for computing the weighting function $w(\bar{R}_i)$ based on the method introduced by Matej and Lewitt (2001). The algorithm works as follow:

1. The proposed approach begins with an initial weighting function that is initialized as a constant function of value 1 as

$$w^{(0)}(\bar{R}_i) = 1 \tag{2.6}$$

2. Then, this weighting function is refined by evaluating Equation 2.5 to compute $c^{(n+1)}(\bar{R})$ at the position of each point on the 3-D regular grid, as below

$$c^{(n+1)}(\bar{R}) = \sum_i w^n(\bar{R}_i) K(\bar{R} - \bar{R}_i) \tag{2.7}$$

where $w^n(\bar{R}_i)$ is the interpolated weight from $w^n(\bar{R})$ of the $n$-th iteration at the position of the $i$-th input sample. Trilinear interpolation is used to interpolate $w^n(\bar{R}_i)$ from $w^n(\bar{R})$. $c^{(n+1)}(\bar{R})$ is the evaluation of Eq. 2.5 when all samples have a value of 1 and there is no CTF.

3. After computing Eq. 2.7 at the position of each frequency point on the 3-D regular grid, the algorithm updates the weights from the previous step by dividing them by the obtained convolution in Eq. 2.7:

$$w^{(n+1)}(\bar{R}) = \frac{w^{(n)}}{c^{(n+1)}(\bar{R})} \tag{2.8}$$

4. Steps 2-3 are repeated until the reconstructed function $F_{3D}(\bar{R})$ is as close to 1 as desired (formally that the Chebyshev norm of the function $F_{3D}(\bar{R}) - 1$ is smaller than $\epsilon$, $\|F_{3D}(\bar{R}) - 1\|_\infty < \epsilon$.

At the end, $w^{(n)}(\bar{R})$ can be used to correct the values of regular samples to overcome the non-uniformity problem.

The algorithm also compensate for the convolution kernel (equivalent to a multiplication in real space of the function $f_{3d}(\bar{r})$ and the inverse Fourier transform of the kernel $k(\bar{r})$) and interpolation in Fourier space (trilinear interpolation to estimate $w(\bar{R}_i)$ from $w(\bar{R})$ in frequency space).
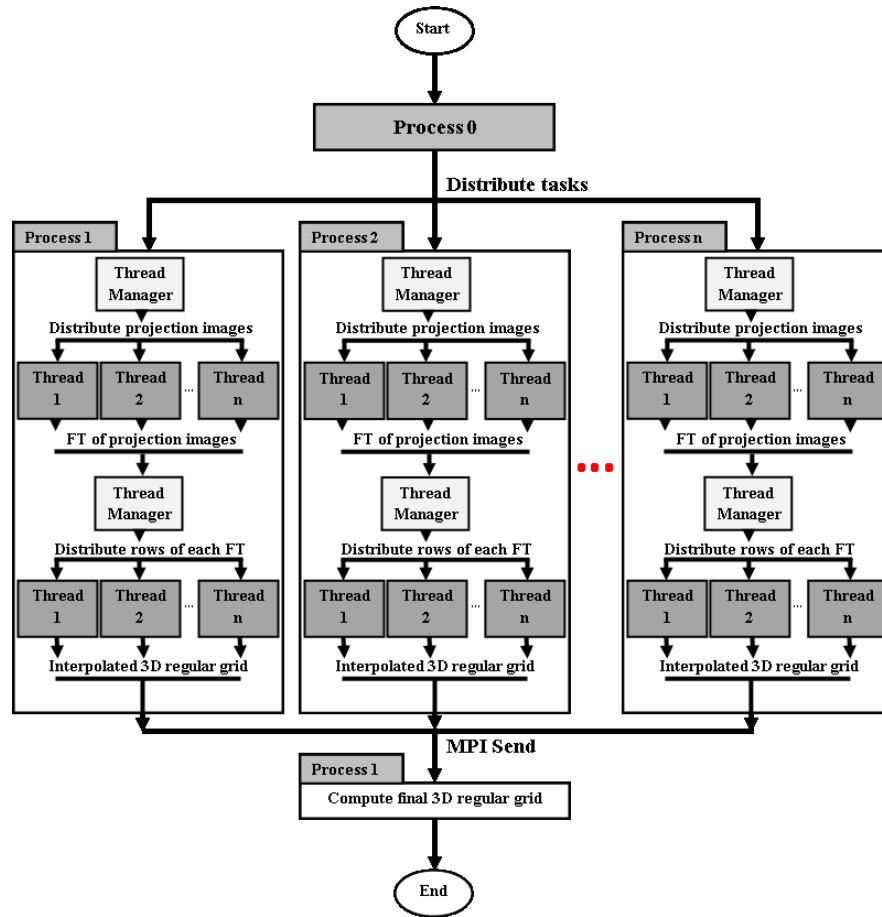
FIGURE 2.12: The two levels of parallelization of the proposed algorithm, including multiple threads and multiple processors are shown. The algorithm can be executed on several nodes of a cluster (multi-CPU parallelism) with several threads (multi-thread parallelism) (from Abrishami *et al.* (2015b)).

**Parallelization of the algorithm**

The processing time of the algorithm depends on the number and size of the projections, which may amount to hundreds of thousands of projections with several hundred pixels per side for SPA. However, most of the processing steps can be accomplished independently on each projection image, which makes the algorithm naturally suitable for parallelization to achieve higher processing speeds. Parallelization of the algorithm has been performed at two levels: at the level of multiple processors distributed over several computing nodes using the Message Passing Interface (MPI) (Gabriel *et al.*, 2004) and at the level of threads executing in the same node using the POSIX threads (Mueller, 1994) (see Figure 2.12).
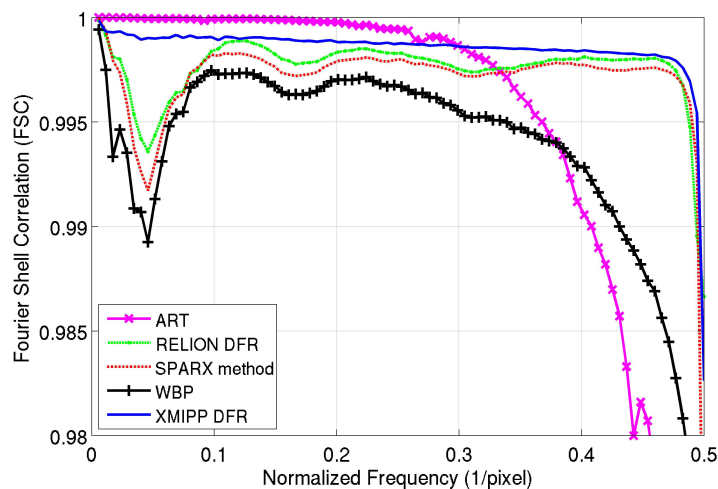
FIGURE 2.13: Comparison of the proposed reconstruction method with RELION method, SPARX method, ART and WBP (from Abrishami *et al.* (2015b)).

## Results

The proposed method was compared with two other well-known methods: (1) NN direct inversion implemented in SPARX (Zhang *et al.*, 2008) and (2) RELION DFR (Scheres, 2012a) in terms of accuracy, memory complexity, and the time required for execution. Two asymmetric test objects were used to achieve this goal: first, a 70S ribosomal subunit as an instance of a small complex and second, a DNA-origami object as a good representative of a large complex. Different data sets (set of projections) with specific properties (i.e., noise in projections, noise in angular assignment, and CTF-affected) were made for each test object to examine the robustness of each method.

## Results for 70S ribosomal subunit

$N = 10,000$ projections of 70S ribosomal subunit (PDB ID: 3V2D) represented on a cube of size $174 \times 174 \times 174$ voxels (the sampling rate was 1.5 Å/pixel) were used.

- Projections without noise: Comparison of the proposed method with RELION method (Scheres, 2012a), SPARX method is shown in Figure 2.13. The proposed method is slightly more accurate at all frequencies than either method. For this reconstruction, the serial implementation of the proposed method required 13 min and 40 s, while it took 5 min and 37 s for RELION and 20 min and 44 s for SPARX using one core of the CPU. In Figure 2.14, it is shown how increasing the number of threads can decrease the required computational time for the proposed method to one-fourth of the SPARX and half of the RELION reconstruction methods. In
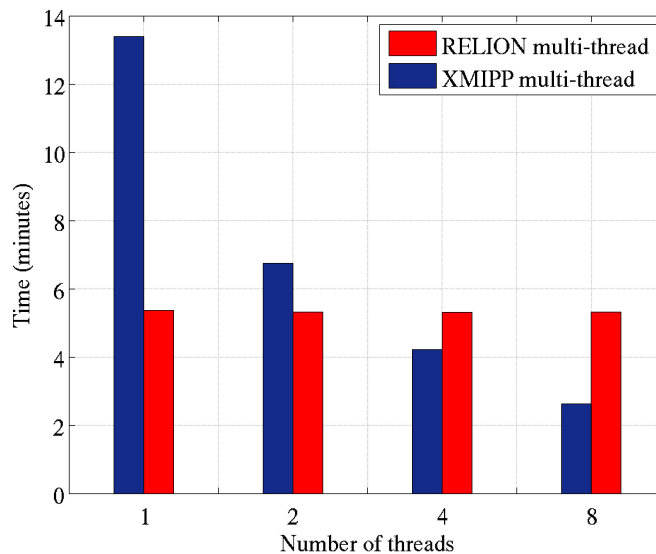
FIGURE 2.14: Computational time required for the 3-D reconstruction of the 70S ribosome from a set of $N = 10,000$ noise-free projections applying the proposed (Xmipp) and RELION reconstructions using different numbers of threads (from Abrishami *et al.* (2015b)).

Figure 2.15, the average execution times of 10 independent runs for the proposed method and the SPARX method using different numbers of nodes were shown. It is obvious from this figure that the parallel implementation of the proposed algorithm is faster than the parallel implementation of SPARX.

- CTF applied projections: The same as the previous experiment, but 10 different sets of CTF parameters were applied to the projections. The comparison results are shown in Figure 2.16. As seen in this figure, the proposed algorithm could properly correct for the CTF.

- Projections with noise: Two different types of noise into the simulated projections to generate two sets of noise-corrupted projections. First, a Gaussian noise added to the pixels of the noise-free projections. Second, a Gaussian noise was applied to each projection direction. The comparisons in Figures 2.17 and 2.18 show that there is a region of low frequency at which the FSC values are larger for the proposed method

**Results for a 3-D DNA-origami object**

The purpose of this experiment was to determine how each reconstruction algorithm can manage memory for a large complex. For this purpose, the model of a discrete DNA
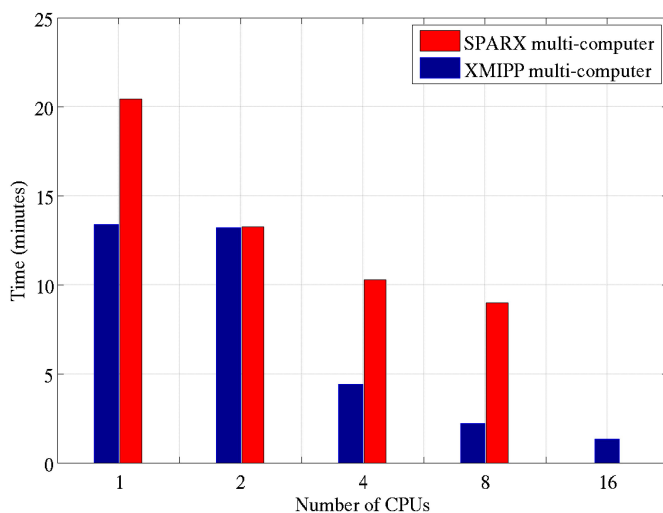
FIGURE 2.15: The required computational time for the parallel execution of the proposed reconstruction method and the SPARX method for reconstructing the 70S ribosome using a set of 10,000 noise-free projections utilizing different numbers of CPUs (from Abrishami *et al.* (2015b)).
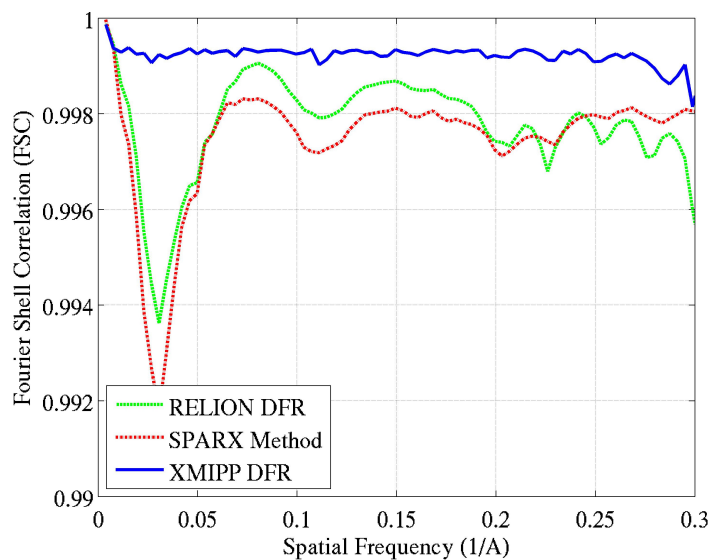


FIGURE 2.16: Comparison of the proposed reconstruction method with RELION method, SPARX method using CTF-affected projections of the 70S ribosome (from Abrishami *et al.* (2015b)).
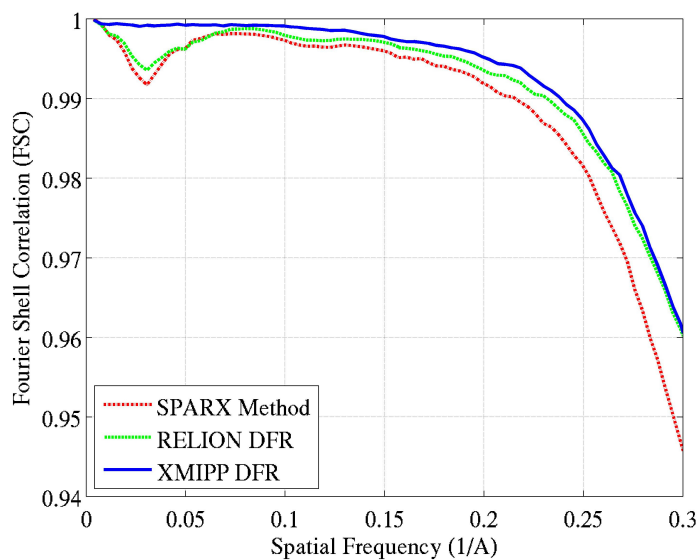
FIGURE 2.17: Comparison of the proposed reconstruction method with RELION method, SPARX method using projections of the 70S ribosome that were contaminated by pixel (from Abrishami *et al.* (2015b)).
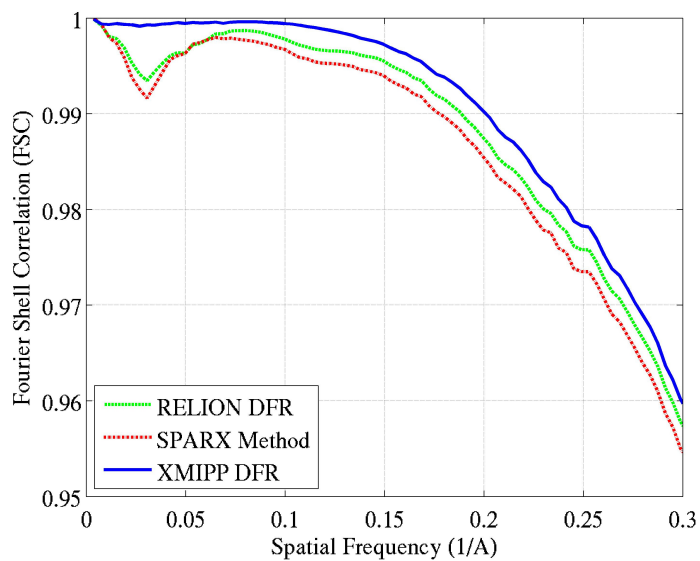


FIGURE 2.18: Comparison of the proposed reconstruction method with RELION method, SPARX method using projections of the 70S ribosome that were contaminated by angular noise using our proposed (from Abrishami *et al.* (2015b)).

object (Bai *et al.*, 2012) (PDB ID: 2YMF) of size $400\times400$ (sampling rate 1.2 Å/pixel) was considered.

As in the previous experiment, $N = 10,000$ projections at random orientations of this complex were generated and used to compare the proposed method with RELION method, and SPARX method. Due to the large size of this complex, neither the SPARX nor RELION reconstruction algorithms (with one CPU core) nor the parallel implementation of the proposed algorithm can be executed on a machine with 16 gigabytes of RAM memory. However, the proposed method can take advantage of the multi-thread implementation to use the process resources efficiently without a lack of memory.

# Chapter 3

# Discussion

The discussion chapter is structured in two sections. In the first one I present a discussion of each included paper in this thesis, in the same order in which they appear in Chapter 2. The second section addresses the contribution of this thesis work to the field of Single Particle Analysis, in the form of a general discussion.

## 3.1 New method for characterization of BIM and its correction by alignment of direct detection device micrographs

Direct electron detectors allow researchers to correct for BIM by providing several frames instead of just one micrograph. Several methods has been proposed to correct for BIM by alignment of the frames (Bai *et al.*, 2013, Campbell *et al.*, 2012, Li *et al.*, 2013, Wang *et al.*, 2014).The method shown in Li *et al.* (2013) consists of a frame drift-like correction with a degree of robustness to local movements provided by the use of subframe alignment; the output of each processed movie is a corrected averaged micrograph obtained from the aligned frames. In Bai *et al.* (2013) the alignment method is quite different, since it works with individual specimens detected and extracted from the frames, as Scheres (2014), also does. This approach allows local movement correction in-plane and out-of-plane, at the expense of requiring individual specimen detection.

The proposed method by Abrishami *et al.* (2015a) can correct for local sample movements in a fast and efficient way using optical flow algorithm, and can provide a movie with the vector field of the movements between frames that can be further analyzed using intuitive 2D representations showing frame to frame changes.

Since the proposed method uses the average of all frames as an initial reference for alignment, it is sensitive to unidirectional drift in-between frame movements. For these cases, the algorithm of Li *et al.* (2013) first can correct for drift-like global movements, and the proposed method can compensate for local changes.

OF method can correct for local in-plane movements at the image level, and it cannot incorporate out-of-plane movements. However, this is not important for small specimens since out-of-plane movements would translate into a small shift at the periphery that is negligible. For larger particles for which rocking might be a problem, the initial use of OF may provide partly corrected frames that might significantly increase the speed of convergence of any individual specimen image-based correction approach.

This method also provides a simple and intuitive way to characterize the BIM pattern. According to obtained results for different detector, Falcon II data presents a larger degree of local distortions than K2 data, making the Optical Flow approach particularly powerful for Falcon II data sets, while for K2 data the additional improvement is reduced. This contrasting behaviour may not be so unexpected, as detection mechanisms of both cameras are different.

Moreover, new figures of merit have been proposed in order to compare the different alignment results Indeed, the effect of summing unaligned frames is an additional modulating term in the Power Spectral Density. Therefore, the CTF envelope as well as the SSNR functions are sensitive to these in-between frames misalignments, typical of DDD's.

## 3.2 A fast and accurate method for automatic selection of particles from electron micrographs

The proposed method for automatic identification of particles from electron micrographs presented in this thesis is based on the previous approach by Sorzano *et al.* (2009), already one of the best performing methods. However, the accuracy and speed of the previous method improved by introducing new discriminative features, limiting the search space, and changing the classifier to SVM.

Three types of features were extracted from particles to generate a robust to noise feature vector. Two types of features are related to the shape of the particle while the other one gives the intensity information. Having two types of features for shape allow the algorithm to distinguish between particle and objects very similar to particle. Intensity information prevents the APP from picking particles from the carbon parts of micrographs.

Two-step classification strategy reduced false positives as much as possible. While the first classifier may not distinguish between particles and very similar to particles, the second classifier just concentrates on discrimination between particles and objects similar to particle, and therefore reduces the false positive rate.

For the KLH dataset, the role of the second classifier was noticeable since noisy top views present features similar to the ones of the particles. For the adenovirus dataset, intensity information helped the classifier to reject those particles located on carbon areas.

The speed of the algorithm was examined for the three datasets. According to the density of micrographs, it can go from 1 s to 50 s. Most computations are related to feature extraction, especially the shape descriptors. In order to eliminate this bottleneck and improve the speed even more, in implementation particle candidates are divided between different threads. In addition to this, the automatic selection of particles from micrographs is performed in parallel.

## 3.3 A fast and accurate direct Fourier three-dimensional reconstruction with correction for the contrast transfer function

In this thesis, a weighting function for gridding-based DFR in SPA based on the method published for PET by Matej and Lewitt (2001) was proposed to discard the artifact connected with the non-uniformity of samples in the Fourier space. The CTF correction is performed during the reconstruction to correct for the effects of the CTF.

The implementation of the proposed algorithm supports two parallelization frameworks: multi-thread parallelization and multi-computer parallelization. These parallelization frameworks allow our approach to efficiently take advantage of all cores of a node and nodes of a cluster to more quickly produce the final 3-D reconstruction. When multi-computers run out of memory due to the large size of the complex (this happens when multiple MPI processes attempt to run on the same node because memory needs are multiplied by the number of processes), multi-thread implementation can still be employed to take advantage of multiple processing resources.

The proposed method was compared with two well-known methods in the field (the method of Zhang *et al.* (2008) implemented in SPARX and the RELION reconstruction method (Scheres, 2012a)) using data sets (sets of projections) of two test objects: the 70S ribosome, as a small complex, and a DNA-origami object, as a large one. The FSC to the known ground truth was used to compare the reconstruction results of these methods. The proposed method achieved slightly higher FSC values compared with the two other approaches for both the noise-free and noise-contaminated projections. Comparing the execution time of each algorithm for reconstructing the 70S ribosome, it is clear that this method is approximately twice as fast as RELION and SPARX when more than two CPUs are used. For the 3-D DNA-origami object with a large size, the SPARX and RELION reconstruction methods ran out of memory in our machine with 16 GB of memory, while our multi-threaded implementation could improve the reconstruction speed by efficiently using processing resources. This lower memory requirement allows the proposed algorithm to work with larger volumes, which is in demand in the field.

## 3.4 Final discussion

As has been mentioned through the whole thesis, single-particle analysis is now widely used in electron microscopy to resolve the structure of many macromolecules. Advances in computational methods and the emergence of new direct electron detectors are two of the main reasons that make SPA a desirable choice for researchers in electron microscopy. Indeed, the excellent perspective of modern electron microscopy has made it the Nature Methods selection for "Method of the year (2015)".

The work done in this thesis is part of the broader objective of increasing the resolution of final maps of 3D-EM while enhancing the technique's throughput. In thsi way, three algorithms were introduced to improve the accuracy and speed of three important steps of SPA. These steps are: movie alignment, automatic particle picking, and 3-D reconstruction. All these algorithms were implemented and are available in Xmipp (http://xmipp.cnb.csic.es).

The proposed Optical Flow method in this thesis (Abrishami *et al.*, 2015a), can correct for local BIM pattern without requiring to an exhaustive alignment of each particle individually Bai *et al.* (2013). The new approach provided a better result than the (Li *et al.*, 2013) method, which is a widely-used global motion correction approach in the field. Moreover, the implementation of the algorithm on GPU enhanced the speed significantly.

Once the corrected micrographs are available from the previous step, it is time to select particles from these micrographs. As previously mentioned in this thesis, the proposed method for automatic selection of particles (Abrishami *et al.*, 2013) showed a better accuracy compared to the previous method in Xmipp package Sorzano *et al.* (2009), already one of the best performing methods. Apart from the parallelization of the algorithm, limiting the search space, fast feature extraction, and using fast classifier make this method really fast.

Finally, a method has been proposed for direct Fourier 3-D reconstruction with correction for the CTF Abrishami *et al.* (2015b). This method was compared with two other well-known methods: (1) NN direct inversion implemented in SPARX (Zhang *et al.*, 2008) and (2) RELION DFR (Scheres, 2012a). Results showed that the proposed method achieved slightly higher FSC values compared with the two other approaches for both noise-free and noise-contaminated projections, while its efficient parallel implementation allowed this algorithm to run much faster than the other two methods. This fast 3-D reconstruction method is significant since it can improve notably the speed of refinement algorithms.

The improvement in detector technology eliminated many barriers toward atomic resolution. However, a good detector is not everything. New sophisticated image-processing methods also are required to have a successful cryo-EM study. In agreement to this goal, this thesis proposed new algorithms for three different steps of SPA that will push resolution. Additionally, the algorithms and their efficient implementation allow the researcher to resolve the structure of biological objects faster than before.

# Chapter 4

# Conclusions and Future Work

- In this thesis, we presented new methods for alignment of direct detection device movies, automatic particle selection of particles from electron micrographs, and 3-D reconstruction. These improvements in SPA workflow can enhance the resolution of a final 3-D map.

- The proposed method for BIM correction can align movie frames locally at the micrograph level without going to the particle picking step, and thus, we can integrate it with the classical processing workflow in a simple manner. The user can check the quality of movies using the movement provided by the algorithm. Results indicate that this algorithm can align frames at the micrograph level better than existing methods when a fixed-noise pattern is not present.

- We proposed a new method for automatic particle selection from electron micrographs that can learn the particles from the user interactively. The algorithm suggests new particles during the learning process and allows the user to correct its results. The user goes to the automatic step when he is satisfied with the results of the training step. Experimental results for three datasets show that this algorithm is able to select particles accurately even from low-contrast and highly dense micrographs.

- A weighting function for gridding-based direct Fourier 3-D reconstruction was proposed in this thesis. This accurate and fast weighting function corrects the values of frequency samples on a 3-D regular grid to compensate for the uneven distribution of projection images before applying the final inverse FFT to recover the original 3-D object. The CTF correction is performed during the reconstruction to correct for the effects of the CTF.The implementation of the proposed algorithm supports two parallelization frameworks: multi-thread parallelization and multi-computer parallelization.

## 4.1 Future Work

- Develop a method to filter each frame with respect to the total dose ratio. In this way, after alignment, different frequencies of each frame will contribute to the final sum with different weights, an approach that can improve the quality of the final average.

- To use the proposed method for alignment of DDD micrographs in cryo-electron tomography to align the frames at each tilt. Since the SNR is so low, the algorithm will have to be modified so that only frame subaverages of sufficient Signal to Noise Ration will be considered for the calculation of the Optical Flow.

# Bibliography

Abrishami, V., Zaldívar-Peraza, A., de la Rosa-Trevín, J., Vargas, J., Otón, J., Marabini, R., Shkolnisky, Y., Carazo, J., and Sorzano, C. (2013). A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs. *Bioinformatics*.

Abrishami, V., Vargas, J., Li, X., Cheng, Y., Marabini, R., Sorzano, C., and Carazo, J. (2015a). Alignment of direct detection device micrographs using a robust optical flow approach. *Journal of Structural Biology*, **189**(3), 163 – 176.

Abrishami, V., Bilbao-Castro, J., Vargas, J., Marabini, R., Carazo, J., and Sorzano, C. (2015b). A fast iterative convolution weighting approach for gridding-based direct fourier three-dimensional reconstruction with correction for the contrast transfer function. *Ultramicroscopy*, **157**, 79 – 87.

Acharya, K. R. and Lloyd, M. D. (2005). The advantages and limitations of protein crystal structures. *Trends in Pharmacological Sciences*, **26**(1), 10 – 14.

Arbeláez, P., Han, B.-G., Typke, D., Lim, J., Glaeser, R. M., and Malik, J. (2011). Experimental evaluation of support vector machine-based and correlation-based approaches to automatic particle selection. *Journal of Structural Biology*, **175**(3), 319 – 328.

Bai, X.-c., Martin, T. G., Scheres, S. H. W., and Dietz, H. (2012). Cryo-em structure of a 3d dna-origami object. *Proceedings of the National Academy of Sciences*, **109**(49), 20012–20017.

Bai, X.-c., Fernandez, I. S., McMullan, G., and Scheres, S. H. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-em particles. *eLife*, **2**.

Bai, X. C., McMullan, G., and Scheres, S. H. (2015). How cryo-em is revolutionizing structural biology. *Trends in Biochemical Sciences*, **40**(1), 49 – 57.

Banaszak, L. J. (2000). *Foundations of structural biology*. Academic Press.

Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A., and Subramaniam, S. (2014). Structure of $\beta$-galactosidase at 3.2-åresolution obtained by cryo-electron microscopy. *Proceedings of the National Academy of Sciences*, **111**(32), 11709–11714.

Booth, C. and Mooney, P. (2013). Applications of electron-counting direct-detection cameras in high-resolution cryo-electron microscopy. *Microscopy and Analysis*, **27**(6), 13–21.

Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.".

Bragg, W. L. (1913). The structure of some crystals as indicated by their diffraction of x-rays. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **89**(610), 248–277.

Briggs, J. A. (2013). Structural biology in situ – the potential of subtomogram averaging. *Current Opinion in Structural Biology*, **23**(2), 261 – 267. Theory and simulation / Macromolecular assemblies.

Campbell, M., Cheng, A., Brilot, A., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S., Potter, C., Carragher, B., and Grigorieff, N. (2012). Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure*, **20**(11), 1823 – 1828.

Carazo, J., Sorzano, C., Otón, J., Marabini, R., and Vargas, J. (2015). Three-dimensional reconstruction methods in single particle analysis from transmission electron microscopy data. *Archives of Biochemistry and Biophysics*, **581**, 39–48.

Castón, J. R., Belnap, D. M., Steven, A. C., and Trus, B. L. (1999). A strategy for determining the orientations of refractory particles for reconstruction from cryo-electron micrographs with particular reference to round, smooth-surfaced, icosahedral viruses. *Journal of Structural Biology*, **125**(2-3), 209–215.

Chatham, J. C. and Blackband, S. J. (2001). Nuclear magnetic resonance spectroscopy and imaging in animal research. *ILAR Journal*, **42**(3), 189–208.

Chen, J. Z. and Grigorieff, N. (2007). Signature: a single-particle selection system for molecular electron microscopy. *J Struct Biol*, **157**(1), 168–173.

Cheng, Y. (2015). Single-particle cryo-em at crystallographic resolution. *Cell*, **161**(3), 450 – 457.

Cheng, Y., Grigorieff, N., Penczek, P., and Walz, T. (2015). A primer to single-particle cryo-electron microscopy. *Cell*, **161**(3), 438–449.

Chong, C. W., Raveendran, P., and Mukundan, R. (2003). Translation invariants of zernike moments. *Pattern recognition*, **36**, 1765–1773.

Crowther, R. A., Amos, L. A., Finch, J. T., De Rosier, D. J., and Klug, A. (1970). Three dimensional reconstructions of spherical viruses by fourier synthesis from electron micrographs. *Nature*, **226**, 421–425.

de la Rosa-Trevín, J., Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J., and Sorzano, C. (2013). Xmipp 3.0: An improved software suite for image processing in electron microscopy. *Journal of Structural Biology*, **184**(2), 321 – 328.

Deptuch, G., Besson, A., Rehak, P., Szelezniak, M., Wall, J., Winter, M., and Zhu, Y. (2007). Direct electron imaging in electron microscopy with monolithic active pixel sensors. *Ultramicroscopy*, **107**(8), 674 – 684.

Doerr, A. (2016). Single-particle cryo-electron microscopy. *Nat Meth*, **13**(1), 23–23.

Eisenstein, M. (2016). The field that came in from the cold. *Nat Meth*, **13**(1), 19–22.

Elmlund, D. and Elmlund, H. (2012). Simple: Software for ab initio reconstruction of heterogeneous single-particles. *Journal of Structural Biology*, **180**(3), 420 – 427.

Elmlund, D., Davis, R., and Elmlund, H. (2010). Ab initio structure determination from electron microscopic images of single molecules coexisting in different functional states. *Structure*, **18**(7), 777–786.

Elmlund, H., Lundqvist, J., Al-Karadaghi, S., Hansson, M., Hebert, H., and Lindahl, M. (2008). A new cryo-em single-particle ab initio reconstruction method visualizes secondary structure elements in an atp-fueled aaa+ motor. *Journal of Molecular Biology*, **375**(4), 934 – 947.

Elmlund, H., Elmlund, D., and Bengio, S. (2013). Prime: Probabilistic initial 3d model generation for single-particle cryo-electron microscopy. *Structure*, **21**(8), 1299–1306.

Faruqi, A., Cattermole, D., and Raeburn, C. (2003). Direct electron detection methods in electron microscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **513**(1-2), 317 – 321. Proceedings of the 6th International Conference on Position-Sensitive Detectors.

Faruqi, A. R. and Subramaniam, S. (2000). Ccd detectors in high-resolution biological electron microscopy. *Quarterly Reviews of Biophysics*, **33**, 1–27.

Flusser, J., Zitova, B., and Suk, T. (1999). Invariant-based registration of rotated and blurred images. In *Proc. of the Geoscience and Remote Sensing Symposium*, volume 2, pages 1262–1264.

Frank, J. (1975). Averaging of low exposure electron micrographs of non-periodic objects. *Ultramicroscopy*, **1**(2), 159 – 162.

Frank, J. (1990). Classification of macromolecular assemblies studied as 'single particles'. *Quarterly Reviews of Biophysics*, **23**, 281–329.

Frank, J. (1996). *Three dimensional electron microscopy of macromolecular assemblies*. Academic Press, San Diego, CA.

Frank, J. (2006). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*. Oxford Univ. Press, New York, USA.

Frank, J., Verschoor, A., and Boublik, M. (1981). Computer averaging of electron micrographs of 40S ribosomal subunits. *Science*, **214**, 1353–1355.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J. Structural Biology*, **116**, 190–9.

Fu, J., Gao, H., and Frank, J. (2007). Unsupervised classification of single particles by cluster tracking in multi-dimensional space. *Journal of Structural Biology*, **157**(1), 226–239.

Gabriel, E., Fagg, G., Bosilca, G., Angskun, T., Dongarra, J., Squyres, J., Sahay, V., Kambadur, P., Barrett, B., Lumsdaine, A., Castain, R., Daniel, D., Graham, R., and Woodall, T. (2004). Open MPI: Goals, concept, and design of a next generation MPI implementation. In *Proceedings of the 11th European PVM/MPI Users' Group Meeting*, pages 97–104.

Gao, H., Valle, M., Ehrenberg, M., and Frank, J. (2004). Dynamics of ef-g interaction with the ribosome explored by classification of a heterogeneous cryo-em dataset. *J. Structural Biology*, **147**, 283–290.

Glaeser, R. M. (2016). How good can cryo-em become? *Nat Meth*, **13**(1), 28–32.

Grant, T. and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-em using a 2.6 å reconstruction of rotavirus vp6. *eLife*, **4**, e06980.

Grigorieff, N. (2007). Frealign: High-resolution refinement of single particle structures. *J. Structural Biology*, **157**, 117–125.

Hall, R. J. and Patwardhan, A. (2004). A two step approach for semi-automated particle selection from low contrast cryo-electron micrographs. *J. Structural Biology*, **145**, 19–28.

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

Henderson, R. (2013). Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proc Natl Acad Sci U S A*, **110**(45), 18037–18041.

Henderson, R., Chen, S., Chen, J. Z., Grigorieff, N., Passmore, L. A., Ciccarelli, L., Rubinstein, J. L., Crowther, R. A., Stewart, P. L., and Rosenthal, P. B. (2011). Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, **413**(5), 1028 – 1046.

Herman, G. T. and Kalinowski, M. (2008). Classification of heterogeneous electron microscopic projections into homogeneous subsets. *Ultramicroscopy*, **108**, 327–338.

Heymann, J. B., Conway, J. F., and Steven, A. C. (2004). Molecular dynamics of protein complexes from four-dimensional cryo-electron microscopy. *J. Structural Biology*, **147**, 291–301.

Hohn, M., Tang, G., Goodyear, G., Baldwin, P. R., Huang, Z., Penczek, P. A., Yang, C., Glaeser, R. M., Adams, P. D., and Ludtke, S. J. (2007). Sparx, a new environment for cryo-em image processing. *J. Structural Biology*, **157**(1), 47–55.

Huang, Z. and Penczek, P. A. (2004). Application of template matching technique to particle detection in electron micrographs. *J. Structural Biology*, **145**, 29–40.

Keeler, J. (2011). *Understanding NMR spectroscopy*. John Wiley & Sons.

Kuijper, M., van Hoften, G., Janssen, B., Geurink, R., De Carlo, S., Vos, M., van Duinen, G., van Haeringen, A., and Storms, M. (2015). Fei's direct electron detector developments: Embarking on a revolution in cryo-tem. *Journal of Structural Biology*, **192**(2), 179 – 187. Recent Advances in Detector Technologies and Applications for Molecular {TEM}.

Ladd, M. F. C. and Palmer, R. A. (1985). *Structure determination by X-ray crystallography*. Springer.

Langlois, R. and Frank, J. (2011). A clarification of the terms used in comparing semi-automated particle selection algorithms in cryo-em. *J Struct Biol*, **175**(3), 348–352.

Langlois, R., Pallesen, J., and Frank, J. (2011). Reference-free particle selection enhanced with semi-supervised machine learning for cryo-electron microscopy. *J Struct Biol*, **175**(3), 353–361.

Leschziner, A. E. and Nogales, E. (2006). The orthogonal tilt reconstruction method: an approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *J. Structural Biology*, **153**, 284–299.

Leschziner, A. E. and Nogales, E. (2007). Visualizing flexibility at molecular resolution: Analysis of heterogeneity in single-particle electron microscopy reconstructions. *Ann. Rev. Biophys. Biomol. Struct.*, **36**, 43–62.

Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D. A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-em. *Nat Meth*, **10**(6), 584–590.

Liao, H., Hashem, Y., and Frank, J. (2015). Efficient estimation of three-dimensional covariance and its application in the analysis of heterogeneous samples in cryo-electron microscopy. *Structure*, **23**(6), 1129–1137.

Liu, X., Jiang, W., Jakana, J., and Chiu, W. (2007). Averaging tens to hundreds of icosahedral particle images to resolve protein secondary structure elements using a multi-path simulated annealing optimization algorithm. *J. Structural Biology*, **160**(1), 11–27.

Lowe, D. (1999). Object recognition from local scale-invariant features. In *Proc. of the ICCV*, pages 1150–1157.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Intl. J. Computer Vision*, **60**, 91–110.

Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *IJCAI*, volume 81, pages 674–679.

Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999). EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J. Structural Biology*, **128**, 82–97.

Mallick, S. P., Zhu, Y., and Kriegman, D. (2004). Detecting particles in cryo-em micrographs using learned features. *J. Structural Biology*, **145**, 52–62.

Matej, S. and Lewitt, R. (2001). 3d-frp: direct fourier reconstruction with fourier re-projection for fully 3-d pet. **48**(4), 1378–1385.

McMullan, G., Faruqi, A., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300 kev of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy*, **147**, 156 – 163.

Milazzo, A.-C., Leblanc, P., Duttweiler, F., Jin, L., Bouwer, J. C., Peltier, S., Ellisman, M., Bieser, F., Matis, H. S., Wieman, H., Denes, P., Kleinfelder, S., and Xuong, N.-H. (2005). Active pixel sensor array as a detector for electron microscopy. *Ultramicroscopy*, **104**(2), 152 – 159.

Moerner, W. E. (2015). Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy (nobel lecture). *Angewandte Chemie International Edition*, **54**(28), 8067–8093.

Mooney, P. (2007). Optimization of image collection for cellular electron microscopy. In *Cellular Electron Microscopy*, volume 79 of *Methods in Cell Biology*, pages 661 – 719. Academic Press.

Mueller, F. (1994). Pthreads library interface.

Nogales, E. (2016). The development of cryo-em into a mainstream structural biology technique. *Nat Meth*, **13**(1), 24–27.

Ogura, T. and Sato, C. (2004). Automatic particle pickup method using a neural network has high accuracy by applying an initial weight derived from eigenimages: a new reference free method for single-particle analysis. *J. Structural Biology*, **145**, 63–75.

Ogura, T. and Sato, C. (2006). A fully automatic 3d reconstruction method using simulated annealing enables accurate posterioric angular assignment of protein projections. *Journal of Structural Biology*, **156**(3), 371 – 386.

Penczek, P. A. (2010). Chapter one - fundamentals of three-dimensional reconstruction from projections. In G. J. Jensen, editor, *Cryo-EM, Part B: 3-D Reconstruction*, volume 482 of *Methods in Enzymology*, pages 1 – 33. Academic Press.

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

Penczek, P. A., Zhu, J., and Frank, J. (1996). A common-lines based method for determining orientations for N¿3 particle projections simultaneously. *Ultramicroscopy*, **63**, 205–218.

Plaisier, J. R., Koning, R. I., Koerten, H. K., van Heel, M., and Abrahams, J. P. (2004). TYSON: Robust searching, sorting, and selecting of single particles in electron micrographs. *J. Structural Biology*, **145**, 76–83.

Ponce, C. and Singer, A. (2011). Computing steerable principal components of a large set of images and their rotations. *IEEE Trans. Image Processing*, **20**(11), 3051–3062.

Pun, C. M. (2003). Rotation-invariant texture feature for image retrieval. *Computer Vision and Image Understanding*, **89**, 24–43.

Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1987). Three-Dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50s ribosomal subunit of *Escherichia coli*. *J. Microscopy*, **146**, 113–136.

Rohou, A. and Grigorieff, N. (2015). Ctffind4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, **192**(2), 216 – 221. Recent Advances in Detector Technologies and Applications for Molecular {TEM}.

Rose, P. W., Prlic, A., Bi, C., Bluhm, W. F., Christie, C. H., Dutta, S., Green, R. K., Goodsell, D. S., Westbrook, J. D., Woo, J., Young, J., Zardecki, C., Berman, H. M., Bourne, P. E., and Burley, S. K. (2015). The rcsb protein data bank: views of structural biology for basic and applied research and education. *Nucleic Acids Research*, **43**(D1), D345–D356.

Roseman, A. (2003). Particle finding in electron micrographs using a fast local correlation algorithm. *Ultramicroscopy*, **94**, 225–236.

Rubinstein, J. L. and Brubaker, M. A. (2015). Alignment of cryo-em movies of individual particles by optimization of image translations. *Journal of Structural Biology*, **192**(2), 188–195.

Ruiz, T. and Radermacher, M. (2006). Three-dimensional analysis of single particles by electron microscopy. In D. Taatjes and B. Mossman, editors, *Cell Imaging Techniques*, volume 319 of *Methods in Molecular Biology*, pages 403–425. Humana Press.

Ruskin, R. S., Yu, Z., and Grigorieff, N. (2013). Quantitative characterization of electron detectors for transmission electron microscopy. *Journal of Structural Biology*, **184**(3), 385 – 393.

Scheres, S. H. (2012a). Relion: Implementation of a bayesian approach to cryo-em structure determination. *Journal of Structural Biology*, **180**(3), 519 – 530.

Scheres, S. H. (2014). Beam-induced motion correction for sub-megadalton cryo-em particles. *eLife*, **3**.

Scheres, S. H. W. (2012b). A Bayesian view on cryo-EM structure determination. *J. Molecular Biology*, **415**(2), 406–418.

Scheres, S. H. W., Valle, M., Núñez, R., Sorzano, C. O. S., Marabini, R., Herman, G. T., and Carazo, J. M. (2005). Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Molecular Biology*, **348**, 139–149.

Scheres, S. H. W., Gao, H., Valle, M., Herman, G. T., Eggermont, P. P. B., Frank, J., and Carazo, J.-M. (2007). Disentangling conformational states of macromolecules in 3d-em through likelihood optimization. *Nature Methods*, **4**(1), 27–29.

Scheres, S. H. W., Núñez Ramírez, R., Sorzano, C. O. S., Carazo, J. M., and Marabini, R. (2008). Image processing for electron microscopy single-particle analysis using xmipp. *Nature Protocols*, **3**, 977–990.

Schmeisser, M. (2009). *New computational methods for 3D structure determination of macromolecular complexes by single particle cryo-electron microscopy*. Ph.D. thesis, Georg-August University Göttingen.

Schmid, C. and Mohr, R. (1997). Local greyvalue invariants for image retrieval. *IEEE Trans. Pattern Analysis & Machine Intelligence*, **19**, 530–535.

Schmid, M. F. (2011). Chapter 2 - single-particle electron cryotomography (cryoet). In S. J. Ludtke and B. V. Prasad, editors, *Recent Advances in Electron Cryomicroscopy, Part B*, volume 82 of *Advances in Protein Chemistry and Structural Biology*, pages 37 – 65. Academic Press.

Shatsky, M., Hall, R. J., Nogales, E., Malik, J., and Brenner, S. E. (2010). Automated multi-model reconstruction from single-particle electron microscopy data. *J Struct Biol*, **170**(1), 98–108.

Sigworth, F. J. (2004). Classical detection theory and the cryo-em particle selection problem. *J. Structural Biology*, **145**, 111–122.

Sorzano, C., Jonić, S., El-Bez, C., Carazo, J., De Carlo, S., Thévenaz, P., and Unser, M. (2004). A multiresolution approach to orientation assignment in 3d electron microscopy of single particles. *Journal of Structural Biology*, **146**(3), 381 – 392.

Sorzano, C., Jonic, S., Núñez Ramírez, R., Boisset, N., and Carazo, J. (2007). Fast, robust, and accurate determination of transmission electron microscopy contrast transfer function. *Journal of Structural Biology*, **160**(2), 249 – 262.

Sorzano, C., Vargas, J., de la Rosa-Trevín, J., Otón, J., Álvarez Cabrera, A., Abrishami, V., Sesmero, E., Marabini, R., and Carazo, J. (2015a). A statistical approach to the initial volume problem in single particle analysis by electron microscopy. *Journal of Structural Biology*, **189**(3), 213 – 219.

Sorzano, C. O. S., Recarte, E., Alcorlo, M., Bilbao-Castro, J. R., San-Martín, C., Marabini, R., and Carazo, J. M. (2009). Automatic particle selection from electron micrographs using machine learning techniques. *J Struct Biol*, **167**(3), 252–260.

Sorzano, C. O. S., Bilbao-Castro, J. R., Shkolnisky, Y., Alcorlo, M., Melero, R., Caffarena-Fernádez, G., Li, M., Xu, G., Marabini, R., and Carazo, J. M. (2010). A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Structural Biology*, **171**, 197–206.

Sorzano, C. O. S., Alcorlo, M., de la Rosa-Trevín, J. M., Melero, R., Foche, I., Zaldívar-Peraza, A., del Cano, L., Vargas, J., Abrishami, V., Otón, J., Marabini, R., and

Carazo, J. M. (2015b). Cryo-em and the elucidation of new macromolecular structures: Random conical tilt revisited. *Scientific Reports*, **5**, 14290–.

Spahn, C. M. T. and Penczek, P. A. (2009). Exploring conformational modes of macromolecular assemblies by multiparticle cryo-em. *Curr Opin Struct Biol*, **19**(5), 623–631.

Steven, A. and Belnap, D. (2001). Electron microscopy and image processing: An essential tool for structural analysis of macromolecules. In *Current Protocols in Protein Science*, pages –. John Wiley & Sons, Inc.

Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007). Eman2: an extensible image processing suite for electron microscopy. *J. Structural Biology*, **157**, 38–46.

Thuman-Commike, P. A. and Chiu, W. (1997). Improved common line-based icosahedral particle image orientation estimation algorithms. *Ultramicroscopy*, **68**, 231–255.

Tuytelaars, T. and Van Gool, L. (1999). Content-based image retrieval based on local affinely invariant regios. *Visual information and information systems*, pages 493–500.

Valle, M., Sengupta, J., Swami, N. K., Grassucci, R. A., Burkhardt, N., Nierhaus, K. H., Agrawal, R. K., and Frank, J. (2002). Cryo-em reveals an active role for aminoacyl-trna in the accommodation process. *The EMBO Journal*, **21**(13), 3557–3567.

van Heel, M. (1984). Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy*, **13**, 165–183.

van Heel, M. (1986). Classification of very large electron microscopical image data sets. *Optik*, **82**, 114–126.

van Heel, M. (1987). Angular reconstitution: *A posteriori* assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, **21**, 111–124.

van Heel, M. and Frank, J. (1981). Use of multivariate statistical statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, **6**, 187–194.

van Heel, M. and Schatz, M. (2005). Fourier shell correlation threshold criteria. *J. Structural Biology*, **151**, 250–262.

van Heel, M. and Stoffler-Meilicke, M. (1985). Characteristic views of *E. coli* and *B. staerothermophilus* 30s ribosomal subunits in the electron microscope. *EMBO J.*, **4**, 2389–2395.

van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. *J. Structural Biology*, **116**, 17–24.

van Heel, M., Gowen, B., and Matadeen, R. (2000). Single-Particle electron cryo-microscopy: Towards atomic resolution. *Quarterly Review of Biophysics*, **33**, 307–369.

Vargas, J., Otón, J., Marabini, R., Jonic, S., de la Rosa-Trevín, J., Carazo, J., and Sorzano, C. (2013a). Fastdef: Fast defocus and astigmatism estimation for high-throughput transmission electron microscopy. *Journal of Structural Biology*, **181**(2), 136 – 148.

Vargas, J., Abrishami, V., Marabini, R., de la Rosa-Trevín, J., Zaldivar, A., Carazo, J., and Sorzano, C. (2013b). Particle quality assessment and sorting for automatic and semiautomatic particle-picking techniques. *Journal of Structural Biology*, **183**(3), 342 – 353.

Vargas, J., Álvarez Cabrera, A.-L., Marabini, R., Carazo, J. M., and Sorzano, C. O. S. (2014). Efficient initial volume determination from electron microscopy images of single particles. *Bioinformatics*, **30**(20), 2891–2898.

Volkmann, N. (2004). An approach to automated particle picking from electron micrographs based on reduced representation templates. *J. Structural Biology*, **145**, 152–156.

Wang, Q., Matsui, T., Domitrovic, T., Zheng, Y., Doerschuk, P. C., and Johnson, J. E. (2013). Dynamics in cryo em reconstructions visualized with maximum-likelihood derived variance maps. *Journal of Structural Biology*, **181**(3), 195–206.

Wang, Z., Hryc, C. F., Bammes, B., Afonine, P. V., Jakana, J., Chen, D.-H., Liu, X., Baker, M. L., Kao, C., Ludtke, S. J., Schmid, M. F., Adams, P. D., and Chiu, W. (2014). An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nat Commun*, **5**, –.

Wong, H. C., Chen, J., Mouche, F., Rouiller, I., and Bern, M. (2004). Model-based particle picking for cryo-electron microscopy. *J. Structural Biology*, **145**, 157–167.

Wu, S., Armache, J., and Cheng, Y. (2015). Single-particle cryo-em data acquisition by using direct electron detection camera. *Microscopy*.

Xuong, N.-H., Milazzo, A.-C., LeBlanc, P., Duttweiler, F., Bouwer, J., Peltier, S., Ellisman, M., Denes, P., Bieser, F., Matis, H. S., Wieman, H., and Kleinfelder, S. (2004). First use of a high-sensitivity active pixel sensor array as a detector for electron microscopy. In M. M. Blouke, N. Sampat, and R. J. Motta, editors, *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V*, volume 5301 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 242–249.

Yu, Z. and Bajaj, C. (2004). Detecting circular and rectangular particles based on geometric feature detection in electron micrographs. *J. Structural Biology*, **145**, 168–180.

Zhang, K. (2016). Gctf: Real-time {CTF} determination and correction. *Journal of Structural Biology*, **193**(1), 1 – 12.

Zhang, W., Kimmel, M., Spahn, C. M., and Penczek, P. A. (2008). Heterogeneity of large macromolecular complexes revealed by 3d cryo-em variance analysis. *Structure*, **16**(12), 1770 – 1776.

Zhao, J., Brubaker, M. A., and Rubinstein, J. L. (2013). Tmacs: A hybrid template matching and classification system for partially-automated particle selection. *Journal of Structural Biology*, **181**(3), 234 – 242.

Zhu, Y., Carragher, B., Mouche, F., and Potter, C. (2003). Automatic particle detection through efficient hough transforms. *IEEE Trans. Medical Imaging*, **22**(9), 1053–1062.

# Appendix A

# Copies of publications

# Alignment of direct detection device micrographs using a robust Optical Flow approach

CrossMark

Vahid Abrishami [a,1], Javier Vargas [a,1], Xueming Li [b], Yifan Cheng [c], Roberto Marabini [d], Carlos Óscar Sánchez Sorzano [a], José María Carazo [a,*]

[a] Biocomputing Unit, Centro Nacional de Biotecnología-CSIC, C/ Darwin 3, 28049 Madrid, Spain
[b] School of Life Sciences, Tsinghua University, Beijing 100084, China
[c] The Keck Advanced Microscopy Laboratory, Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158-2517, USA
[d] Escuela Politécnica Superior, Universidad Autónoma de Madrid, C/ Francisco Tomás y Valiente, 28049 Cantoblanco, Madrid, Spain

## ABSTRACT

The introduction of direct detection devices in cryo-EM has shown that specimens present beam-induced motion (BIM). Consequently, in this work, we develop a BIM correction method at the image level, resulting in an integrated image in which the in-plane BIM blurring is compensated prior to particle picking. The methodology is based on a robust Optical Flow (OF) approach that can efficiently correct for local movements in a rapid manner. The OF works particularly well if the BIM pattern presents a substantial degree of local movements, which occurs in our data sets for Falcon II data. However, for those cases in which the BIM pattern corresponds to global movements, we have found it advantageous to first run a global motion correction approach and to subsequently apply OF. Additionally, spatial analysis of the Optical Flow allows for quantitative analysis of the BIM pattern. The software that incorporates the new approach is available in XMIPP (http://xmipp.cnb.csic.es).

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

The single-particle analysis (SPA) technique is able to yield three-dimensional (3D) structural information for biological complexes at near atomic resolution by combining many thousands of projection images obtained using transmission electron microscopy (TEM) (Frank, 1996). To achieve high-resolution results in SPA, the characteristics of the image-recording medium are of great importance. Traditionally, electron microscopy images were either recorded on photographic film or with scintillator-based charge-coupled device (CCD) cameras. Each of these two types of detectors offered certain advantages and disadvantages. Film was the preferred recording medium for high-resolution information but required manual scanning of the micrographs, thus limiting automation of processing. Scintillator-coupled CCDs permitted high-throughput image acquisition, allowing full integration between the electron microscope and image-processing software packages. However, scintillator-coupled CCDs record photons, not electrons, and the conversion of electrons to photons comes at the expense of resolution loss at high spatial frequencies (Frank, 2006). This "status quo" limitation has been overcome recently by the new generation of "direct detection devices" (DDDs), first introduced as academic prototypes in 2005 (Milazzo et al., 2005) and offered commercially in 2010 (Jin and Bilhorn, 2010). These sensors detect electrons directly and provide sharper images and higher signal-to-noise ratios (SNRs) (Bammes et al., 2012). Additionally, the fast image acquisition rate of these DDD detectors, ranging from 16 to 400 images per second (Bai et al., 2013; Bammes et al., 2012; Li et al., 2013b), makes it possible to study the behavior of frozen hydrated specimens as a function of electron dose and rate. Therefore, it has become clear that biological specimens in a solid matrix of amorphous ice move during imaging, resulting in "beam- induced motion" (BIM) (Brilot et al., 2012), which is a critical experimental "resolution barrier" in cryo-electron microscopy (Glaeser and Hall, 2011). The subsequent introduction of DDDs has cleared a path to obtaining reconstructions at close-to-atomic resolution for a broad range of specimens. However, the number of reported works that use DDDs is currently not large, and therefore, certain basic questions on BIM characterization remain unanswered. In general, BIM is expected to induce patterns of local movement, although the degree of locality and the

extent of the movement itself are varying widely in different reports (Bai et al., 2013; Booth et al., 2004; Campbell et al., 2012; Li et al., 2013a,b).

In Bai et al. (2013), results for the Falcon II (FEI) are presented using two test samples: the prokaryotic and eukaryotic ribosome. The authors report a movie processing workflow in which an initial 3D map (which disregards BIM) is first obtained and subsequently used to estimate an initial alignment from different frames by applying a statistical refinement method (Scheres, 2012). In their work, the authors report a high degree of local sample movements (see Fig. 1b in Bai et al. (2013)). The algorithm proposed in Bai et al. (2013) produces high-resolution information, but it requires the specimen images to be detected and picked up from the initial video frames, which is a challenge for small particles. However, in Li et al. (2013b), the authors showed results for the K2 Summit (Gatan) direct electron-detection camera and achieved ∼3.6-Å resolution in their 3D map of an archaeal 20S proteasome (∼700 kDa and dihedral D7 point group symmetry) using 10,000 particles. The alignment method used in Li et al. (2013b) consists of a pure in-plane drift correction in which a step of the sub-frame translational alignment is introduced by dividing each frame into a number of sub-frames (normally 3 × 3 sub-frames, each of 2000 × 2000 pixels). This approach is fast if running on GPUs (typically, it takes approximately 10–20 s to process 16 frames of 3876 × 3876 pixels), and at the end, an "average" micrograph is generated for each movie via the summation of all corrected frames. The output is easy to connect with standard processing workflows in use in the field because the DDD "video" is transformed into a "micrograph". The method is certainly appropriate for global sample movements but is not the best option if the sample motion is local. However, by achieving close to atomic resolution, Li et al. (2013b) convincingly showed that for their data, the majority of the motion to be corrected was global, especially if the first few frames were discarded (see Fig. 4b and e in Li et al. (2013b)). At the same time, we note the difference from the results in Bai et al. (2013) and Campbell et al. (2012) in which the authors report a high degree of local sample movements, as previously mentioned. Recently, in Scheres (2014), another method for BIM compensation was introduced that operates over sets of previously picked particles by fitting them to

line trajectories along the frames of the stack. Moreover, Wang et al. (2014) introduced an approach that corrects the motion of boxed particles using the running averages of the frames and calculating the cross-correlations between each frame and the sum of the previously aligned frames, starting at the end of the exposure and working backwards towards the beginning. In this work, we do not enter into a discussion of the nature of BIM itself, but instead, we concentrate on a new image- processing approach that aims to achieve the following objectives: (1) obtain an in-plane "BIM corrected" image that integrates all frames and is computed directly from the stack without performing a particle picking step, and (2) provide fast, objective and quantitative characterization of BIM that accounts for both global and local BIM patterns. Naturally, this integrated image can be used in any of the standard image-processing workflows in cryo-EM as if it were a traditional micrograph.

The proposed method is based on an advanced Optical Flow approach (abbreviated as "OF" in this work) using a pyramidal implementation of the Lucas–Kanade (LK) algorithm (Lucas and Kanade, 1981) with iterative refinement (Bouguet, 2001), which makes the approach quite robust to high levels of noise (Vargas et al., 2014). In essence, OF works best at a local level and is therefore particularly suited for those cases in which the BIM pattern presents a high degree of local movements, as in the Falcon II data sets used in this work. If the BIM pattern is characterized primarily by global movements, OF will have only a minor effect on the final average. Still, even for those latter cases, we have found it advantageous to use the Li et al. (2013b) method combined with OF, by running the Li method followed by the second method to obtain an additional level of refinement and a highly intuitive graphical representation of the total BIM pattern.

## 2. Methods

Our proposed method is based on a regularized Optical Flow approach. The input is a video composed of a set of unaligned low dose frames, and the output is a single image obtained by averaging the resulting motion-corrected frames.
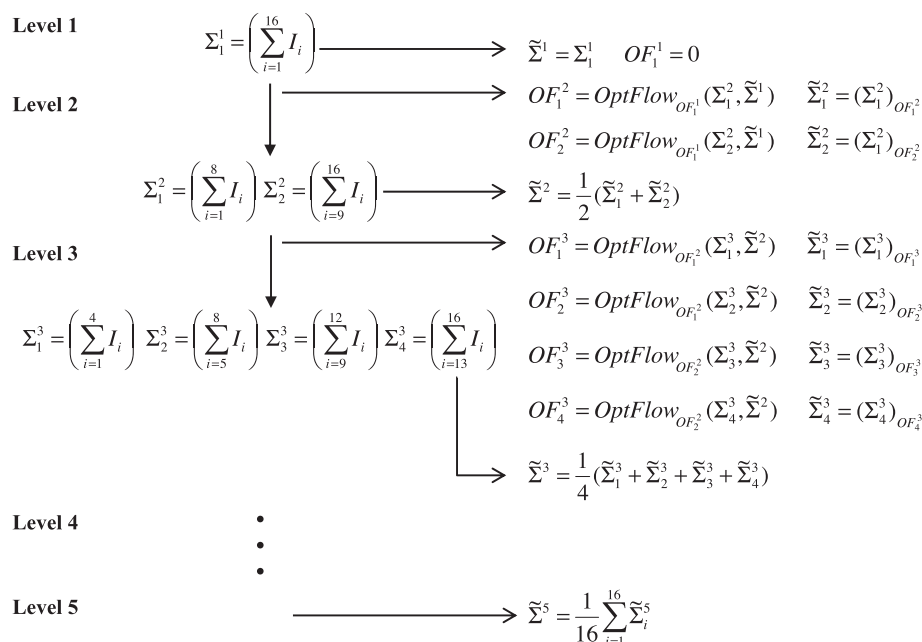


**Fig.1.** Diagram of the proposed Optical Flow method: The required steps for alignment of a video with 16 frames using the proposed Optical Flow method.

## 2.1. Optical Flow approach

Optical Flow is a method used in computer vision to obtain the local motion field between two image frames taken at times $t$ and $t + \Delta t$ (Horn and Schunck, 1981). Let us assume that an arbitrary pixel at time $t$, with coordinates given by $(x, y)$ and intensity $I(x, y, t)$ moves between two consecutive frames to a new position $(x + \Delta x, y + \Delta y)$ with changed intensity $I(x + \Delta x, y + \Delta y, t + \Delta t)$. Assuming that the movement is small, we can expand the intensity map at $t + \Delta t$ as given by:

$$I(x + \Delta x, y + \Delta y, t + \Delta t) \cong I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t \qquad (1)$$

We assume that brightness constancy equation holds, which is given by:

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) \qquad (2)$$

This restriction assumes that the brightness of an object does not significantly change between two consecutive frames, an assumption that is not problematic in our case because the exposure of each frame within the movie is the same. Expression (1) can be rewritten using the brightness constancy equation as:

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \qquad (3)$$

where $u = \frac{\Delta x}{\Delta t}$ and $v = \frac{\Delta y}{\Delta t}$ correspond to the velocity components or local shift components per unit time (assuming that $t$ is only a frame index in our case, but taking into account the temporal information, we can actually retrieve the velocity components). The LK method assumes that the displacements between images are small and approximately constant within a neighborhood of the point $(x_0, y_0)$ under consideration. Thus, the Optical Flow equation can be assumed to apply for all pixels $(\widetilde{x}, \widetilde{y})$ within a window or neighborhood centered at $(x_0, y_0)$, and in that case, the motion (velocity) vector $(u, v)$ at the point $(x_0, y_0)$ must satisfy:

$$I_x(\widetilde{x}, \widetilde{y})u + I_y(\widetilde{x}, \widetilde{y})v = -I_t(\widetilde{x}, \widetilde{y}) \qquad (4)$$

with $I_x, I_y$ and $I_t$ equal to $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ and $\frac{\partial I}{\partial t}$, respectively. Note that for a window or neighborhood composed of $N_w = N_x \times N_y$ pixels with $N_w$ larger than two, Eq. (3) gives an over-determined set of $N_w$ equations. Therefore, we obtain the displacements between images $t$ and $\Delta t$ at $(x_0, y_0)$ by:

$$\begin{pmatrix} I_x(\widetilde{x}_1, \widetilde{y}_1) & I_y(\widetilde{x}_1, \widetilde{y}_1) \\ \dots & \dots \\ I_x(\widetilde{x}_N, \widetilde{y}_N) & I_y(\widetilde{x}_N, \widetilde{y}_N) \end{pmatrix} \begin{pmatrix} u(x_0, y_0) \\ v(x_0, y_0) \end{pmatrix} = \begin{pmatrix} -I_t(\widetilde{x}_1, \widetilde{y}_1) \\ \dots \\ -I_t(\widetilde{x}_N, \widetilde{y}_N) \end{pmatrix} \qquad (5)$$

which is of the form $A\overline{x} = \overline{b}$. We can obtain $\overline{x}$ through the Moore–Penrose pseudo-inverse of $A$, described concisely as:

$$\overline{x} = (A^T A)^{-1}(A^T \overline{b}) \qquad (6)$$

We observe that Eq. (6) is valid only if $A^T A$ is invertible, which is true if $I_x$ and $I_y$ are different from zero in the neighborhood of the point $(x_0, y_0)$, a condition that is satisfied in our case because of the high noise in our frames. From a practical point of view, to avoid instabilities in the inversion of $A^T A$, it is recommended that a small number (approximately 0.1) is added to the diagonal elements (Bouguet, 2000).

## 2.2. Optical Flow with pyramidal decomposition and iterative refinement

The Optical Flow approach presented is the standard LK approach (Lucas and Kanade, 1981). This approach works well if the pixel displacements between the frames are sufficiently small to allow approximation using the first-order Taylor expansion shown in Eq. (3). However, this restriction is severe and can almost never be verified. A possible method for overcoming this limitation is to use the LK approach multiple times, leading to a new Lucas–Kanade Optical Flow approach with iterative refinement. The conceptual workflow of this iterative process is presented in the following for the case of two consecutive images $I_1$ and $I_2$:

(1) Estimate the local shifts $\Delta x$ and $\Delta y$ at each pixel using the LK Optical Flow approach presented in Section 2.1.
(2) Warp one image toward the other using the previously estimated shifts.
(3) Repeat steps (1) and (2) $N_i$ times, where $N_i$ is the number of iterations selected.

Suppose that the LK approach is applied between two consecutive images $I_1$ and $I_2$ and that shifts $\Delta x$ and $\Delta y$ are computed between them. Next, we focus on a certain pixel $(i, j)$, and the corresponding shifts at that pixel are $\Delta x(i, j) = 2$ and $\Delta y(i, j) = 3$ pixels, for example. In this case, applying warping to $I_2(i, j)$ corresponds to changing this intensity value to $I_2(i + \Delta x(i, j), j + \Delta y(i, j)) = I_2(i + 2, j + 3) = I_2^1(i, j)$, where the superindex $I$ denotes one Optical Flow iteration. Obviously, this warping process must be performed for all pixels using the associated shifts $\Delta x$ and $\Delta y$, which is practically performed using cubic interpolation. If the so-obtained values of $\Delta x$ and $\Delta y$ are accurate, after applying the second step (warping), the two images $I_1$ and $I_2^1$ will be approximately equal. Therefore, if the LK algorithm is applied again on these two images (step 3), we obtain $\Delta x \cong 0$ and $\Delta y \cong 0$ at every pixel. Note that following our notation, we must rewrite these shift maps as $\Delta x^1 \cong 0$ and $\Delta y^1 \cong 0$ because we have applied one extra LK iteration. Moreover, we observe that in experimental cases, it is usually not possible to obtain $\Delta x^1 \cong 0$ and $\Delta y^1 \cong 0$, and additional iterations will be necessary. Therefore, after the $k$th iteration, the brightness constancy equation is expressed as follows:

$$I((x + \Delta x^k) + \Delta x^{k+1}, (y + \Delta y^k) + \Delta y^{k+1}, t + \Delta t) = I(x, y, t) \qquad (7)$$

Without loss of generality, we assume that $\Delta t = 1$ and that $\Delta x^k = u^k \Delta t = u^k$ and $\Delta y^k = v^k \Delta t = v^k$. The final displacement vectors computed from this iterative scheme correspond to:

$$\Delta x = \sum_{k=1}^{N_i} \Delta x^k, \quad \Delta y = \sum_{k=1}^{N_i} \Delta y^k \qquad (8)$$

with $N_i$ as the number of iterations used. The LK with iterative refinement allows the algorithm to provide good results in cases in which the movements are not sufficiently small to directly use the Taylor expression shown in Eq. (1).

An important consideration for the LK Optical Flow algorithm with iterative refinement presented is that the window or neighborhood size must be specified. As intuitively expected, the choice of this window size introduces a trade-off between accuracy and robustness of the approach. In this work, accuracy relates to the local sub-pixel systematic error that is achieved, and robustness refers to the insensitivity of the approach with respect to disturbances, i.e., noise and outliers. In a nutshell, large window sizes provide more robustness but less accuracy. Additionally, to handle large motions between the images, it is necessary to use large integration windows. To solve this trade-off, Bouguet (2000) presented a pyramidal implementation of the iterative LK algorithm introduced above that substantially reduces the dependency on the window size. This approach is based on performing an iterative LK Optical Flow approach recursively over different resolution representations of the input images, known as pyramids, which are

obtained by successively down-sampling the images. We rank these image pyramid representations from coarse to fine resolutions (note that the images with highest resolution are the input images) and subsequently perform the iterative LK approach with a fixed window size, beginning with the images with lowest resolution and moving towards those with highest resolution. This approach first provides a coarse and rather robust estimation of the motion (velocity) vectors $(u_0, v_0)$ from the low-resolution pyramids. These vector fields are used as an initial guess for the next Optical Flow estimation using the next higher-resolution pyramidal representation images. This process is followed recursively until the highest-resolution images (input images) are used. Observe that this pyramidal implementation of the LK Optical Flow algorithm with iterative refinement provides a clear advantage because the integration window size can remain fixed and small, thus ensuring high accuracy and robustness. In Section 3.3, we present the details of the strategy we use in this work to select the window size for each case.

### 2.3. Proposed Optical Flow approach for DDD frame alignment

In cryo-EM, the types of images in which we are interested are characterized by notably low SNRs. As a consequence, we must add additional robustness against noise to the standard LK Optical Flow with the pyramidal decomposition and iterative refinement (LKPDI) presented in Section 2.2 (Bouguet, 2000). To this end, we have designed a sequential processing workflow, which attempts to maximize the signal and minimize the noise at each step of Optical Flow estimation.

Because the movie frame images are characterized by notably low SNRs, it is not possible to directly obtain reliable shift maps between consecutive frames. To improve the Optical Flow results, we follow a coarse-to-fine alignment process. To this end, we first obtain the average of all frames within the movie, which is initially established as the first estimation of the corrected movie average image. Subsequently, the frames are divided into two sets by grouping frames that are consecutive in time, and the average image of each set is computed. Observe that the first estimation of the corrected movie average as well as the obtained average images of each set have high SNRs at low resolution. As a consequence, if we align these average images with respect to the available corrected movie average, we obtain robust but coarse shift estimations. Using these shifts, a new corrected average can be estimated, and each of the frame sets can be divided in two again. Therefore, we once again perform the same process between the obtained average images of each group and the available corrected movie average using the previously obtained shifts as initial guesses. As a result, we obtain improved shift estimation in each refinement process. Therefore, the proposed alignment approach consists of a coarse-to-fine estimation of the shifts, which improves the robustness of the method with respect to noise in the images.

In the following, we introduce a more detailed description of the proposed method from a mathematical point of view. We first introduce the notation and subsequently describe the algorithm. Let $I_i$ be the $i$th frame in the stack, $\Sigma_j^k$ be the unaligned average of the $j$th subgroup of the frames in the $k$th iteration (how these subsets are formed will be described later), $\widetilde{\Sigma}^k$ be the aligned average in the $k$th iteration obtained using all images and, finally, let $\widetilde{\Sigma}_j^k$ be the aligned average of the images belonging to the subset $j$ in iteration $k$. Next, the Optical Flow algorithm is described as follows:

1. For $k = 1$, the algorithm computes the average of all $N$ unaligned frames $\Sigma_1^1 = \left( \sum_{i=1}^{N} I_i \right)$ and initializes the refined average

estimation as $\widetilde{\Sigma}_1^1 = \Sigma_1^1$. Next, $k$ is updated as $k = k + 1$, and for $k > 1$, the following iterative process is executed.

2. Images are split into $2^{k-1}$ subgroups, and the unaligned average of each group is computed. Therefore, there are $2^{k-1}$ averages with $N/2^{k-1}$ images in each average such that $\Sigma_j^k = \left( \sum_{i=(j-1)s^k+1}^{js^k} I_i \right)$, where $s^k = N/2^{k-1}$. Note that for $k = 2$, the set is divided into two groups, and for each group, we obtain the unaligned averages $\Sigma_1^2$ and $\Sigma_2^2$.

3. Next, we compute the LKPDI optical flows between the new set of unaligned averages $\Sigma_j^k$ and the corresponding set of aligned images computed in the previous iteration $\widetilde{\Sigma}^{k-1}$. The result, which is a map of shift vectors centered at each pixel that aligns $\Sigma_j^k$ with $\widetilde{\Sigma}^{k-1}$, is known as $OF_j^k$. The estimation of $OF_j^k$ requires an initial guess; our algorithm uses $OF_{j/2+j \mod 2}^{k-1}$ for $k > 2$ and a zero-valued matrix of displacement or shift vectors for $k = 2$.

4. Using the obtained shift vector matrix ($OF_j^k$), the averages $\Sigma_j^k$ are aligned; we refer to these aligned averages as $\widetilde{\Sigma}_j^k$. A new corrected average is computed as $\widetilde{\Sigma}^k = \sum_{j=1}^{2^k} (\widetilde{\Sigma}_j^k)/2^k$. If the groups contain more than one frame, Steps 2 to 4 are repeated.

Fig. 1 shows a diagram of the proposed approach for a movie composed of 16 frames. This procedure is applicable to movies with a number of images that is not a power of 2 by simply readjusting the appropriate indices. Note that this alignment strategy improves the robustness of the approach against noise, and at the same time, it reduces other potential issues, such as the validity of the brightness constancy equation due to radiation damage or changes in brightness resulting from mass loss during exposure.

## 3. Results

This section presents a collection of tests using both simulated and experimental data. To obtain the optical flow, we used the OpenCV (Bradski, 2000) (Open Source Computer Vision Library) implementation of the Lucas–Kanade dense Optical Flow (in which the Optical Flow is computed for all pixels), which provides both CPU and GPU execution modes. For the LK algorithm, we set the number of iterations to 10 and the number of scales to 6 as constant values for all datasets and also separately optimized the window size for each dataset (the details of how to obtain the window size are described in Section 3.3). The algorithm was executed on a single GeForce GTX 690 GPU card with two Kepler GPUs as well as in CPU mode. Typically, the GPU-based execution time is only approximately three times faster than in CPU execution mode, clearly showing that I/O has a large detrimental impact on the GPU performance.

### 3.1. Simulations

We used computer simulations to test a subset of the basic features of the newly proposed algorithm. In particular, in this simplified system, we aimed to study two different situations: (1) global versus local movements, i.e., the ability of the method to compensate for both global and local movements, and (2) random versus drift-like movements, i.e., the ability to recover information when the movement in between frames is random and when it displays a certain directionality (a drift-like effect).

In all cases, we compared the results obtained by OF with the results provided by the Li et al. (2013b) alignment approach for two different cases. Note that because the output of our method is an integrated image, we cannot easily compare the proposed method with that of Bai et al. (2013), for which the output is a

3D reconstruction. In the first case, we applied the Li et al. (2013b) alignment method using the entire collection of image frames (Li Global), and in the second, we divided the images (i.e., the frames) into nine overlapping sub-images and aligned each of them independently, as suggested in Li et al. (2013b), to render the method more robust with respect to local sample movements (Li Local). For the simulation, we used one frame of a movie provided in Bai et al. (2013) with a size of $4096 \times 4096$ pixels that were cropped to $3876 \times 3876$ pixels (because of the presence of artifacts at the borders). Depending on the type of the simulation, we applied global or local shifts (random or drift-like) to this frame, and added a random Gaussian noise to each simulated frame, thus creating a series of movies. In all cases, the shifts were randomly drawn from a uniform distribution with zero mean and a variance of 2 pixels. The SNR of each frame in the different simulated movies is 0.5.

### 3.1.1. Global versus local movements

*3.1.1.1. Global movements.* In this case, we added the same movement to each pixel of the simulated frames. Thus, we composed a simulated movie formed by sixteen unaligned frames. The different applied shifts are shown in Table S1 with reference to the shifts of the tenth frame, i.e., $\Delta \tilde{x}_i = (\Delta x_i - \Delta x_{10})$ and $\Delta \tilde{y}_i = (\Delta y_i - \Delta y_{10})$. Note that for a stack of size 16, Li's implementation considers the tenth frame as the reference by default and provides all other shifts with respect to this frame. For this movie, we applied OF using a window size of 25 pixels as an input parameter. In turn, the alignment approach presented in Li et al. (2013b) was executed using the entire image and also dividing the frames into nine overlapping sub-images of size $2048 \times 2048$ pixels, as suggested in Li et al. (2013b). The resulting shifts (mean shifts) are quite similar to the theoretical shifts, as shown in Table S1. An analysis of the root-mean-square error (*rms*) between the computed and the theoretical displacements provides average *rms* values of 0.0048 and 0.011 for Li Global, 0.0027 and 0.011 for Li Local, and 0.032 and 0.042 for OF, respectively, in pixel units. As expected, in this case, Li's method provides slightly better results than OF because it is essentially a global alignment approach, whereas OF is a local approach. However, we observe that even in this unfavorable case, the proposed OF approach provides notably accurate and precise results (on the order of $10^{-2}$ pixels), indicating that the two methods behave the same in practical terms.

*3.1.1.2. Local movements.* For this case, we used the same starting image as in the previous case but introduced local shifts between frames. For simplicity, the local shifts were introduced only in *x* as given by the expression:

$$\Delta x_i = \begin{cases} 0, & i = 1, \\ 4 \sin\left(2\pi\left(\frac{2x}{N_C} + \frac{i}{N_R}\right)\right), & i \neq 1 \end{cases} \quad (9)$$

where $\Delta x_i$ are the local shifts, $(N_C, N_R)$ are the number of rows and columns, *i* is the number of the frame within the movie, and *N* is the number of frames that compose the movie. The reason for this simulation with only local movements in the *x*-axis was to provide a precision estimation for the alignment approaches (how well the shifts are recovered when the theoretical shifts were exactly zero), thus avoiding such synthetic effects as quantization and interpolation while maintaining the added noise.

Fig. 2(A) displays a graphical plot of the theoretical shifts introduced by Eq. (9), for $i = 1$, with reference to the shift of the tenth frame given by $\Delta \tilde{x}_1 = (\Delta x_1 - \Delta x_{10})$. In the same manner, Fig. 2B presents the shifts $\Delta \tilde{x}_i$ recovered by OF, and Fig. 2C displays the results from Li Local. In this simple graphical representation, it is clear that 2B is rather similar to 2A, but 2C is markedly different. In more quantitative terms, we computed the mean of the *rms* errors between the theoretical and obtained shifts, obtaining 5.26 pixels for Global Li, 4.04 pixels for Local Li; 0.1 pixels for OF along *x*; and 0.19 pixels for Global Li, 0.20 pixels for Local Li, and 0.02 pixels for OF along *y*. As observed from this analysis and Fig. 2, in this local-shift case, the results obtained from OF are considerably better than those provided by the Li approach, as expected by the relationship between the two methods. Additionally, we observe that the OF results along the *y*-axis are significantly better than the results along the *x*-axis because of quantization and interpolation effects and because the OF approach assumes the shifts are the same within a window, which is obviously true for the shifts along the *y*-axis but not true along the *x*-axis.

### 3.1.2. Random versus drift-like movements

The goal of this test was to evaluate the performance of the Optical Flow approach in those cases in which the movement in between frames was random compared with the cases in which the movement had a certain level of directionality along the stack. We were particularly interested in this question because both the OF and Bai et al. (2013) method use a first step of averaging the sets of frames without correction, and the quality of this uncorrected average would be quite different if the frames moved randomly than if the frame movement had a certain directionality. In the former case, the uncorrected average would "only" appear as an isotropically blurred version of the ideal image, a rather intuitive type of degradation that is easy to address computationally, whereas in the latter case, the result would be a rather artifactual image that might preclude further steps of processing. For simplicity, we considered global movements (i.e., in which the entire frame moves in the same direction) such that the only variable remaining was the directionality of the movement along the stack. We studied three cases: random movements, movements with restricted directionality, and drift-like movements.

The first case addressing random movements was discussed under Section 3.1.1.1, and clearly, our newly proposed method was shown to perform quite well. The second case (movements with restricted directionality) was simulated by applying different shifts between frames, where the directions of the shifts were
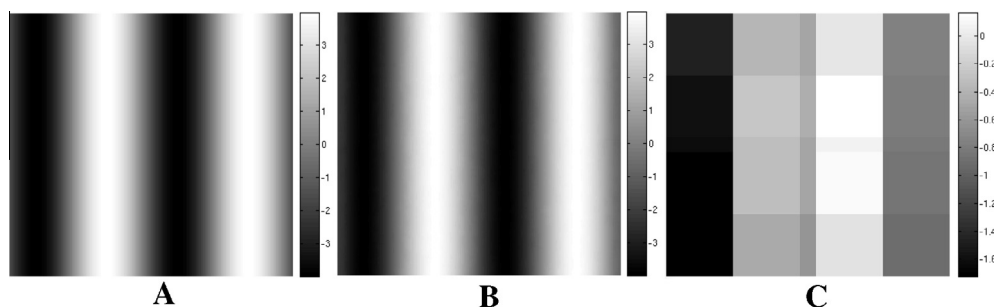


**Fig.2.** Recovered simulated local shifts using optical flow: graphical representation of theoretical and recovered local shifts along *x* referred to the shift of the tenth frame $(\Delta x_1 - \Delta x_{10})$. Theoretical shifts in *x* (A). Shifts obtained by OF (B) and Li Local (C), respectively.

drawn from a Gaussian distribution with a zero mean and a standard deviation of 0.4 radians or 23 degrees to generate a stack of 16 frames. Note that the shifts along the stack are cumulative; each shift is the sum of the previous shifts and the current shift. The Optical Flow method with different window sizes was applied to this stack. For a window size of 50 pixels, we obtained the best *rms* on the *x*- and *y*-axes of 0.2518 and 0.0146 pixels, respectively. Note that although we were still able to recover the shifts with good accuracy, the results were worse than in the previous case. Finally, in the third case, we considered the situation in which all shifts accumulated in exactly the same direction, and in this case, we clearly were not able to recover the movements, even if large window sizes were applied (note that the Li method was not affected by this drift-like effect and performed well in this case).

### 3.2. Experimental results

In this section, we present the results obtained by applying our proposed algorithm to different datasets, each taken from different specimen samples imaged on different microscopes and cameras, to assess the robustness of the proposed method under experimental conditions.

A key issue in comparing the performance of different algorithms is the use of a good comparison metric. This metric should provide objective results while ideally offering simple and rapid calculation such that multiple tests can be performed. Thus far, most of the published works on DDD have concentrated on exploring the best resolution achievable from a certain dataset when processed in a number of different ways. Indeed, this ideal describes the "ultimate" comparison metric, but it is clearly neither simple nor fast to calculate. Additionally, it is difficult to distinguish whether an improvement originates from the movie processing procedure or the 3D reconstruction iterative process. As a method of mitigating this problem, we analyzed the effect of misalignment (initial or "residual" misalignment) on the power spectral density (PSD) of the final "corrected" average. We present this analysis in the [Appendix]. An interesting result is that the shapes of the CTF envelope and of the spectral signal-to-noise ratio (SSNR) represent notably good performance metrics. Indeed, residual errors in the alignment correction translate into a dampening of these metrics such that we can judge the quality of the correction (how small the residual misalignment was) by comparing them and looking for the case in which this dampening is most reduced. Assuming axial astigmatism to be minimal in these high-quality datasets, we calculated frequency radial averages, resulting in 1D curves that can be easily analyzed. In all cases, we compared the results of OF with those of both the Local and Global Li approaches in terms of alignment accuracy and execution speed.

### 3.2.1. Ribosome dataset

In this case, we used two different movie stacks previously reported in [Bai et al. (2013)] and made publicly available at ([http://www.ebi.ac.uk/~ardan/aspera/em-aspera-demo.html]). The names of the movie stacks as deposited are "15_movie_gc_window.mrcs" and "205_movie_gc_window.mrcs", and both correspond to images of the *Saccharomyces cerevisiae* 80S ribosome. The images were recorded on a FEI Falcon II direct electron detector with nominal magnification of 59,000× at a working voltage of 300 kV, resulting in a pixel size of 1.77 Å (the size of the ribosome at this sampling rate is approximately 150 pixels).

We processed these movies with both the Li Global and Li Local approaches as well as with the proposed OF method using a window size of 150 pixels. The results are presented in [Fig. 3]A and B, which show the 1D profiles of the CTF envelope and SSNR functions obtained for "205_movie_gc_window.mrcs" and "15_movie_gc_window.mrcs", respectively. As expected, both the Li and OF

approaches improve the results compared with the direct average of the unaligned frames. However, OF clearly provides better results (higher envelope and SSNR) than the Li approach for low and high frequencies. In [Fig. 3]C, we present a comparison between the periodogram (i.e., the modulus-squared of the discrete Fourier transform) of the initial uncorrected average (left) and that obtained by the application of the Optical Flow approach (right).

As a new piece of information, in [Fig. 4], we present the vector fields obtained by OF corresponding to the movements that each pixel undergoes between frame 1 and frame 2 for the two test cases (see [Movie S1] for a representation of the vector fields for all consecutive frames). In this figure, the information of the shifts is coded using the Hue, Saturation and Value (HSV) of the color. As the hue value varies from 0 to 1.0, the corresponding colors vary from red, through yellow, green, cyan, blue, and magenta and back to red such that red values actually occur at both 0 and 1.0. As the saturation varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value or brightness varies from 0 to 1.0, the corresponding colors become increasingly brighter. In our case, the hue gives information on the direction of the movement, and the saturation gives information for the magnitude. As observed from these movies, the orientation map is locally smooth, and the displacement pattern is clearly non-global, as originally reported ([Bai et al., 2013]). A simple and intuitive analysis of the movements among all of the movie frames is presented in [Fig. 5] (as applied to the results shown in [Movie S1]). The information on the pixel movement is coded both in Cartesian and polar representations in which a number of circles are shown. The position of the center of each circle corresponds to the mean displacement between two consecutive frames calculated by averaging the pixel's displacements in *x* and *y* over all of the pixels within the frame as provided by the corresponding vector fields. In turn, the area of the circle represents the compound standard deviation in *x* and *y*, i.e., $\sqrt{\sigma_x^2 + \sigma_y^2}$, where $\sigma_x^2$ and $\sigma_y^2$ are the variances in the *x*- and *y*-axes, respectively. Note that for the sake of this representation, we consider that the shifts along the *x*- and *y*-axes behave as random variables that are independent such that we can obtain the standard deviation of the sum as the square root of sum of the variances. Consequently, the position of the center of the circle represents the "global" (mean) displacement, and its area represents the pixel-to-pixel differences with respect to this mean displacement. In other words, the area is related to the amount of local versus global movements between two frames. It is clear from [Fig. 5]A and B that the global movements between frames are within a fraction of a pixel and that there is a certain directionality in the between-frame movement. We comment on this observation and subsequent plots later in this manuscript together with the results for additional cases presented.

The execution time for a stack of this dataset is 123 s using OF, 50 s using the Li Local approach, and 20 s using the Li Global method.

### 3.2.2. Influenza virus RNP dataset

In this experiment, we used a stack of influenza virus ribonucleoprotein (RNP) (unpublished data) consisting of 70 frames collected with a FEI Falcon II direct electron detector at the MRC with a pixel size of 2.26 Å. The total exposure time was 4 s.

As in the previous experiment, we obtained the integrated image using the simple average, Li Local, Li Global, and OF. A comparison of these methods using the CTF envelope as well as the SSNR is shown in [Fig. 6]A. As shown in this figure, OF presents a higher accuracy compared with that of the other methods. In [Fig. 6]B, the uncorrected (left) and corrected (right) periodograms are shown together, and in [Fig. 6]C, an analysis of inter-frame
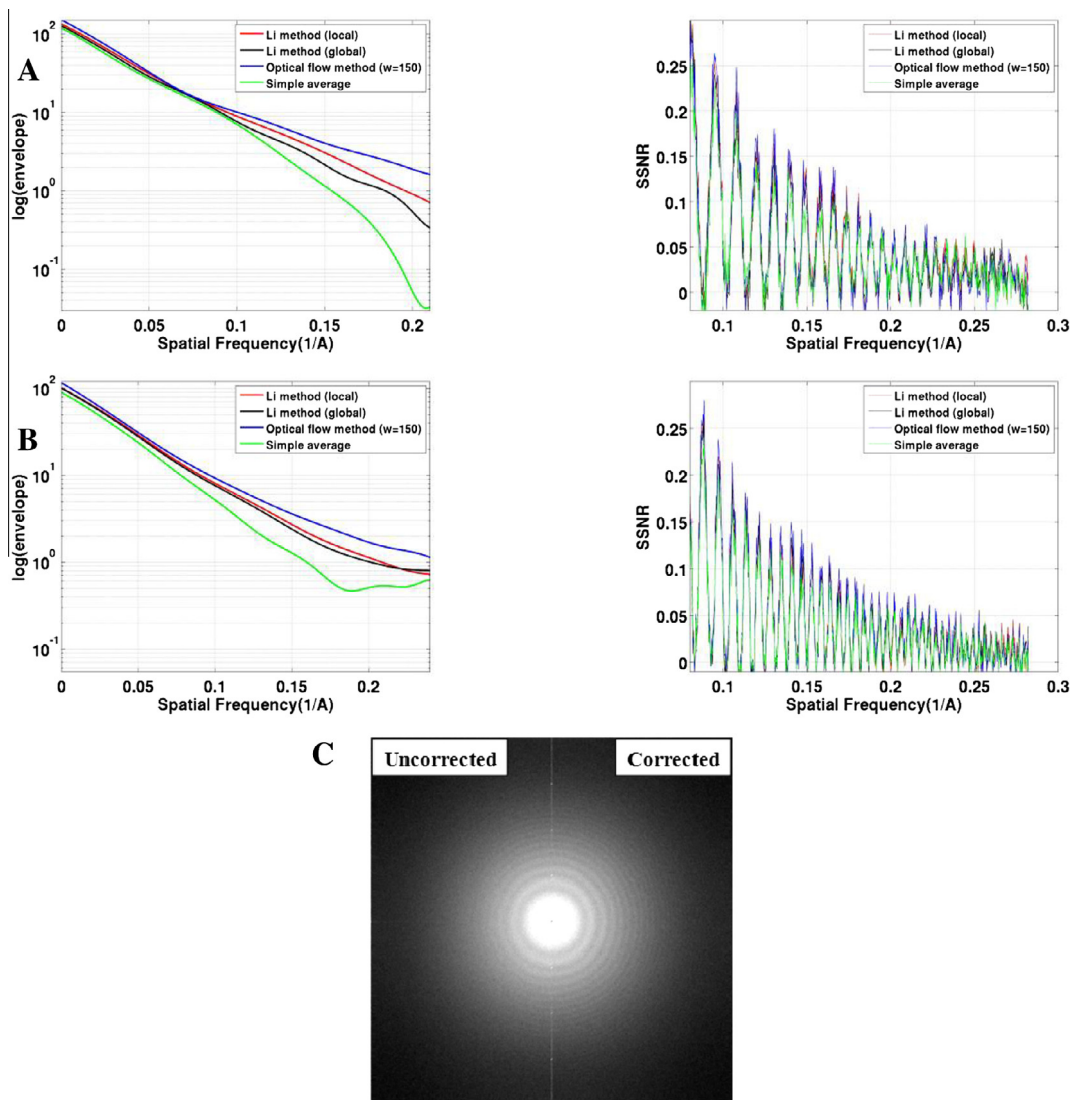
**Fig.3.** Comparison of the alignment results of OF and Li method for two videos of Bai et al. (2013): envelopes of the contrast transfer function (in logarithmic scale) as well as SSNR functions obtained from the unaligned average and the aligned averages using Li Global, Li Local and Optical Flow for the movies "205_movie_gc_window.mrcs" (A) and "15_movie_gc_window.mrcs" (B). Comparison of the periodograms (i.e., the modulus-squared of the discrete Fourier transform) of the uncorrected (left) and Optical Flow aligned average image (right) corresponding to "15_movie_gc_window.mrcs" (C).
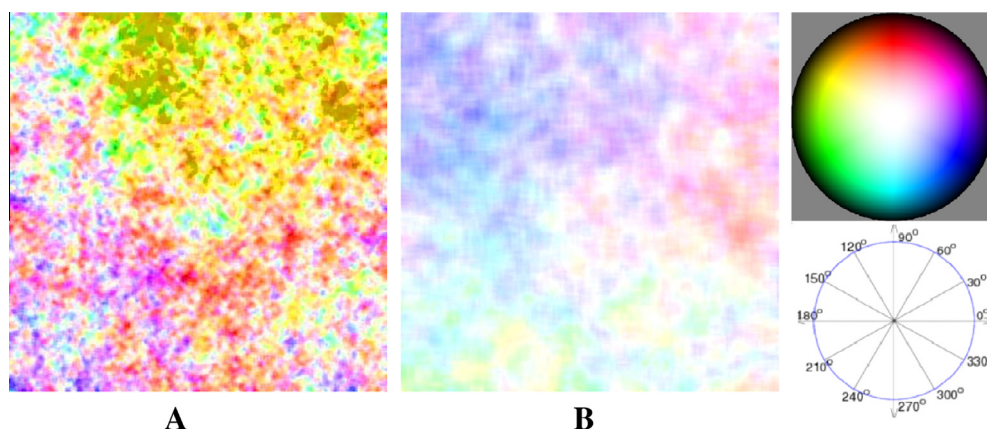


**Fig.4.** Vector fields for two videos of Bai et al. (2013): vector fields obtained by the Optical Flow approach corresponding to the movements for each pixel between frame 1 and frame 2 for the stacks "205_movie_gc_window.mrcs" (A) and "15_movie_gc_window.mrcs" (B) (see also Movie S1). The information of the shifts is coded using the Hue, Saturation and Value (HSV) of the color. The hue gives information on the direction of the movement, and the saturation gives information for the magnitude.
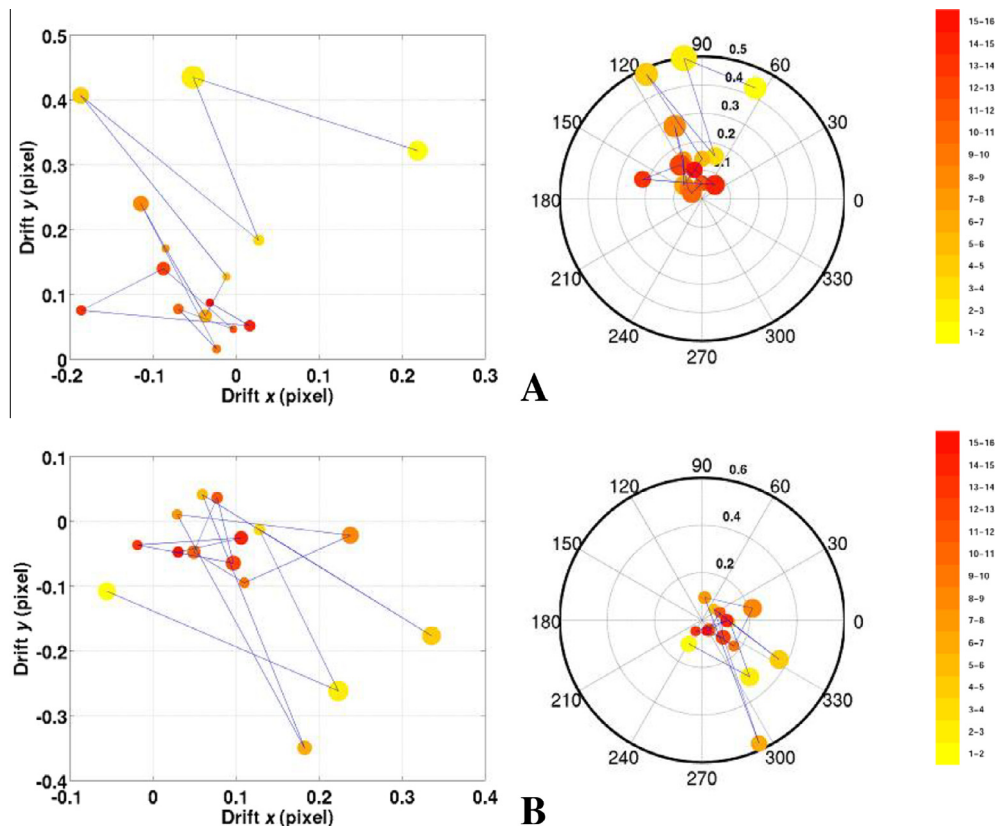
**Fig.5.** Movement analysis for two videos of Bai et al. (2013): analysis of the movements between consecutive frames of "205_movie_gc_window.mrcs" (A) and "15_movie_gc_window.mrcs" (B). In these Cartesian and polar representations, the displacements correspond to frame shifts as measured in pixels such that the center of each circle corresponds to the mean displacement between two consecutive frames, and the area of each point represents the compound standard deviation in the $x$- and $y$-axes.

movements is presented. Fig. 6C clearly shows that in this stack, the movements between frames are virtually random.

The total processing time for this stack is 5 min and 17 s for OF, 1 min and 25 s for Li Global, and 4 min and 17 s for Local Li.

*3.2.3. Archeal 20S proteasome*

In this work, we also tested the OF procedure on image stacks recorded using a K2 Summit camera operated in super-resolution counting mode. Images were recorded from frozen-hydrated archaeal 20S proteasome, as described previously (Li et al., 2013a). Briefly, each image stack consists of 24 sub-frames with a pixel size of 1.22 Å. The total exposure time is 5 s. Because it is an electron counting camera, the K2 Summit requires that images are recorded with a low dose rate in the camera, resulting in a long exposure time (Li et al., 2013a). Different from the previously presented images recorded with Falcon II, images recorded with K2 Summit in this work often contain drift-like motion in addition to charge-induced motion.

As in all previous cases, the stacks were directly processed with the newly proposed Optical Flow approach. However, in this case, we obtained results quite similar to those presented in Section 3.1.2 for the case of drift-like movements. In essence, the shape of the envelope function did not improve despite the use of rather large windows. At this stage, we explored the use of the Li et al. (2013b) method as a pre-processing step prior to OF (note that Li method is not perturbed by the possible accumulation of shifts in similar directions, as shown in Section 1.3.2). Furthermore, we incorporated into OF the information on inter-frame global movements provided by the Li et al. (2013b) method, resulting in a combined new method that provides a clear and intuitive tool for representing the complete BIM pattern from the vector fields.

We extended this analysis to the processing of a large set of K2 frame stacks and consistently found that the corresponding BIM pattern had a substantially higher level of global versus local movements. Naturally, for those cases in which the degree of local BIM pattern was higher, the improvements due to the application of OF were more significant. Still, in all cases, the application of OF after the Li method provided a quantitative understanding of the BIM pattern of the stack via analysis of the corresponding vector fields, as presented in previous plots in polar and Cartesian coordinates.

Following the previous reasoning, we first applied the Li Global method to remove global movements, followed by OF to correct for possible remaining local movements. The resulting shapes of the CTF envelope and SSNR functions are shown in Fig. 7A for a window size of 150 pixels, and clearly, the application of OF after Li Global improved the results only slightly. A comparison between the periodograms corresponding to the uncorrected (left) and fully corrected (right) stack average is shown in Fig. 7B. Furthermore, this dataset illustrates the power of the vector fields as a new and convenient diagnostic tool for analysis of BIM patterns. Additionally, in Fig. 7C, we show Cartesian and polar plots in which it is immediately noticeable that there is a strong directionality of the frame-to-frame global movement (the points are concentrated in a small slice in polar coordinates) and that the amount of this movement is larger than in any of the previous cases. Additionally, the area of the circles is notably small, even considering the change of scale in the radial variable required to properly represent larger shifts, further indicating that the ratio of global versus local movements was higher in this case than in previous ones.

Application of Li Global followed by Optical Flow on a typical stack of this dataset required 5 and a half minutes of processing.
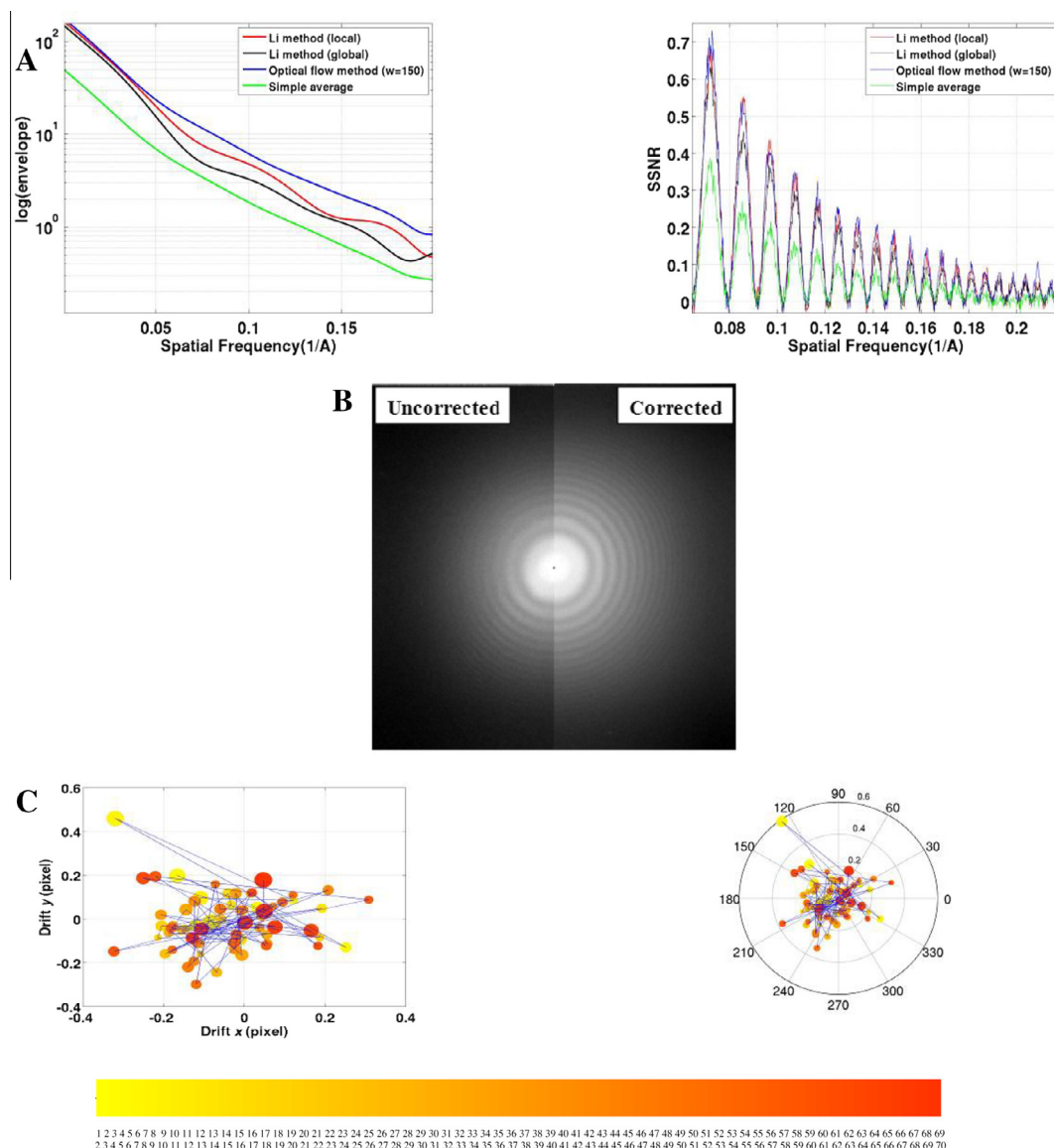
**Fig.6.** Comparison and analysis of the alignment result for the influenza virus RNP dataset: comparison of the envelopes of the CTF (in logarithmic scale) as well as the SSNR functions corresponding to the average image obtained from the summation of unaligned frames and aligned frames using Li Local, Li Global and Optical Flow with a window size of 150 pixels (A). The periodogram for the uncorrected (left) and corrected averages (right) (B), and the analysis of the movements between consecutive frames (C).

### 3.2.4. Specimen images presenting fixed-noise patterns

A fixed-noise pattern is present in a non-negligible proportion of the stacks (Li et al., 2013b; Shigematsu and Sigworth, 2013), making pertinent the analysis of the behavior of the newly proposed Optical Flow approach under these circumstances. According to first principles, we expect that a pattern present in the same manner in all frames would pose problems for the Optical Flow approach because it could drive the local movement detection towards the detection of no movement at all. We envisioned two possible solutions: (1) perform a type of filtering to attenuate the fixed-noise pattern or (2) first use Li Global to perform a coarse frame alignment that would "unalign" the fixed-noise pattern between the frames and subsequently apply Optical Flow. We concentrated on the second approach and obtained good results. Fig. 8A presents a comparison of the CTF envelopes and SSNR using Li Global, Li Local and Li Global + OF (window size = 150 pixels), and Fig. 8B shows the periodogram of an uncorrected stack average that presents a clear fixed-noise pattern that is removed by the combined procedure Li Global + OF. Finally, Fig. 8C displays an

analysis of the movements between frames of this stack. Note that in cases with strong fixed pattern noise but without a significant global motion in each stack, our approach will not provide good results.

### 3.3. Selection of the window size parameter

All methods contain parameters, and their selection generally has important effects on the method's behavior. For our Optical Flow approach, and as initially introduced in the mathematical presentation, the most important parameter is the window size. Indeed, the window size represents the (square) neighborhood used to calculate the Optical Flow at that point, and movements are expected to be constant within that window.

The practical approach to calculating the window size for each set of videos has been to first run the algorithm on several videos using different window sizes ranging from 25 to 500 pixels, and second, to choose the window size that provides the highest integral over the 1D profile of the CTF envelope function. As an example, we consid-
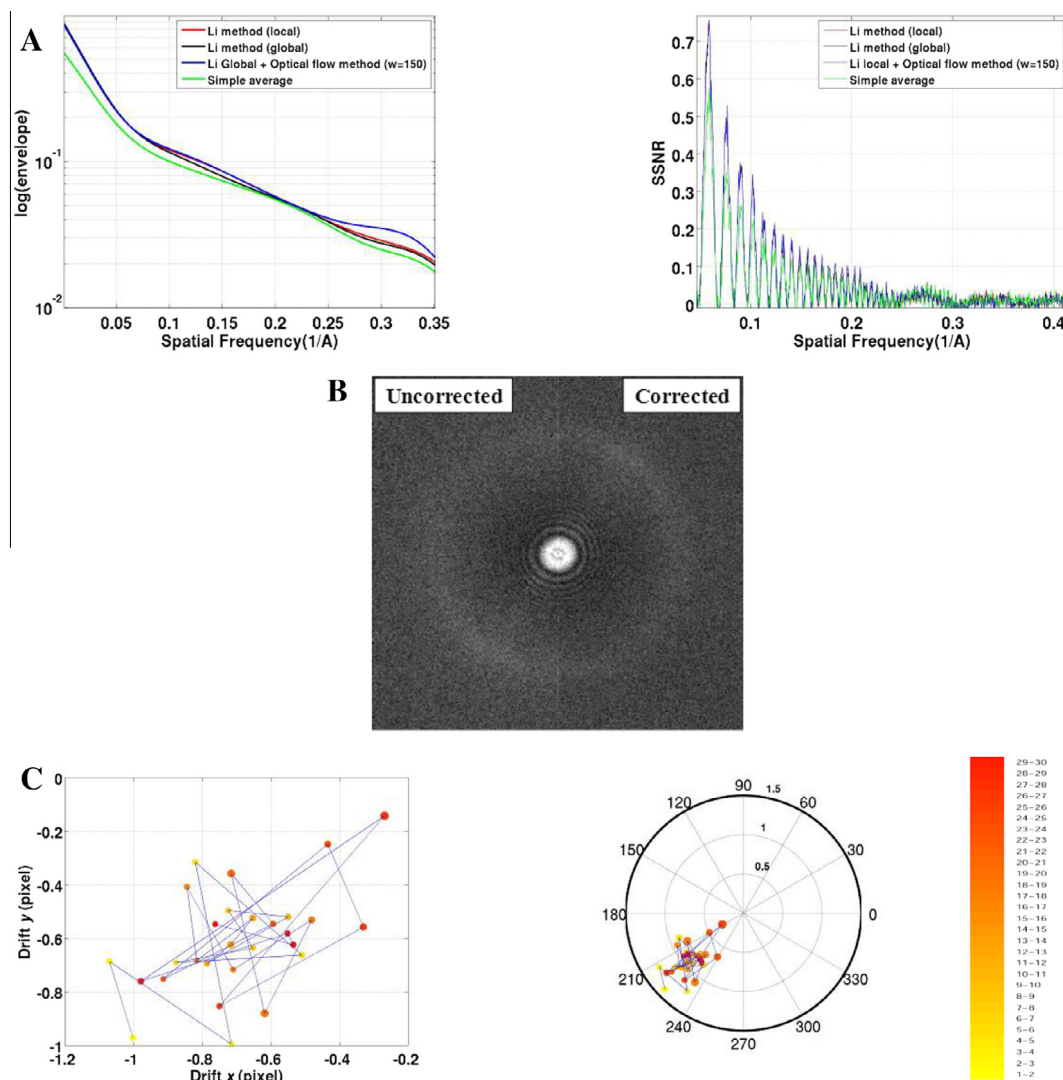
**Fig.7.** Analysis of 20S proteasome images obtained on a K2: 1D profiles corresponding to the envelope of the CTF function (in logarithmic scale) as well as the SSNR functions for the average image of a representative stack obtained using simple average (green), Li Local (red), Li Global (black) and, finally, Li Global + Optical Flow (blue) with a window size of 150 (A). Comparison of the periodograms of the uncorrected (left) and Li Global + Optical Flow (right) aligned averages (B), and analysis of the movements in between consecutive frames of the stack (C).

ered one micrograph ("205_movie_gc_window.mrcs") of the dataset from Bai et al. (2013). In Fig. 9, we show the values of this integral for different window sizes. As observed in this figure, the maximum integral value can be obtained for a window size of 150, which is set as the "default" window size of the method because it works well in most practical cases.

## 4. Discussion and conclusion

Direct detection devices represent a significant step forward for 3DEM, and at the same time, they allow high-resolution and automation to be achieved. These sensors provide sharper images with higher SNRs. Additionally, the rapid image acquisition rates of these DDDs detectors allow us to "see" biological specimens moving in their solid matrix of amorphous ice, a phenomenon referred to as beam-induced motion (BIM). Several approaches have been proposed to correct for these sample movements (Bai et al., 2013; Campbell et al., 2012; Li et al., 2013b; Wang et al., 2014). The method shown in Li et al. (2013b) consists of a frame drift-like correction with a degree of robustness to local movements provided by the use of sub-frame alignment; the output of each processed movie is a corrected averaged micrograph obtained from

the aligned frames. In (Bai et al., 2013) the alignment method is quite different because it works with the individual specimens detected and extracted from the frames, as does the method of Scheres (2014). This approach allows in-plane and out-of-plane local movement correction at the expense of requiring individual specimen detection.

In this work, we proposed a frame alignment approach that can correct for local sample movements in a rapid and efficient manner. The approach is based on a robust Optical Flow workflow, which has been tested with simulated and experimental data and compared with the method of Li et al. (2013b) and obtained satisfactory results. The input of the proposed algorithm is a set of unaligned frames (movie), and the output is a corrected averaged micrograph, as in the Li et al. (2013b) approach. The approach assumes that the particle movement is local, invertible and smooth, which in mathematics is known as a diffeomorphism. Note that the proposed model assumes that the displacements between images are equal within a window that typically has a value of approximately 150 pixels.

Additionally, to characterize the BIM pattern in the sample, a movie with the vector field of the movements between frames is also provided, which can be further analyzed using intuitive 2D
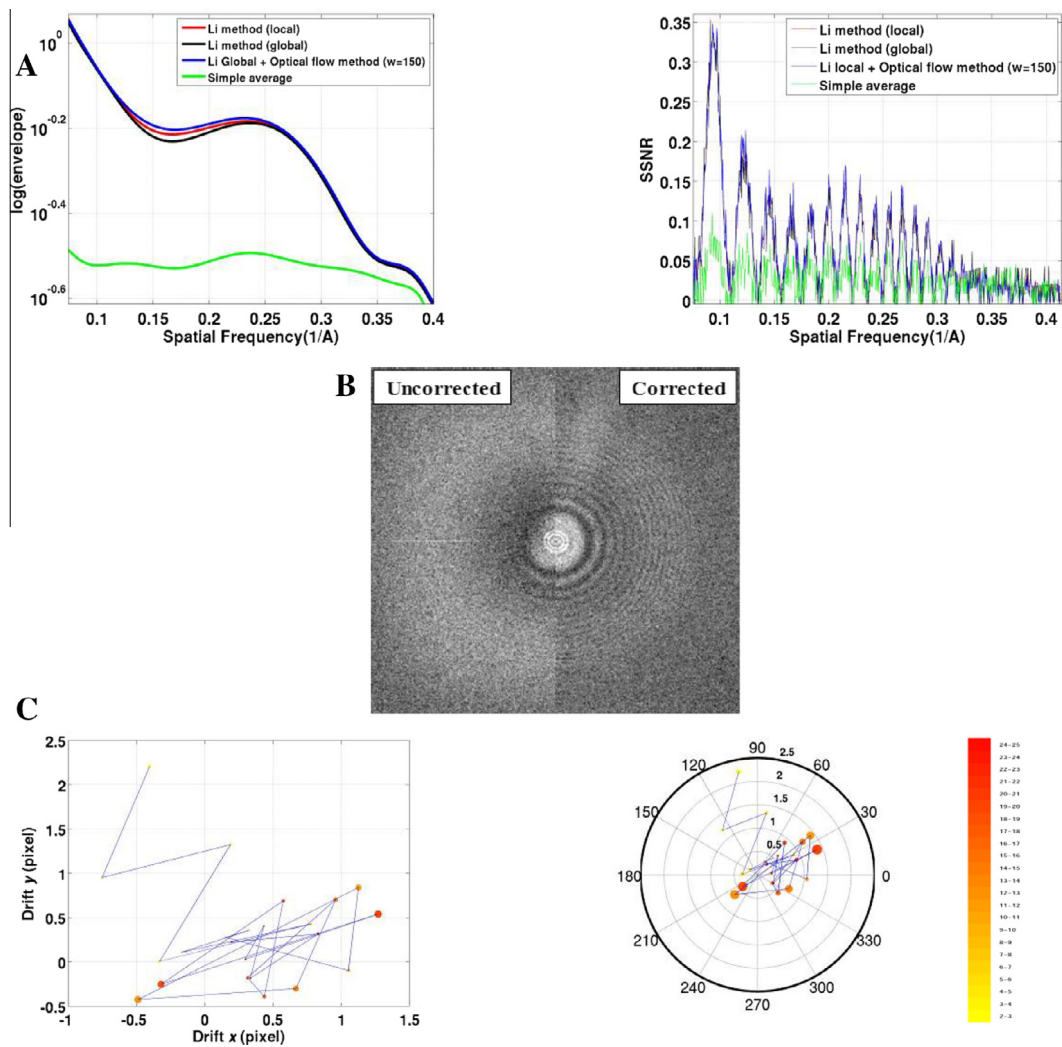
**Fig.8.** Analysis of fixed-noise pattern case from K2 data: Envelope of the CTF (in logarithmic scale) as well as the SSNR function corresponding to the corrected average after applying simple average (green), Li Local (red), Li Global (black) and Li Global + Optical Flow (blue) with a window size of 150 pixels (A). Comparison of the periodograms of the uncorrected (left) and Li Global + Optical Flow (right) aligned averages (B), and the analysis of the movements between consecutive frames (C).
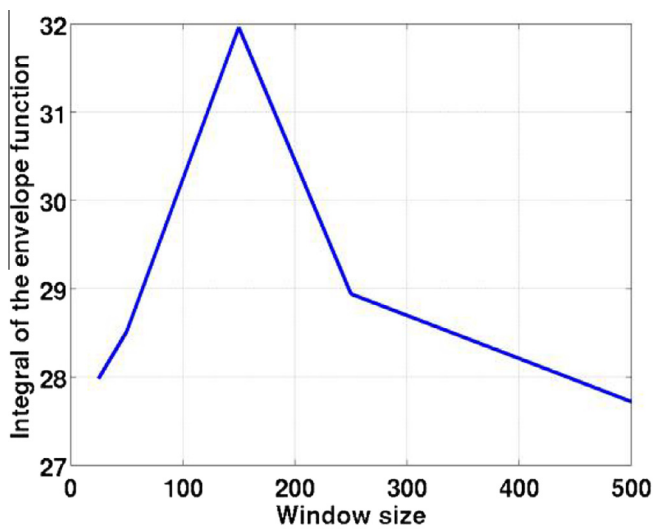


**Fig.9.** Obtaining the window size for optical flow: integrated values of the 1D profile of the CTF envelope function for different window sizes for the case of "15_movie_gc_window.mrcs". The maximum integral value is obtained for a window size of 150 pixels.

representations that show frame-to-frame movements, where the location of the center of each circle is given by the mean of the movements in the $x$- and $y$-axes and the area is described by the square root of the sum of variances along these axes.

A shortcoming of the proposed method is that it is especially sensitive to the consistency of the frame-to-frame direction of movement. Indeed, if the movements between frames are similar to a unidirectional drift, then the average of the uncorrected frames creates an effect that might translate into averaging of a particle in a frame with background in another frame in the extreme case of small particles and large movements. Still, this effect can be recovered by combining the methods of Li and Optical Flow such that by applying the Li method, we first calculate the global movement between frames, and using the Optical Flow approach, we (1) compensate for local movements and (2) calculate a BIM pattern characterization with the use of the vector fields.

Note that our approach (as well as Li's) corrects for local in-plane movements at the image level, and it cannot incorporate out-of-plane movements (i.e., known in practical terms as rocking). However, the expected extent of out-of-plane rotations and their impact on resolution may not be large, especially for small specimens. Indeed, for a specimen whose outer diameter was 150 Å, rocking of ±2.5 degrees would translate into a 3-Å shift at the periphery. Furthermore, the experience presented in Li et al.

(2013b) indicates that for small specimens, it may be impossible to apply particle-based correction methods (i.e., Bai et al. (2013)) due to the inability to detect the particles in the individual frames, making the approach presented in this work the method of choice to: (1) correct for in-plane local displacements prior to specimen detection, especially for local movements, and (2) characterize the BIM pattern with the aid of the vector fields. In much the same way, for larger particles for which rocking might be a problem, the initial use of OF might provide partially corrected frames that could significantly increase the speed of convergence of any individual specimen image-based correction approach.

In addition to its ability to efficiently correct for local movements, OF also provides a rather direct, simple and intuitive characterization of the BIM pattern with the aid of the vector fields and their associated graphical polar representations. Indeed, this analysis can indicate the amount of global versus local movements as well as its magnitude and directionality. For instance, a simple comparison of the plots shown in Figs. 5–7 shows that the ratio of global versus local movements is much larger for the 20S proteasome recorded on a K2 than for both the ribosomes or the influenza RNPs recorded on a Falcon II. Whether this observation represents a general trend associated with certain imaging conditions cannot be determined from this reduced set, but we are certainly provided with the appropriate tools to perform this type of wide-range analysis in a systematic and clear manner. Still, the accumulated experience gathered in the different laboratories contributing to this work tends to indicate that indeed, the Falcon II data present a larger degree of local distortions than the K2 data, making the Optical Flow approach particularly powerful for Falcon II data sets, whereas for K2 data, the additional improvement is reduced. We observe that this contrasting behavior may not be so unexpected because the detection mechanisms of both cameras are different. The K2 camera operates with low dose rates and large acquisition times, whereas Falcon II works with high doses obtained at short camera acquisition times.

Moreover, new figures of merit have been proposed to compare the different alignment results (uncorrected average, Li-Global and Local, and OF). Indeed, and as presented in the Appendix, the effect of summing the unaligned frames is an additional modulating term in the PSD. Therefore, the contrast transfer function (CTF) envelope and the SSNR functions are sensitive to these between frame misalignments typical of DDDs.

Finally, in future work, we will explore the possibility of using the proposed alignment approach in tomography, thus improving the quality of each tilt-pair image by alignment of the frames composing the movie. Note that this case is especially challenging because of the reduced SNR of these images.

All of the methods presented are publically available as a component of XMIPP (http://xmipp.cnb.csic.es) (de la Rosa-Trevín et al., 2013).

## Author contribution

V.A. and J.V. developed the OF method and obtained the results, Y.C. and X.L. provided K2 images and contributed to results analysis, C.S. and R.M. contributed to algorithmic developments, and J.M.C. focused on analysis of results and writing.

## Acknowledgments

## Appendix A

In this Appendix, we analyze the effect of frame misalignment on the PSD of the image resulting from summation of all frames in a stack. Because different noise models could be considered, we differentiate between "deterministic" and "random" noise. In the first case, we show that the envelope of the CTF is a notably good indicator of the extent of misalignment, whereas in the second case, both the envelope of the CTF and the SSNR are good indicators of frame misalignment.

*Deterministic noise*

In this section, we analyze the effect of frame misalignment for the case in which noise is modeled as a deterministic signal that affects all frames, e.g., the situation for certain of the ice contributions to the frame image. We consider an ideal "micrograph" given by $I(x, y)$. However, $I(x, y)$ is never experimentally measured in a DDD movie; instead, its recording is fractionated into $N$ frames with local and smooth shifts between frame $i$ and $i + 1$ given by $g_{x, i}(x, y)$ and $g_{y, i}(x, y)$. Therefore, we can recover $I(x, y)$ using a corrected average of the different frames as:

$$I(x, y) = (1/N)\sum_{i=1}^{N}I(x - g_{x, i}^{-1}(x, y), y - g_{y, i}(x, y)). \quad (A.1)$$

Note that if the shifts are smooth, we consider that they are constant in small image patches. Therefore, if we divide the images into $M$ small pieces, each image patch given by index $m \in [1, M]$ can be calculated as

$$I^m(x, y) = (1/N)\sum_{i=1}^{N}I_i^m(x - x_{0i}^m, y - y_{0i}^m). \quad (A.2)$$

Next, we calculate the Power Spectrum Density (PSD) of each patch

$$PSD^m(X, Y) = \left|(1/N)\sum_{i=1}^{N}FT\left[I_i^m(x - x_{0i}, y - y_{0i})\right]\right|^2$$

$$= \left|(1/N)\xi(X, Y)\sum_{i=1}^{N}e^{-j\left(\frac{2\pi}{N_C}Xx_{0i}+\frac{2\pi}{N_R}Yy_{0i}\right)}\right|^2 \quad (A.3)$$

where we assumed that $FT\left[I_i^m(x, y)\right] = \xi(X, Y)$, $N_C$ and $N_R$ are the number of columns and rows of the image, respectively, and $(X, Y)$ are the frequency coordinates. If we further develop Eq. (A.3), we obtain

$$PSD^m(X, Y) = \left((1/N)\xi(X, Y)\sum_{i=1}^{N}e^{-j\left(\frac{2\pi}{N_C}Xx_{0i}+\frac{2\pi}{N_R}Yy_{0i}\right)}\right)$$

$$\left((1/N)\xi(X, Y)\sum_{i=1}^{N}e^{-j\left(\frac{2\pi}{N_C}Xx_{0i}+\frac{2\pi}{N_R}Yy_{0i}\right)}\right)^* = (1/N^2)|\xi(X, Y)|^2$$

$$\left(N + \sum_{i=1}^{N}\sum_{i'=1, i'\neq i}^{N}e^{j\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)}\right) = (1/N^2)|\xi(X, Y)|^2$$

$$\left(N + 2\sum_{i=1}^{N}\sum_{i'>i}^{N}\cos\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)\right) \quad (A.4)$$

In Periodogram averaging, we assume that the PSD of the entire micrograph is the average of the PSD of all patches:

$$PSD(X, Y) = (1/M)\sum_{m=1}^{M} PSD^m(X, Y) = 1/M \sum_{m=1}^{M}$$
$$\left((1/N^2)|\xi(X,Y)|^2\left(N + 2\sum_{i=1}^{N}\sum_{i'>i}^{N}\cos\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)\right)\right)$$
$$= (1/MN^2)|\xi(X,Y)|^2\sum_{m=1}^{M}\left(\left(N + 2\sum_{i=1}^{N}\sum_{i'>i}^{N}\cos\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)\right)\right)$$

(A.5)

If there is no shift between frames, the ideal PSD is given by $\widetilde{PSD}(X, Y) = |\xi(X, Y)|^2$.

Thus, we can rewrite the actual PSD (Eq. (A.5)) as:

$$PSD(X, Y) = \Psi(X, Y)\widetilde{PSD}(X, Y) \quad (A.6)$$

where $\Psi(X, Y)$ collects all of the terms that are different from the ideal PSD. Note that if there are no shifts between the different frames, all cosines in Eq. (A.5) are equal to 1, and $\Psi(X, Y) = 1$. Otherwise, if there are shifts, then necessarily $0 \leqslant \Psi(X, Y) \leqslant 1$. Taking into account that the ideal PSD is given by

$$\widetilde{PSD} = \widetilde{PSD}_b(X, Y)|CTF(X, Y)|^2 + \widetilde{PSD}_a(X, Y) \quad (A.7)$$

where $CTF(X, Y)$ corresponds to the CTF, and $\widetilde{PSD}_b(X, Y)$ and $\widetilde{PSD}_a(X, Y)$ refer to the PSD before and after CTF, respectively (Sorzano et al., 2007; Vargas et al., 2013), by substituting Eq. (A.7) into Eq. (A.6), we obtain:

$$PSD(X, Y) = \Psi(X, Y)\left(\widetilde{PSD}_b(X, Y)|CTF(X, Y)|^2 + \widetilde{PSD}_a(X, Y)\right)$$
$$= \widetilde{PSD}_b\Psi(X, Y)|CTF(X, Y)|^2$$
$$+ \Psi(X, Y)\widetilde{PSD}_a(X, Y)$$

(A.8)

From Eq. (A.8), we can extract several conclusions. The first and most obvious is that the misalignments have a direct effect on the CTF consisting of an additional modulation term. The effective CTF is given by:

$$CTF(X, Y) = \sqrt{\Psi(X, Y)}E(X, Y)\sin\chi(X, Y) \quad (A.9)$$

Moreover, observe that $\Psi(X, Y)$ is not isotropic in general. Additionally, we note from Eq. (A.8) that the misalignment modulation function $\Psi(X, Y)$ affects the CTF envelope and the background noise $\widetilde{PSD}_a$ at the same time, although there is no effect on the CTF phase term $\chi(X, Y)$.. We observe that the SSNR (Booth et al., 2004), which is estimated from the radially averaged CTF envelope divided by the noise curve (the baseline that passes through the zeroes in the CTF) is not adequate for characterizing good or bad frame alignments because it is insensitive to $\Psi(X, Y)$. Consequently, a good approach to analyzing the frame alignment quality is to study the decay in an "effective" CTF envelope function that corresponds to

$$\hat{E}(x, y) = \sqrt{\Psi(X, Y)}E(X, Y) \quad (A.10)$$

In the remainder of the paper, we refer to this "effective" CTF envelope as the CTF envelope

*Random noise*

We consider this case if noise is modeled as a random event added to each frame. Up to (A.2), there is no difference between the deterministic and the random case. However, when we calculate the PSD of the $m$th patch, we should calculate it in the correct manner for random signals. Let us assume that each patch has a

deterministic component $I_{i,d}^m(x, y)$ and a random component $\varepsilon_i^m(x, y)$,

$$I^m(x, y) = (1/N)\sum_{i=1}^{N}I_{i,d}^m(x - x_{0i}^m, y - y_{0i}^m) + (1/N)\sum_{i=1}^{N}\varepsilon_i^m(x - x_{0i}^m, y - y_{0i}^m).$$

(A.3')

Thus

$$PSD^m(X, Y) = \left|(1/N)\sum_{i=1}^{N}FT[I_{i,d}^m(x - x_{0i}, y - y_{0i})]\right|^2$$
$$+ (1/N)PSD_\varepsilon(X, Y) = (1/N^2)|\xi(X, Y)|^2$$
$$\left(N + 2\sum_{i=1}^{N}\sum_{i'>i}^{N}\cos\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)\right)$$
$$+ (1/N)PSD_\varepsilon(X, Y)$$

(A.4')

Next, we calculate the PSD of the entire micrograph

$$PSD(X, Y) = (1/M)\sum_{m=1}^{M}PSD^m(X, Y) = (1/MN^2)|\xi(X, Y)|^2$$
$$\sum_{m=1}^{M}\left(\left(N + 2\sum_{i=1}^{N}\sum_{i'>i}^{N}\cos\left(\frac{2\pi}{N_C}(x_{0i'}^m - x_{0i}^m)X + \frac{2\pi}{N_R}(y_{0i'}^m - y_{0i}^m)Y\right)\right)\right)$$
$$+ (1/N)PSD_\varepsilon(X, Y)$$

(A.5')

Compared with the deterministic case, we write

$$PSD(X, Y) = \Psi(X, Y)\widetilde{PSD}(X, Y) + (1/N)PSD_\varepsilon(X, Y)$$
$$= \widetilde{PSD}_b\Psi(X, Y)|CTF(X, Y)|^2$$
$$+ \Psi(X, Y)\widetilde{PSD}_a(X, Y) + (1/N)PSD_\varepsilon(X, Y) \quad (A.6')$$

The SSNR becomes

$$SSNR(X, Y) = \frac{\widetilde{PSD}_b\Psi(X, Y)|CTF(X, Y)|^2}{\Psi(X, Y)\widetilde{PSD}_a(X, Y) + (1/N)PSD_\varepsilon(X, Y)}$$
$$= \frac{\widetilde{PSD}_b(X, Y)|CTF(X, Y)|^2}{\widetilde{PSD}_a(X, Y) + \frac{1}{N\Psi(X, Y)}PSD_\varepsilon(X, Y)} \quad (A.7')$$

The degree of misalignment is encoded in the $\Psi(X, Y)$ term. As the misalignment increases, $\Psi(X, Y)$ approaches 0, as does the $SSNR(X, Y)$. In this case, both the envelope and the SSNR serve as good measures of misalignment. As suggested by Wang et al. (2014), note that the noise appears to be a mix of deterministic and random noise for DDD movies.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2015.02.001.

## References

Bai, X.-C., Fernandez, I.S., McMullan, G., Scheres, S.H., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. eLife 2 (e0046).

Bammes, B.E., Rochat, R.H., Jakana, J., Chen, D.-H., Chiu, W., 2012. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. J. Struct. Biol. 177, 589–601.

Booth, C.R., Jiang, W., Baker, M.L., Hong Zhou, Z., Ludtke, S.J., Chiu, W., 2004. A 9 Å single particle reconstruction from CCD captured images on a 200 kV electron cryomicroscope. J. Struct. Biol. 147, 116–127.

Bouguet, J.-Y., 2000. Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the Algorithm.

Bouguet, J.-Y., 2001. Pyramidal Implementation of the Affine Lucas Kanade Feature Tracker Description of the Algorithm, Vol. 2. Intel Corporation (3).

Bradski, G., 2000. {The OpenCV Library}. Dr. Dobb's Journal of Software Tools.

Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Henderson, R., Grigorieff, N., 2012. Beam-induced motion of vitrified specimen on holey carbon film. J. Struct. Biol. 177, 630–637.

Campbell, M.G., Cheng, A., Brilot, A.F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. Structure 20, 1823–1828.

de la Rosa-Trevín, J.M., Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J.M., Sorzano, C.O.S., 2013. Xmipp 3.0: an improved software suite for image processing in electron microscopy. J. Struct. Biol. 184, 321–328.

Frank, J., 1996. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Academic Press.

Frank, J., 2006. Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State. Oxford University Press.

Glaeser, R.M., Hall, R.J., 2011. Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects. Biophys. J. 100, 2331–2337.

Horn, B.K., Schunck, B.G., 1981. Determining optical flow. In: Technical Symposium East. International Society for Optics and Photonics, pp. 319–331.

Jin, L., Bilhorn, R., 2010. Performance of the DDD as a direct electron detector for low dose electron microscopy. Microsc. Microanal. 16, 854–855.

Li, X., Zheng, S.Q., Egami, K., Agard, D.A., Cheng, Y., 2013a. Influence of electron dose rate on electron counting images recorded with the K2 camera. J. Struct. Biol. 184, 251–260.

Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013b. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590.

Lucas, B.D., Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. In: IJCAI, Vol. 81, pp. 674–679.

Milazzo, A.-C., Leblanc, P., Duttweiler, F., Jin, L., Bouwer, J.C., Peltier, S., Ellisman, M., Bieser, F., Matis, H.S., Wieman, H., 2005. Active pixel sensor array as a detector for electron microscopy. Ultramicroscopy 104, 152–159.

Scheres, S.H., 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Scheres, S.H., 2014. Beam-induced motion correction for sub-megadalton cryo-EM particles. eLife 3 (e03665).

Shigematsu, H., Sigworth, F., 2013. Noise models and cryo-EM drift correction with a direct-electron camera. Ultramicroscopy 131, 61–69.

Sorzano, C., Jonic, S., Núñez-Ramírez, R., Boisset, N., Carazo, J., 2007. Fast, robust, and accurate determination of transmission electron microscopy contrast transfer function. J. Struct. Biol. 160, 249–262.

Vargas, J., Otón, J., Marabini, R., Jonic, S., De La Rosa-Trevín, J., Carazo, J., Sorzano, C., 2013. FASTDEF: Fast defocus and astigmatism estimation for high-throughput transmission electron microscopy. J. Struct. Biol. 181, 136–148.

Vargas, J., Restrepo, R., Belenguer, T., 2014. Shack-Hartmann spot dislocation map determination using an optical flow method. Opt. Express 22, 1319–1329.

Wang, Z., Hryc, C.F., Bammes, B., Afonine, P.V., Jakana, J., Chen, D.-H., Liu, X., Baker, M.L., Kao, C., Ludtke, S.J., 2014. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. Nat. Commun. 5.

# A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs

V. Abrishami[1,2,*], A. Zaldívar-Peraza[1,2], J. M. de la Rosa-Trevín[1,2], J. Vargas[1], J. Otón[1], R. Marabini[2], Y. Shkolnisky[3], J. M. Carazo[1] and C. O. S. Sorzano[1,4,*]

[1]Biocomputing Unit, National Center of Biotechnology (CSIC), [2]Department of Computer Science, University Autonoma de Madrid, Campus Universidad Autonoma s/n, 28049 Cantoblanco, Madrid, Spain, [3]Department Applied Mathematics, Tel Aviv University, Ramat Aviv, Tel Aviv 69978 Israel and [4]Bioengineering Lab, Escuela Politecnica Superior, University San Pablo CEU, 28668 Boadilla del Monte, Madrid, Spain

Associate Editor: Jonathan Wren

**ABSTRACT**

**Motivation:** Structural information of macromolecular complexes provides key insights into the way they carry out their biological functions. Achieving high-resolution structural details with electron microscopy requires the identification of a large number (up to hundreds of thousands) of single particles from electron micrographs, which is a laborious task if it has to be manually done and constitutes a hurdle towards high-throughput. Automatic particle selection in micrographs is far from being settled and new and more robust algorithms are required to reduce the number of false positives and false negatives.

**Results:** In this article, we introduce an automatic particle picker that learns from the user the kind of particles he is interested in. Particle candidates are quickly and robustly classified as particles or nonparticles. A number of new discriminative shape-related features as well as some statistical description of the image grey intensities are used to train two support vector machine classifiers. Experimental results demonstrate that the proposed method: (i) has a considerably low computational complexity and (ii) provides results better or comparable with previously reported methods at a fraction of their computing time.

**Availability:** The algorithm is fully implemented in the open-source Xmipp package and downloadable from http://xmipp.cnb.csic.es.

**Contact:** vabrishami@cnb.csic.es or coss@cnb.csic.es

**Supplementary Information:** Supplementary data are available at *Bioinformatics* online.

## 1 INTRODUCTION

Electron Microscopy (EM) is a key tool to study the structure and function of biological macromolecules at a medium–high resolution. Single particle analysis is an EM modality in which multiple copies of the same macromolecule are simultaneously imaged into a single micrograph. Particles from several hundreds or even thousands of micrographs are commonly employed in a structural study. The standard data processing workflow of single particle reconstruction includes: particle selection, particle alignment, particle classification, three-dimensional (3D)

reconstruction and model refinement (Chen and Grigorieff, 2007; Sorzano *et al.*, 2012). Different views of a specimen are required for the 3D reconstruction of a complex, but these views suffer from low signal-to-noise ratio (SNR) (due to low-dose imaging) (Glaeser, 1971), low contrast (due to close to focus conditions), and image deformations (due to the microscope aberrations). It is generally accepted that high resolution can only be achieved with thousands of projection images, so that the 3D reconstruction algorithm can compensate for these challenging imaging conditions. In particular, there is generally a direct relationship between the number of selected particles and the maximum achievable resolution (Henderson, 1995). Manually identifying that number of particles is not just time consuming and laborious, but also an error-prone process. A robust automatic particle picker (APP) algorithm is, therefore, indispensable to enhance the technique's throughput.

As the selection of several thousands of particles from low-dose micrographs is the first and one of the crucial steps towards a high-resolution reconstruction, a large amount of effort has been made by researchers to develop accurate methods for APP. These methods have been classified into groups by different authors (Mallick *et al.*, 2004; Nicholson and Glaeser, 2001; Zhu *et al.*, 2004). Among them, the classification by Nicholson and Glaeser (2001), suggests a general categorization into template matching-based and feature-based approaches. Template matching-based methods (Chen and Grigorieff, 2007; Huang and Penczek, 2004; Ludtke *et al.*, 1999; Plaisier *et al.*, 2004; Roseman, 2003; Sigworth, 2004; Wong *et al.*, 2004) calculate the cross-correlation (or any other measure of similarity) between a set of templates and a micrograph image to seek for particle candidates. Templates are obtained either from different projections of an initial 3D volume (Huang and Penczek, 2004; Wong *et al.*, 2004) or from a number of manually picked particles (Hall and Patwardhan, 2004; Roseman, 2003). Instead of using all templates (either from different projections of an initial 3D volume or from a number of manually picked particles), which severely increases the processing time, other alternatives can be employed, such as eigenimages of templates (Sigworth, 2004) or some form of an average of each template cluster (Wong *et al.*, 2004). In feature-based approaches, particles are sought through the calculation of some prominent geometric and/or statistical features of the particle images (Arbeláez *et al.*, 2011; Hall and Patwardhan, 2004; Langlois *et al.*, 2011; Mallick

---

*To whom correspondence should be addressed.

*et al.*, 2004; Ogura and Sato, 2004; Sorzano *et al.*, 2009; Volkmann, 2004; Yu and Bajaj, 2004; Zhao *et al.*, 2013; Zhu *et al.*, 2003). Feature-based methods can be reference-free or learning-based. In the former, features corresponding to particles are known to fall within a certain region of the feature space and, therefore, no training is necessary and the algorithm can start picking particles straightaway; in the latter, however, a set of particles and non-particles are required to train a classifier which is then able to distinguish between particles and non-particles based on the training features. Although reference-free methods require less effort from the user, they are of limited applicability, because the space region corresponding to particle features has to be known *a priori*.

We previously introduced a feature-based APP method (Sorzano *et al.*, 2009) that learns features from the user selected particles via a continuous learning phase (the algorithm is available in the open-source package Xmipp 2.4). In a manual picking step, a small dataset of particle and non-particle images is formed to train an ensemble naive Bayesian classifier. Once the classifier is trained, it suggests new particles in a new micrograph. The user supervises this result by discarding the wrongly picked particles and identifying the disregarded ones. This feedback information is then submitted to the classifier which is updated to accommodate this new information. This semi-automatic picking is continued on several micrographs until the user is satisfied by the results. At this point, the trained classifier carries out the selection of the particles in the remaining micrographs in a fully automatic way.

In this article, we introduce an APP method that follows the general learning structure of Sorzano *et al.* (2009), but major improvements are made to increase speed and accuracy. Our new feature vector is completely different from the previous method and consists of a number of geometrical and statistical features; it is robust to noise, very fast to compute and most of its features are rotationally invariant. Instead of Naive Bayesian (NB), we now use a support vector machine (SVM) as the base classifier due to its interesting properties, like high generalization capabilities and small training/classification time. In order to reduce the number of false positives, two SVM classifiers are used: one for discriminating between particle candidates and non-particle objects and the other for checking if a particle candidate recognized by the first classifier is a real particle or not. In contrast to Sorzano *et al.* (2009), which explores a big search space for particles, the proposed method limits the search space to the peaks obtained from the cross-correlation of the micrograph with some pre-computed templates. Templates are generated during the manual picking step by clustering [the clustering algorithm is described in Sorzano *et al.* (2010)] the hand-picked particles and selecting the average of each cluster. As the correlation with all orientations of the template noticeably increases the computation time, a rotationally averaged template can be used instead. The algorithm has been successfully tested on three experimental datasets and compared with our previous algorithm (which had, in turn, been compared with other approaches), resulting in a more than an order of magnitude decrease in computing time while achieving even better performance.

## 2 METHODS

There are three crucial steps in the proposed algorithm: identifying initial possible locations of particles, locally characterizing the image by means

of a low-dimensional feature vector and classifying each location as particle or non-particle. In this section, we go through the details of each step.

### 2.1 Identification of possible particle centres

In principle, each pixel of a micrograph has the potential to be the centre of a particle. In practice, checking each pixel is rather time consuming, and so the number of possible candidates to evaluate must be reduced. Therefore, we cross-correlate the input micrograph with suitable templates to make an initial guess about the candidate particle positions. In this way, not only the search space is noticeably reduced, but also the rate of false positives. During the training phase, we cluster manually picked up particles into a few number of classes using an algorithm similar to the one described in Sorzano *et al.* (2010). The number of clusters can be increased or decreased dynamically to facilitate user's ongoing control on the quality of the templates. These templates are correlated with the input micrograph at all possible orientations (at each location of the correlation map, we should keep the maximum observed correlation for all templates and all orientations). If the particle is relatively globular, the calculation of the correlation map can be further accelerated by substituting the template by its rotational average, avoiding in this way the need to correlate with all possible orientations.

Local maxima of the correlation map are possible candidate locations for being at the centre of a particle. We further reduce this candidate set by ranking local maxima according to their correlation values and keeping only a portion of the local maxima with a value higher than a small threshold. This portion may not be the same for all datasets, and basically depends on the density of particles within the micrographs.

We should note that the cross-correlation with a template in our approach is just a first step to have a set of possible candidates, but that in no way do we rely on this cross-correlation for further processing of these candidates into particles and non-particles. In fact, the chosen threshold for local maxima is small enough such that even particles that were related to excluded templates could still be detected at this initial step (Fig. 1). In most practical cases we have found that the use of only one template is generally enough. For instance, in Figure 1, only a single class representative has been azimuthally averaged and cross-correlated with the micrograph, to identify the local maxima. In this figure, picking
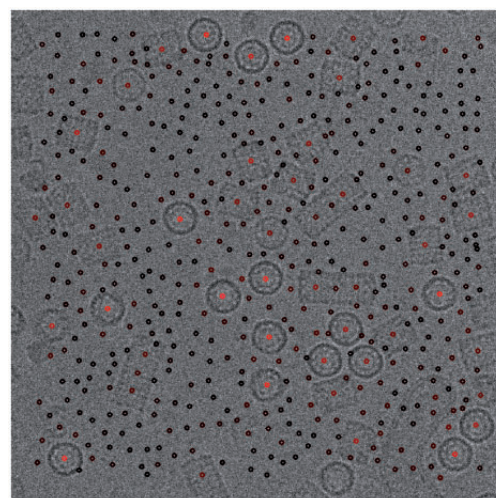


**Fig. 1.** Micrograph from the KLH training dataset with cross-correlation peaks superimposed. The peaks have been obtained by cross-correlation of a single class representative and the micrograph. Colours show the energy of each peak by black as lowest and red as highest

**V.Abrishami et al.**

10% of local maxima (with the highest cross-correlation values) results in local maxima with minimum value of 0.27 and maximum value of 0.98. As can be seen, at the location of each particle a local maximum with a specific energy can be observed. Still, this behaviour depends on the practical case at hand, and consequently, the number of templates is dynamic, so if the result is not satisfactory the user can request for more templates.

## 2.2 Feature extraction

We compute a feature vector at each location previously identified as a candidate for particle centre. This vector is used by the classifier in the next step to distinguish between particles and non-particles. Two properties have been sought for the feature vector: being robust to noise (due to the low SNR in the micrographs) and being rotationally invariant (to save computational time and avoid having to look for the particles in all possible orientations). We use a feature vector that is robust to noise and most of its features are rotationally invariant. it is made of three subsets of features: the first two feature subsets are sensitive to the particle shape, whereas the third one encodes the particle grey intensities. The first and third feature subsets are rotational invariant but not the second one.

*2.2.1 Particle shape description at different frequencies*  Micrographs are submitted to a filter bank with $N_h$ raised cosine band pass filters to decompose them into several sub-band images in order to being able later on to extract features associated to particular frequencies (Fig. 2). The filtered micrograph by the $k$th ($k = 1, \ldots, N_h$) filter, $M_k(\mathbf{r})$ (i.e. $k$th sub-band image) is computed by

$$M_k(\mathbf{r}) = FT^{-1}\{M(\mathbf{R})H_k(\mathbf{R})\}, \tag{1}$$

where $FT^{-1}$ is the inverse Fourier transform and $M(\mathbf{R})$ and $H_k(\mathbf{R})$ show the Fourier transform of the micrograph and $k$th Fourier filter in the filter bank, respectively. $\mathbf{R}$ is the two-dimensional (2D) spatial frequency, and $\mathbf{r}$ is the spatial coordinate within the image. Each raised cosine band pass filter has a particular width $\Delta_R$ and a decay $\delta_R$ (see the last image in column B of Fig. 2). The low-pass filter in the filter bank is defined by a transfer function given by

$$H_1(\mathbf{R}) = \begin{cases} 1, & R \le \Delta_R \\ \frac{1}{2}\left(1 + \cos\left(\frac{\pi(R - \Delta_R)}{\delta_R}\right)\right), & \Delta_R < R \le \Delta_R + \delta_R \\ 0, & R > \Delta_R + \delta_R \end{cases} \tag{2}$$

where $R$ is the modulus of $\mathbf{R}$. The $k$th band pass filter ($k = 2, 3, \ldots, N_h$) is defined by the transfer function $H_k(\mathbf{R})$ defined as

$$\begin{cases} 0, & R \le (k-1)\Delta_R - \delta_R \\ \frac{1}{2}\left(1 + \cos\left(\frac{\pi(R - (k-1)\Delta_R)}{\delta_R}\right)\right), & (k-1)\Delta_R - \delta_R < R \le (k-1)\Delta_R \\ 1, & (k-1)\Delta_R < R \le k\Delta_R \\ \frac{1}{2}\left(1 + \cos\left(\frac{\pi(R - k\Delta_R)}{\delta_R}\right)\right), & k\Delta_R < R \le k\Delta_R + \delta_R \\ 0, & R > k\Delta_R + \delta_R \end{cases} \tag{3}$$

Let us concentrate now on a given particle within a micrograph. We will refer to this boxed image as $I(\mathbf{r})$. Let us call $I_k(\mathbf{r})$ the corresponding boxed image extracted from $M_k(\mathbf{r})$. By considering $r$ as the distance from the image centre in the direction of $\theta$, we express this boxed, filtered image in polar coordinates, $I_k(r, \theta)$ and compute the cross-correlation function of pairs of $I_k(r, \theta)$ images [this step has been partially inspired by Schatz and van Heel (1990)]. Note that the cross-correlation is rotationally invariant in 2D polar coordinates (see proof in Supplementary material). Let us define $\psi_{kk'}(\Delta r, \Delta\theta)$ as the cross-correlation function between polar, sub-band images $I_k(r, \theta)$ and $I_{k'}(r, \theta)$, as below

$$\psi_{kk'}(\Delta r, \Delta\theta) = \sum_r \sum_\theta I_k(r, \theta)I_{k'}(r + \Delta r, \theta + \Delta\theta), \tag{4}$$

where $I_{k'}(r, \theta)$ is shifted over $I_k(r, \theta)$ by $\Delta r$ ([0, $r_p$], where $r_p$ is the radius of the particle), and by $\Delta\theta$ ([$-180°, 180°$]) along $r$ and $\theta$, respectively (see column D in Fig. 2). In particular, for each $k$, we calculate $\psi_{kk'}$ for $k' = k$ (autocorrelation; $k = 1, \ldots, N_h$), $k' = k + 1$ ($k = 1, \ldots, N_h - 1$) and $k' = k + 2$ ($k = 1, \ldots, N_h - 2$). The autocorrelation of a given band is related to the particle shape, and the cross-correlation between sub-bands reveals the linear relationships between the shapes at two different frequency bands. We can see in Figure 3 an example of how these cross-correlation functions can, indeed, distinguish between particles and non-particles. From this figure, it is clear that cross-correlation functions for particles (first row) are quite different from those for non-particles (second row). It is worth mentioning that, for the sake of speed up, we just consider the 2D projections and not the original 3D object. According to this, cross-correlation functions are not rotationally invariant in 3D, and therefore we need to have enough projections from different orientations to fully cover the projection space.

The cross-correlation functions of the training set of particles for a particular combination $kk'$ are highly redundant and can be easily compressed using principal component analysis (PCA) (Pearson, 1901). There are as many $\psi_{kk'}$ images as candidate particles, which is a number in the many thousands. This is the set from which the PCA basis is extracted (there are, therefore, as many PCA's as all combinations of $k$ and $k'$). For each $\psi_{kk'}$ image, we keep its projection onto the first $N_b$ PCA vectors as features to be used during the classification step. Note that the PCA basis is calculated for the $kk'$ cross-correlations of the training dataset of particles. This means that the $kk'$ cross-correlations of non-particles will be poorly represented by this PCA basis. In this way, the dimensionality reduction itself is presumed to have a positive impact on the classification accuracy. An example of these bases can be seen in Figure 4, where four eigenvectors are shown for each cross-correlation function.
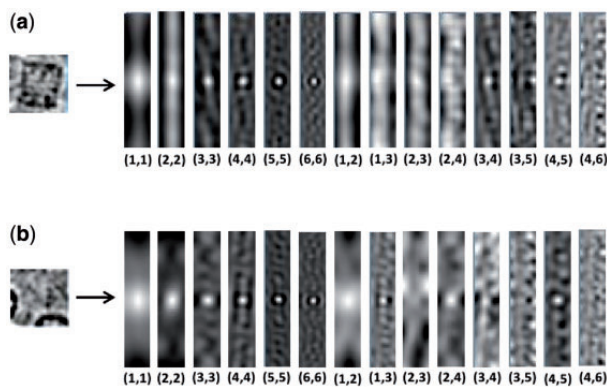
*2.2.2 Particle shape description in a particle-adapted rotational invariant subspace*  Ponce and Singer (2011) proposed to calculate an image basis that is adapted to the kind of images being studied and their in-plane rotations. They do this by calculating the PCA of a set of images and all their possible rotations. We apply this principle to extract some features from boxed particles or particle candidates [this idea is partially the same as in Dube *et al.* (1993)]. Given the training particles provided by the user, we first align them into a few templates using a process similar to that described in Sorzano *et al.* (2010). Then, we compute the PCA basis associated to these templates (Fig. 5) to form a rotational invariant subspace (the basis of this subspace is able to reproduce any 2D rotation of particle images). We keep the first $N_{rb}$ projection coefficients onto this basis as part of the feature vector. Note that the subspace that is spanned by these vectors is rotationally invariant, but not the basis itself. Therefore, particles with different 2D orientations have different coefficients but still can be efficiently approximated using this basis. Again, this basis has been especially designed to represent good particle images; consequently, non-particles will be poorly represented by the basis helping the classifier to perform its task.

*2.2.3 Particle intensity*  Desired particles are assumed to follow a specified pattern of intensity distribution once the boxed image is normalized to have zero mean and unit power (Sorzano *et al.*, 2004). For instance, particles in the carbon region are normally discarded. To capture the intensity features desired by the user we calculate the mean, the SD as well as $N_i$ equidistributed deciles of the intensity histogram.

*2.2.4 Feature vector*  We collect all these features into a feature vector that characterizes the boxed image to be classified. The vector size is $3N_b(N_h - 1) + N_{rb} + N_i + 2$. By default, we suggest to use $N_b = 4$, $N_h = 6$ ($\Delta_R = 0.025$, $\delta_R = 0.02$ in digital frequencies normalized to 0.5), $N_{rb} = 20$ and $N_i = 9$. This produces a feature vector of dimension 91. In practice, we have observed that these choices provide generally good results on all the tested datasets. The values of the parameters depend on
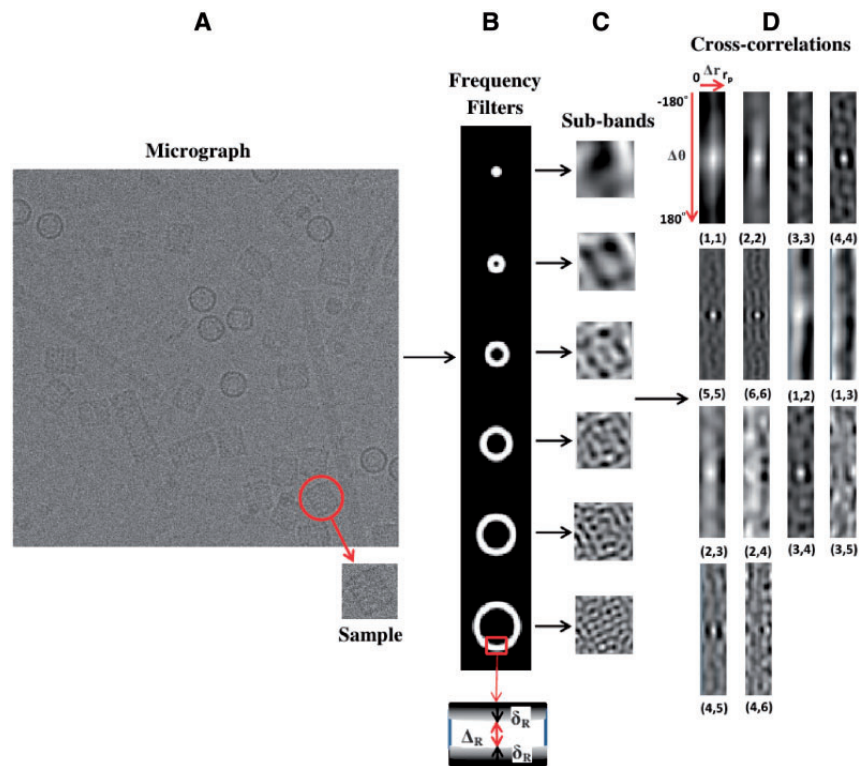
**Fig. 2.** Particle shape description at different frequencies. First, $N_h$ bandpass filters (column B), with a width of $\Delta_R$ and a decay of $\delta_R$ (see the last image in this column), are applied to the input micrograph (column A) and boxed images are extracted at the location of particles or particle candidates (column C). Sub-band images for the particle (column C) are converted to polar form, and cross-correlations between different sub-bands are calculated (column D). $\Delta\theta$ [$(-180^\circ, 180^\circ)$] and $\Delta r$ ([$0, r_p$], where $r_p$ is the radius of particle) are shifting parameters to slide one polar image on one another. The indexes below each cross-correlation (column D) refer to the polar images that form it. Parameters for this figure are $N_h = 6$, $\Delta_R = 0.025$, $\delta_R = 0.02$, $r_p = 25$ (see the main text)
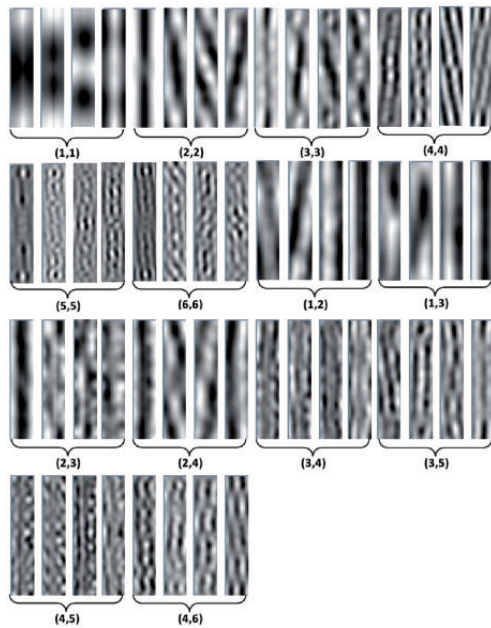


**Fig. 3.** Particle shape description at different frequency bands: cross-correlation among sub-bands for a particle (**a**) and non-particle (**b**). The indexes below each cross-correlation refer to the $k$ and $k'$ sub-bands. Note the important differences between particles and non-particles for the functions (1, 1), (2, 2), (3, 3), (1, 2), (1, 3), (2, 3) and (2, 4)

both properties of the micrographs and particles. For instance, for a low defocus set of micrographs, we need a higher value of $N_h$ to extract the information of higher frequencies, whereas for smaller particles we should increase $N_b$ and $N_{rb}$ to be able to regenerate them from the basis.
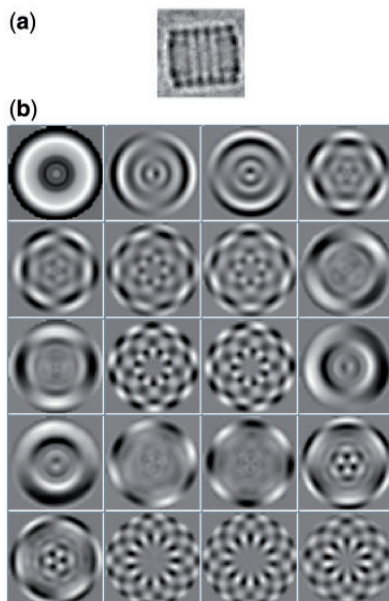
## 2.3 Classification

Distinguishing between particles and non-particles is, in general, a complex task due to the low contrast and low SNR present in the micrographs. Additionally, the problem is complicated by having to distinguish between particles and damaged particles, contaminated particles, partly formed particles, etc., which, in general we will denote as 'errors'. We use non-linear, binary support vector machines (SVMs) (Boser *et al.*, 1992; Cortes and Vapnik, 1995) as classifiers due to their good performance in other classification problems [it is particularly used successfully in APP by Arbeláez *et al.* (2011); Zhao *et al.* (2013)], their robustness to noise and their speed. The general idea of the SVM classifier is to find an optimum hyperplane in an n-dimensional space by which two different classes are distinguishable. LIBSVM (Chang and Lin, 2011) is an efficient and widely used implementation of the SVM. This package suggests a variety of kernels to perform the non-linear classification. We use this package with a radial basis function (RBF) kernel to gain a high accuracy in our classification (see Supplementary material for an introduction to SVM).

We use two SVM classifiers. The first classifier is responsible to discriminate between particle objects and any other kind of objects (non-particles and errors). Because of the similarity between errors and particles, the output of the first classifier is not so accurate, and some errors are labelled as particles. Therefore, to reduce the false positives to a feasible extent, the second classifier is dedicated to just focus on distinguishing particles from errors. This strategy was already used by Sorzano *et al.* (2009). Figure 6 shows the behaviour of the classifier for three types of objects. As can be seen, the first classifier passes the error, but the
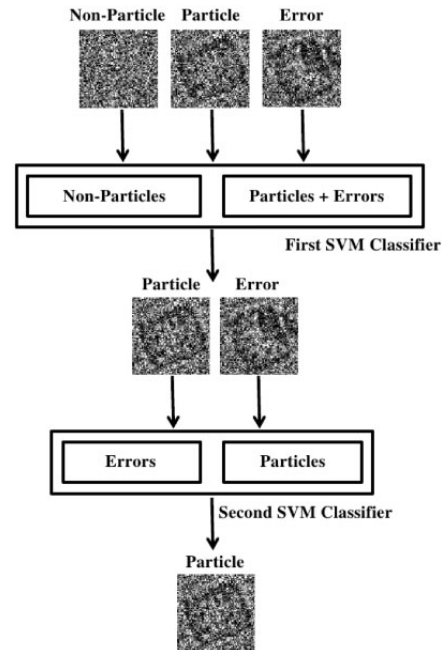
*V.Abrishami et al.*

**Fig. 4.** First four eigenvectors corresponding to the PCA of the cross-correlation functions between two sub-band images for manually picked particles of the KLH dataset. The index under each group of four images shows the contributed sub-band images ($kk'$, in the text)



**Fig. 6.** Behaviour of the classifiers for three types of objects. The first classifier may classify errors as particles, but the second classifier is designed to remove errors from the final result

the classifiers. This manual step can be continued with more micrographs to expand the training set. It is worth pointing out that none of the classifiers are trained at this point.

Once the training set is large enough (empirically, at least 30 training particles are required), the process enters a supervised phase. The rotationally invariant subspace as well as the PCAs for the polar, sub-band cross-correlations are calculated on the training particles. Then, feature vectors are calculated for the manually selected particles and non-particles. Finally, the first classifier is trained using this data. At this point, the algorithm tries to automatically pick the next micrograph in the list of micrographs that was not previously manually picked. After suggesting possible particle locations, the user can correct the results by adding those missed particles (false negatives) and removing wrongly picked particles (false positives). After being corrected by the user, the first classifier is retrained using all previous information plus the new set of false negatives and false positives. The second classifier is trained to distinguish between all particles known so far and the set of false positives. This process can be repeated several times on more micrographs till the performance of the classifiers is not further improved by user corrections. This ongoing learning process is particularly interesting because the classifiers carefully adapt to the user's preferences.

When the user is satisfied with the performance of the classifier during the semi-supervised phase, he can go to a fully automatic particle picking mode, in which all micrographs that have not been picked yet are automatically picked (in parallel). At the end of this process, the user can supervise the result and eliminate wrongly picked particles or add missed particles.



**Fig. 5.** Eigenvectors (Ponce and Singer, 2011) corresponding to a given template, (**a**) Given particle template. (**b**) First 20 eigenvectors of the template that generate a rotational invariant subspace

second classifier rejects it. In fact, the learning process follows the same steps as in the algorithm by Sorzano *et al.* (2009). First, the user picks all particles from the first micrograph and they are clustered in order to construct particle templates. All non-selected locations are assumed to correspond to non-particles and some of them will be later used to train

## 3 RESULTS

We applied our APP method on three datasets to assess the speed and accuracy of the proposed algorithm for micrographs with different contrast, density and particle shape and size. These

datasets are: KLH (keyhole limpet haemocyanin), adenovirus and helicase. All the experiments were done with the same fixed parameters ($N_b = 4$, $N_h = 6$, $N_{rb} = 20$ and $N_i = 9$) on a single core of a CPU Intel Core i5, 64 bits, 2.53 GHz (of a standard laptop with 4 gigabytes of RAM).

To show the results in a quantitative way, we have used the performance metrics introduced by Langlois and Frank (2011). If TP is the number of true positives, FP the number of false positives and FN the number of false negatives, then these metrics are defined as

- Precision $= \frac{TP}{FP+TP}$

- Recall $= \frac{TP}{FN+TP}$

- $F$-measure $= 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$

Precision shows the fraction of picked particles by the algorithm that is real particles, and recall indicates the fraction of true particles that are picked by the algorithm. $F$-measure is a harmonic mean to summarize both recall and precision.

### 3.1 KLH dataset

This dataset was produced as a general benchmark for APP by Zhu *et al.* (2003). It includes 82 micrographs of KLH particles. A Phillips CM200 TEM was used to record the micrographs on a $2\,K \times 2\,K$ CCD Tietz camera at a magnification of $66\,000\times$ and a voltage of 120 kV. The sampling rate at this magnification was $2.2\,\text{Å/pixel}$.

In the 3DEM Benchmark site (http://i2pc.cnb.csic.es/3dem benchmark) the dataset has been split into two datasets: one with 30 micrographs for training and another one with 50 micrographs for testing. We set the size of particles to 200 pixels, and processed 10% of the local maxima of the correlation map.

To evaluate the connection between the number of particles and accuracy metrics, we calculated precision, recall and $F$-measure for the test dataset after training the classifier with different numbers of true particles of the training dataset. Figure 7 shows how these three parameters change as the number of training particles is increased. As can be seen in this figure, precision, recall and $F$-measure of the algorithm are increased if we keep training the classifier with more and more particles. Training with more than 173 particles (13 micrographs), a precision range of [80.0, 85.2] and a recall range of [73.1, 77.8] are achievable. $F$-measure ranges as a summary of precision and recall in the interval [77.53, 80.06]. Since there is a trade-off between recall and precision, $F$-measure looks smooth, and changes in a small range.

Training the algorithm with 39 manually picked particles took 3.5 s. After training, the algorithm needs 1 s to suggest new particles on each new micrograph and 2.5 s to retrain after being corrected by the user. Picking particles from 50 micrographs of the test dataset took 51 s, without any parallelization (however, current Xmipp implementation can benefit from multiple CPUs by concurrently picking different micrographs).

According to the 3DEM benchmark, our previous APP method was capable of selecting particles with an average time of 47 s, precision rate 80.94% and recall rate 68.59%. For the new algorithm, the precision and recall rates are reported as
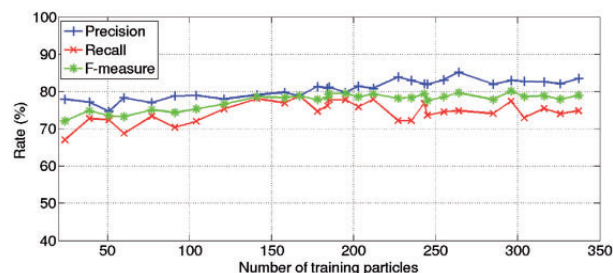


**Fig. 7.** Three accuracy metrics for the automatic selection of particles from the KLH test dataset according to the number of particles used to train. The vertical axis shows the values for precision (blue), recall (red) and $F$-measure and the horizontal axis shows the number of training particles
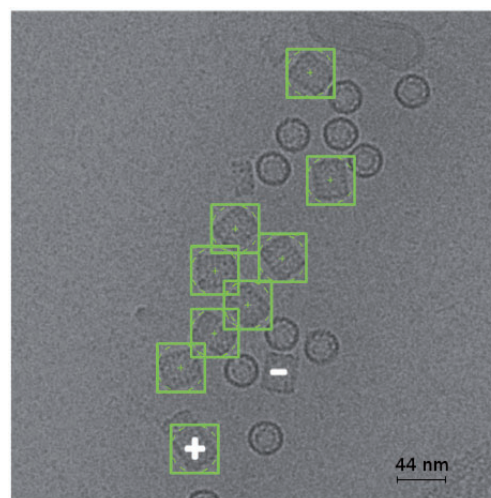


**Fig. 8.** The result of the algorithm for one micrograph of the KLH dataset with nine particles. Nine objects are boxed in green as particles. There is one false positive and one false negative, which are identified by + and −, respectively

85.16% and 74.81%, respectively, and the average processing time for each micrograph is 1 s. Therefore, the proposed algorithm is very fast and produces accurate results compared with our previous algorithm (Sorzano *et al.*, 2009) as well as with the reported results in the APP challenge (Zhu *et al.*, 2004).

Figure 8 shows the result of the algorithm for one micrograph of the test dataset with nine particles. In this figure, nine green squares show the selected objects, from which one long particle with '+' symbol is a false positive. One particle, marked with '−' symbol, is missed from the final result.

### 3.2 Adenovirus dataset

In this experiment we examined the reliability of our method by means of a set of micrographs (Pérez-Berná *et al.*, 2009) that has lower contrast but higher density and a larger particle size than the KLH dataset (Fig. 9).
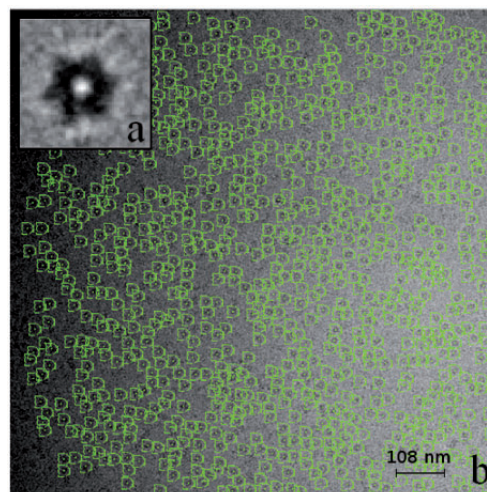
**V.Abrishami et al.**
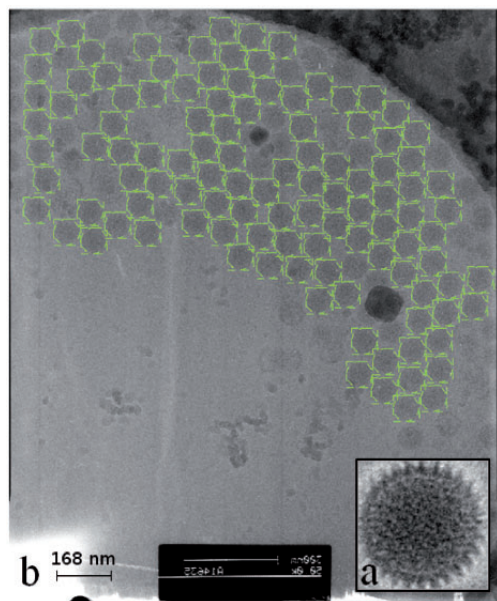
**Fig. 9.** Automatic selection for one micrograph of the adenovirus dataset, (**a**) The template for cross-correlation. (**b**) The micrograph with 115 automatically selected particles



**Fig. 10.** Automatic selection for the third micrograph of the helicase dataset, (**a**) The template for cross-correlation; (**b**) 874 automatically picked particles from the micrograph

Adenovirus type 2 is an icosahedral mammalian virus with a molecular mass of 150 MDa, and consists of genomic DNA and structural proteins. Samples were vitrified in liquid ethane, and a FEI Tecnai G2 FEG microscope at working voltage of 200 kV was used to analyse the samples. Micrographs were recorded on a film at a magnification of 5000×, and digitized in a Zeiss Photoscan TD scanner using a step size of 7 μm, which provided a sampling rate of 1.4 Å (particle size at this sampling rate is 600 pixels). True particles for this dataset were manually identified by Pérez-Berná *et al.* (2009).

Like in the previous experiment, the original 305 micrographs are divided into a training dataset with 30 micrographs and a test dataset with 275 micrographs (these two datasets can be obtained from http://i2pc.cnb.csic.es/3dembenchmark/). The sizes of the micrographs are not the same, but we know that $4000 < height < 15\,000$ and $4000 < width < 12\,000$ pixels. The dimensions of the micrographs were internally reduced by a factor of 4. We examined 90% of the local correlation maxima, and assumed that one template was enough for the cross-correlation.

We used 334 particles and 690 non-particles from the first 14 micrographs to fully train the classifier. The classifier was used to automatically pick the particles of the test dataset. This experiment resulted in picking 9216 particles with precision rate 92.17% and recall rate 90.24%. Figure 9b shows the result of the automatic picking for one micrograph of this dataset. As it can be seen in this figure, the algorithm picks 115 particles from this micrograph and particles close to or inside the carbon parts are rejected.

The average processing time per micrograph was 34 s, which is more than for the KLH dataset because it is more dense (90% of the local maxima are checked) and the size of the micrographs

and particles were twice and three times larger than the KLH dataset ones, respectively.

### 3.3 Helicase dataset

In this section we work with the complex of a helicase and its loading factor with a total molecular mass of 0.39 MDa (V.A. *et al.,* unpublished data). This dataset consists of 160 micrographs of size $4046 \times 4046$. The CCD of a JEOL JEM-2200FS microscope with magnification 50 000× and voltage 200 kV provided digital micrographs with pixel size 2.16 Å.

The rather dense micrographs of this dataset (the average number of particles in each micrograph is 572) (Fig. 10b) as well as the small particle size make them particularly difficult. Additionally, contrast is especially low due to the cryo-EM conditions. This in-house dataset is not published yet, and no benchmark is available in order to check the accuracy quantitatively, therefore the result is given qualitatively.

We set the size of particles to 100 pixels, and processed 90% of the local maxima to ensure that no particle was missed. To train the classifier, 1037 positive and 2510 negative samples were extracted from the first two micrographs, and then for the third micrograph the algorithm automatically picked 874 particles in 50 s. Figure 10 shows the result for the third micrograph of this dataset, qualitatively in agreement with user expectations.

### 4 DISCUSSION AND CONCLUSION

In this article, we proposed an APP algorithm that identifies particles from electron micrographs more accurately and faster compared to our previous method by Sorzano *et al.* (2009), already one of the best performing methods. A set of robust shape-related and statistical features are extracted from particle candidate locations (that are distinguished by cross-correlation between the templates and micrograph) and a two-stage classifier decides if each feature vector is a particle or not. Like in

Sorzano *et al.* (2009), the learning process of the classifier is continuous to grab the features of the user desired particles, and the second stage of the classifier is used to concentrate just on making distinctions between particles and errors, which are very similar to particles. Experimental results for three datasets show that this algorithm is able to select particles accurately even from low-contrast and highly dense micrographs.

The feature vector includes three types of features. This feature vector is robust to noise and has a high discrimination power, so that the classifier can distinguish between particles and non-particles. Having two types of features for shape description at the same time helps us to reduce the probability of producing partially similar feature vectors for particles and non-particles. For instance, in case of the KLH dataset, the descriptors at different frequencies are not adequate to properly distinguish between side and top views of the particle. On the other hand, the calculated shape descriptors from rotational invariant subspace are not sufficient to separate errors very similar to the side views of the particle. As a summary, these two types of features conspire efficiently to decrease the false positive rate. Statistical features are also important to catch the properties of the intensity distribution, and prevent the APP from picking particles from the carbon parts of micrographs. The role of this type of feature is clearer for helicase and adenovirus datasets, where more particles lie in dark areas of the micrographs. In Supplementary material, the classification power of the individual features is assessed in depth.

To decrease false positives as much as possible, a two stage SVM classifier is used to classify the feature vectors: the first stage to separate particles and non-particles, and the second stage to remove errors (very similar to particles) from the output of the previous stage. Each SVM component of the classifier is capable of performing the separation with an accuracy of >90% (see Supplementary material). The second component of the classifier plays an important role in reducing the false positive rate (e.g. 15% of wrongly selected particles were discarded by the second classifier in the KLH dataset).

There are six parameters that can be set by the user: number of sub-bands, PCA basis, rotational PCA basis, consecutive sub-bands to be correlated, templates and percentage of local maxima in correlation map to keep. Although these parameters help the user to achieve the highest possible accuracy, they can result in complexity. To moderate this complexity, the value of the first four parameters is set by default, so that the user can focus on adjusting the last two parameters (number of templates and percentage of local maxima to keep).

Regarding the particularities of the datasets, KLH is a dataset with highly contrasted particles, but there are two sources of errors that make it challenging. First, a few background objects and also noisy top views that present features similar to the ones of the particles. Second, KLH can polymerize to some degree. The second classifier is in charge of removing these polymerized particles, and keeping, at the same time, the recall rate at a reasonable level.

For the adenovirus dataset, our algorithm achieved a better accuracy than for the KLH dataset. The reason is that, although the micrographs are denser and have samples with lower contrast, the similarity between particles and non-particles is not as high as in the KLH dataset. The second classifier rejects just 3%

of the outputs of the first one. Difficult cases in this dataset are those particles located on carbon areas, which are efficiently distinguished by statistical features.

The speed of the algorithm was examined for the three datasets. According to the density of micrographs, it can go from 1 s to 50 s. Most computations are related to feature extraction, especially the shape descriptors. In order to eliminate this bottleneck and improve the speed even more, in our implementation particle candidates are divided between different threads. In addition to this, the automatic selection of particles from micrographs is performed in parallel. In this way, our APP is extremely fast and can be executed in a very short time.

This algorithm is included in Xmipp 3.0 and downloadable from http://xmipp.cnb.csic.es. The APP is accessible through the protocols described by Scheres *et al.* (2008).

*Conflict of Interest*: none declared.

## REFERENCES

Arbeláez,P. *et al.* (2011) Experimental evaluation of support vector machine-based and correlation-based approaches to automatic particle selection. *J. Struct. Biol.*, **175**, 319–328.

Boser,B.E. *et al.* (1992) A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory, COLT '92.* ACM, New York, NY, USA, pp. 144–152.

Chang,C.-C. and Lin,C.-J. (2011) Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, **2**, 27:1–27:27.

Chen,J.Z. and Grigorieff,N. (2007) Signature: a single-particle selection system for molecular electron microscopy. *J. Struct. Biol.*, **157**, 168–173.

Cortes,C. and Vapnik,V. (1995) Support-vector networks. *Mach Learning*, **20**, 273–297.

Dube,P. *et al.* (1993) The portal protein of bacteriophage SPP1: A DNA pump with 13-fold symmetry. *EMBO J.*, **12**, 1303–1309.

Glaeser,R.M. (1971) Limitations to significant information in biological electron microscopy as a result of radiation damage. *J. Ultrastruct. Res.*, **36**, 466–482.

Hall,R.J. and Patwardhan,A. (2004) A two step approach for semi-automated particle selection from low contrast cryo-electron micrographs. *J. Struct. Biol.*, **145**, 19–28.

Henderson,R. (1995) The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *J. Mol.r Biol.*, **247**, 726–738.

Huang,Z. and Penczek,P.A. (2004) Application of template matching technique to particle detection in electron micrographs. *J. Struct. Biol.*, **145**, 29–40.

Langlois,R. and Frank,J. (2011) A clarification of the terms used in comparing semi-automated particle selection algorithms in cryo-em. *J. Struct. Biol.*, **175**, 348–352.

Langlois,R. *et al.* (2011) Reference-free particle selection enhanced with semi-supervised machine learning for cryo-electron microscopy. *J. Struct. Biol.*, **175**, 353–361.

**V.Abrishami et al.**

Ludtke,S.J. *et al.* (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, **128**, 82–97.

Mallick,S.P. *et al.* (2004) Detecting particles in cryo-em micrographs using learned features. *J. Struct. Biol.*, **145**, 52–62.

Nicholson,W. and Glaeser,R. (2001) Review: Automatic particle detection in electron microscopy. *J. Struct. Biol.*, **133**, 90–101.

Ogura,T. and Sato,C. (2004) Automatic particle pickup method using a neural network has high accuracy by applying an initial weight derived from eigen-images: a new reference free method for single-particle analysis. *J. Struct. Biol.*, **145**, 63–75.

Pearson,K. (1901) On lines and planes of closest fit to systems of points in space. *Philos. Mag.*, **2**, 559–572.

Plaisier,J.R. *et al.* (2004) TYSON: Robust searching, sorting, and selecting of single particles in electron micrographs. *J. Struct. Biol.*, **145**, 76–83.

Ponce,C. and Singer,A. (2011) Computing steerable principal components of a large set of images and their rotations. *IEEE T. Image Process.*, **20**, 3051–3062.

Pérez-Berná,A.J. *et al.* (2009) Structure and uncoating of immature adenovirus. *J. Mol. Biol.*, **392**, 547–557.

Roseman,A. (2003) Particle finding in electron micrographs using a fast local correlation algorithm. *Ultramicroscopy*, **94**, 225–236.

Schatz,M. and van Heel,M. (1990) Invariant classification of molecular views in electron micrographs. *Ultramicroscopy*, **32**, 255–264.

Scheres,S.H.W. *et al.* (2008) Image processing for electron microscopy single-particle analysis using xmipp. *Nat. Protoc.*, **3**, 977–990.

Sigworth,F.J. (2004) Classical detection theory and the cryo-em particle selection problem. *J. Struct. Biol.*, **145**, 111–122.

Sorzano,C.O.S. *et al.* (2004) Normalizing projection images: A study of image normalizing procedures for single particle three-dimensional electron microscopy. *Ultramicroscopy*, **101**, 129–138.

Sorzano,C.O.S. *et al.* (2009) Automatic particle selection from electron micrographs using machine learning techniques. *J. Struct. Biol.*, **167**, 252–260.

Sorzano,C.O.S. *et al.* (2010) A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Struct. Biol.*, **171**, 197–206.

Sorzano,C.O.S. *et al.* (2012) Semiautomatic, high-throughput, high-resolution protocol for three-dimensional reconstruction of Single Particles in Electron Microscopy. In: *Nanoimaging: Methods and Protocols. Methods in Molecular Biology.* Humana Press Inc., New York, NY, pp. 171–193.

Volkmann,N. (2004) An approach to automated particle picking from electron micrographs based on reduced representation templates. *J. Struct. Biol.*, **145**, 152–156.

Wong,H.C. *et al.* (2004) Model-based particle picking for cryo-electron microscopy. *J. Struct. Biol.*, **145**, 157–167.

Yu,Z. and Bajaj,C. (2004) Detecting circular and rectangular particles based on geometric feature detection in electron micrographs. *J. Struct. Biol.*, **145**, 168–180.

Zhao,J. *et al.* (2013) Tmacs: A hybrid template matching and classification system for partially-automated particle selection. *J. Struct. Biol.*, **181**, 234–242.

Zhu,Y. *et al.* (2003) Automatic particle detection through efficient hough transforms. *IEEE T. Med. Imaging*, **22**, 1053–1062.

Zhu,Y. *et al.* (2004) Automatic particle selection: results of a comparative study. *J. Struct. Biol.*, **145**, 3–14.

Contents lists available at ScienceDirect

# Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

# A fast iterative convolution weighting approach for gridding-based direct Fourier three-dimensional reconstruction with correction for the contrast transfer function

V. Abrishami [a], J.R. Bilbao-Castro [c], J. Vargas [a], R. Marabini [a,b], J.M. Carazo [a],
C.O.S. Sorzano [a,d,*]

[a] *Biocomputing Unit, Centro Nacional de Biotecnología-CSIC, C/Darwin 3, Cantoblanco, 28049 Madrid, Spain*
[b] *Escuela Politécnica Superior, Universidad Autónoma de Madrid, C/Francisco Tomás y Valiente, Cantoblanco, 28049 Madrid, Spain*
[c] *Unidad Asociada al CSIC, 'Supercomputación y Algoritmos', Universidad Almería & Málaga, Spain*
[d] *Bioengineering Lab, Escuela Politecnica Superior, University San Pablo CEU, Boadilla del Monte, 28668 Madrid, Spain*

## ARTICLE INFO

## ABSTRACT

We describe a fast and accurate method for the reconstruction of macromolecular complexes from a set of projections. Direct Fourier inversion (in which the Fourier Slice Theorem plays a central role) is a solution for dealing with this inverse problem. Unfortunately, the set of projections provides a non-equidistantly sampled version of the macromolecule Fourier transform in the single particle field (and, therefore, a direct Fourier inversion) may not be an optimal solution. In this paper, we introduce a gridding-based direct Fourier method for the three-dimensional reconstruction approach that uses a weighting technique to compute a uniform sampled Fourier transform. Moreover, the contrast transfer function of the microscope, which is a limiting factor in pursuing a high resolution reconstruction, is corrected by the algorithm. Parallelization of this algorithm, both on threads and on multiple CPU's, makes the process of three-dimensional reconstruction even faster. The experimental results show that our proposed gridding-based direct Fourier reconstruction is slightly more accurate than similar existing methods and presents a lower computational complexity both in terms of time and memory, thereby allowing its use on larger volumes. The algorithm is fully implemented in the open-source Xmipp package and is downloadable from http://xmipp.cnb.csic.es.

## 1. Introduction

Single-Particle Analysis (SPA) is an Electron Microscopy (EM) method wherein the three-dimensional (3D) structure of a biological complex is determined from projections at random orientations of multiple instances of the specimen. Each projection is a two-dimensional (2D) projection of the 3D complex with a random spatial orientation that is additionally modulated by the Contrast Transfer Function (CTF) of the microscope. Upon determination of the orientation parameters, an inversion procedure yields a 3D volume that is compatible with the original projections. However, noisy imaging conditions, CTF effects, errors in orientation parameters, and a finite number of discrete projections not covering the whole spatial domain under study makes this

inversion problem nontrivial [1]. Many approaches have been proposed to solve this ill-posed inversion, which can be categorized into three classes: algebraic, Weighted Back-Projection (WBP), and direct Fourier methods.

Algebraic methods treat this inversion problem as a system of linear equations where well-established algebra methods are employed to find the solution. In other words, the problem is formulated as $\bar{p} = W\bar{v}$, where the 3D object is decomposed into a finite set of basis functions whose coefficients are lexicographically stored in a vector $\bar{v}$, $\bar{p}$ is a vector with the values of all of the projections' pixels, and $W$ encodes the weight of each of the basis functions onto each pixel. The Algebraic Reconstruction Technique (ART) [2] and Simultaneous Iterative Reconstruction Technique (SIRT) [3] are iterative approaches to solve this system of equations. The general idea behind these methods is to iteratively improve an initial volume by comparing each of the experimental projections with the projections from the current volume, thereby attempting to compensate for this difference. Although algebraic

* Corresponding author at: Biocomputing Unit, Centro Nacional de Biotecnología-CSIC, C/Darwin 3, Cantoblanco, 28049 Madrid, Spain. Fax: +34 91 585 4506.
*E-mail address:* coss@cnb.csic.es (C.O.S. Sorzano).

methods have the potential to be applied to many different types of reconstructions [1] and to incorporate a variety of constraints [4], they suffer from a high computational complexity. In the SPA field, the introduction of blobs as spherically symmetric basis functions by [5] was one of the most efficient efforts towards a fast algebraic reconstruction, but it is still more computationally expensive compared to the two other aforementioned groups.

WBP [6,7] is the most intuitive method used to reconstruct a 3D object from its 2D projections based on the concept of the Fourier slice theorem. For a set of projections with known orientations, each projection is back-projected across the objective volume from its position in the projection space defined by its orientation. The superposition of all of these back-projections provides an estimation of the original 3D object. To account for the angular distribution of projections, these methods use a weighting function in Fourier space. Weighted back-projection methods are faster than algebraic methods, but they perform poorly in cases where large angular gaps exist, and they have been shown to underperform algebraic methods in a number of cases [5,8].

Based on the central slice theorem, direct Fourier reconstruction (DFR) methods try to obtain the 3D Fourier transform of an object directly from the 2D Fourier transform of its projections, so that an estimation of the original 3D object can be quickly obtained through an inverse 3D Fast Fourier Transform (FFT). In practice, the irregularity of the spatial distribution of the frequency of samples in the set of projections in experimental SPA studies makes the direct use of the inverse FFT unfeasible. Thus, an additional interpolation step is required to obtain the 3D Fourier transform of the object on a regular grid. The so-called gridding algorithm is an alternative method introduced by Penczek et al. [9] into the SPA field; this method was originally developed by Jackson et al. [10] to efficiently estimate the 3D Fourier transform in a regular grid of points using irregularly distributed samples in Fourier space. This algorithm uses an interpolation kernel; in our case this kernel is a modified Kaiser-Bessel (MKB) window function (also known as a blob). The gridding-based direct Fourier method can yield resolutions higher than the algebraic methods (and clearly weighted back-projection methods) in a fraction of their computing time.

Frequency samples from different projections mainly concentrate at the center of the 3D frequency domain, and their sparsity increases as we move away from the center. If no weighting scheme were employed, samples close to the 3D Fourier origin would be over-represented with respect to points away from the origin. In general, the sample values must be corrected by a weight function before inverting the Fourier transform of the volume. An accepted method in SPA to perform this weighting task is to use the volume of the Voronoi region [9] around each Fourier sample. This region is a polyhedron associated with each sample such that the distance between this sample and any point in the polyhedron is shorter than the distance from these points to any of the remaining samples. However, computing Voronoi cells is time consuming (particularly for SPA, where the algorithm has to handle millions of Fourier samples).

An alternate algorithm for obtaining the weighting function was proposed by Matej and Lewitt [11] for Positron Electron Tomography (PET). This approach seeks appropriate weights at each Fourier sample so that it participates with the right weight during the interpolation of regular points. The algorithm begins with the initial weights of the samples and uses convolution with a kernel to iteratively refine these weights. Their proposed method is practical for PET where the geometry of data is known and the number of sampling points is not large, but not for SPA where such conditions are not met.

In this paper, we introduce a gridding-based direct Fourier three-dimensional reconstruction in SPA following the method suggested by Matej and Lewitt [11] in PET. Our method follows the same iterative scheme for computing the weights, but we estimate the weights at each Fourier sample by evaluating a function instead of storing the collection of weights (which would become impractical in SPA). The proposed approach follows the general idea of Scheres [12], but differs in the way that the weights are calculated. Finally, our algorithm has an additional novel step to compensate for the trilinear interpolation of weights in Fourier space (described in Section 2.4), which improves the resolution of the final reconstruction. The CTF correction is applied during 3D reconstruction and is crucial for a high-resolution structure determination.

We compared the proposed reconstruction algorithm with the algorithms from the SPARX package [13] and the RELION package [12]. In the method by Zhang et al. [13], the projections are first padded to 2 times their original size (default value in the implementation), and then the Nearest Neighbor (NN) interpolation is used to calculate the target 3D Fourier volume. Finally, a weighting function using Bracewell's "local density" [14] is computed to correct the value for each voxel of the 3D Fourier volume.

The experiments showed that our approach is a superior method for 3D reconstruction in terms of accuracy, speed, and memory usage. The new algorithm is fully implemented in the open-source Xmipp package and is downloadable from http://xmipp.cnb.csic.es.

## 2. Material and methods

### 2.1. Preliminaries

The goal of the gridding-based direct Fourier method is to approximate frequency samples on a regular 3D Cartesian lattice $F_{3D}(\bar{R})$ from the measured samples of the 3D frequency domain $\hat{F}_{3D}(\bar{Q})$ as

$$F_{3D}(\bar{R}) = \int CTF^{-1}(\bar{Q})\hat{F}_{3D}(\bar{Q})K(\bar{R} - \bar{Q})d\bar{Q}, \qquad (1)$$

where $\bar{R}$ is the frequency coordinate within the regular 3D grid and $K$ is the kernel function by which the integration is accomplished. We recommend using a kernel function with some appealing attributes, such as finite size, bell-shaped decay, and differentiability at the borders. The modified Kaiser-Bessel (MKB or blob) is considered to be the best kernel for gridding interpolation by several authors [10,11,15,16]. Matej and Lewitt [16] generally assessed the optimal values for the parameters of the MKB to achieve a reconstruction with good quality. We use an MKB with the same parameter values suggested in their paper. CTF correction is incorporated during interpolation by dividing each irregular sample by $CTF^{-1}(\bar{Q})$, where $CTF(\bar{Q})$ is the value of the CTF at frequency $\bar{Q}$ (as a practical implementation issue, this division is performed as long as the CTF is above a given threshold).

Under experimental conditions, a limited number of projections from the specimen are available. Therefore, the discrete form of the integral in Eq. (1) should be considered because measurements are only available at a finite set of frequencies $\overline{Q} \in \{\bar{R}_i\}$. To obtain a discrete form of Eq. (1), the integral is substituted by a summation as shown

$$F_{3D}(\bar{R}) = \sum_i CTF^{-1}(\bar{R}_i)\hat{F}_{3D}(\bar{R}_i)K(\bar{R} - \bar{R}_i)w(\bar{R}_i), \qquad (2)$$

where $w(\bar{R}_i)$ is the weighting factor for the $i$-th irregular sample. It is important to note that the weighting function is a substitution for $d\bar{Q}$ within Eq. (1). In fact, the value of each irregular sample should be corrected by the weighting factor and the related CTF

value before participating in the interpolation of the points on the 3D regular grid because the distribution of the samples in Fourier space is non-uniform. If no weighting scheme was applied, the regular points close to the origin would contribute more to the summation and the points further away would contribute less, generating undesired artifacts and yielding the wrong reconstructions. In Penczek et al. [9], this weight is proportional to the volume of the Voronoi regions around each experimental Fourier sample.

In the following sections, we describe in detail our iterative algorithm for obtaining the weighting function.

## 2.2. Calculation of the weighting function

In this section, we propose a numerical iterative approach for computing the weighting function $w(\bar{R})$ based on the method introduced by Matej and Lewitt [11] for PET. In their method, the weighting coefficients are computed for each sample in the 3D frequency space. Because the PET scanner geometry is fixed, this calculation can be performed once and the result can be stored for subsequent reconstructions. Additionally, due to the scanner characteristics the number of Fourier samples is relatively low compared to the number of samples in 3DEM. However, this algorithm is not applicable to SPA, where hundreds of thousands of projections of several hundred pixels may exist per side with no prior knowledge about their angular distribution. Thus, the algorithm has to compute the weighting coefficients for a large number of samples (in the order of billions) for each new set of projections, which is not efficient in terms of both time and storage complexity. Our proposed method is adapted to SPA to overcome these issues.

The proposed approach begins with an initial weighting function that is initialized as a constant function of value 1 as

$$w^{(0)}(\bar{R}) = 1. \tag{3}$$

Then, this weighting function is refined through a number of iterations by evaluating Eq. (2) with $F_{3D}(\bar{R}_i) = 1$ at the locations specified by the projection directions until the reconstructed function $F_{3D}(\bar{R})$ is as close to 1 as desired (formally that the Chebyshev norm of the function $F_{3D}(\bar{R}) - 1$ is smaller than $\varepsilon$, $\|F_{3D}(\bar{R}) - 1\|_\infty < \varepsilon$). Then, the refined weighting function can efficiently remove the effect of non-uniformity by multiplying it by the estimated Fourier Transform of the 3D object.

To iteratively refine the weights, we first evaluate Eq. (2) to compute $c^{(n+1)}(\bar{R})$ at the position of each point on the 3D regular grid, as below

$$c^{(n+1)}(\bar{R}) = \sum_i w^{(n)}(\bar{R}_i) K(\bar{R} - \bar{R}_i), \tag{4}$$

where $w^{(n)}(\bar{R}_i)$ is the interpolated weight from $w^{(n)}(\bar{R})$ of the $n$-th iteration at the position of the $i$-th input sample. We use trilinear interpolation to interpolate $w^{(n)}(\bar{R}_i)$ from $w^{(n)}(\bar{R})$. $c^{(n+1)}(\bar{R})$ is the evaluation of Eq. (2) when all samples have a value of 1 and there is no CTF.

After computing Eq. (4) at the position of each frequency point on the 3D regular grid, the algorithm updates the weights from the previous step by dividing them by the obtained convolution in Eq. (4):

$$w^{(n+1)}(\bar{R}) = \frac{w^{(n)}(\bar{R})}{c^{(n+1)}(\bar{R})}. \tag{5}$$

After a number of iterations, the points in $c^{(n)}(\bar{R})$ converge to values close to one and the weighting function reaches stability. At this point, $w^{(n)}(\bar{R})$ can be used to correct the values of regular

samples to overcome the non-uniformity problem. Note that $w^{(1)}(\bar{R}) = \frac{1}{\sum_i K(\bar{R} - \bar{R}_i)}$. Substituting this value into Eq. (2) results in

$$F_{3D}(\bar{R}) = \frac{\sum_i F_{3D}(\bar{R}_i) K(\bar{R} - \bar{R}_i)}{\sum_i K(\bar{R} - \bar{R}_i)}$$

that is the well-known formula for a kernel interpolator. In this way, we see that our method goes beyond kernel interpolation by further guaranteeing that the different weights result in a constant interpolated function when the input samples are constant.

Note that Eq. (5) is similar to Eq. (A3) in Scheres [12]. However, there are some important differences: (1) the motivation of Eq. (5) and Eq. (A3) of Scheres [12] is different. In our algorithm the weights are constructed as a way of discretizing the convolution in Eq. (1), while in Eq. (A3) of Scheres [12] the weights are meant to represent the inverse of the denominator in Eq. (3) of Scheres [12]; (2) Eq. (A3) has a term $\Omega$ that is far from trivial and is not present in Eq. (5); (3) the starting values of both sets of weights ($w^{(0)}(\bar{R})$ in our algorithm and the corresponding starting point in Relion) are different, so that both sets of weights become increasingly dissimilar as the iterations increase (due to reason 2, the iterative step is also different); and (4) our iterative scheme normally converges in 2–3 iterations, while the scheme in Relion is normally run for 10 iterations, probably reflecting a larger distance between the starting and finishing sets of weights.

## 2.3. Compensation for the convolution kernel

Eq. (1) represents a convolution in Fourier space of a function given by its samples $F_{3D}(\bar{R}_i)$ and the kernel $K(\bar{R})$. This is equivalent to a multiplication in real space of the function $f_{3d}(\bar{r})$ and the inverse Fourier transform of the kernel $k(\bar{r})$. This means that we need to divide by $k(\bar{r})$ to recover the function representing the macromolecular complex. This is a well-known effect that is already handled by the standard gridding algorithm. In our implementation, the kernel function in Fourier space is given by the MKB function with parameters $a$ (the width of the kernel), $\alpha$ (a smoothness parameter), $m$ (the order of the Bessel function), and $n$ (the dimension in which the kernel is defined, in our case $n=3$), whose definition is

$$K(R) = \begin{cases} \dfrac{\left(\sqrt{1 - \left(\frac{R}{a}\right)^2}\right)^m I_m\left(\alpha\left(\sqrt{1 - \left(\frac{R}{a}\right)^2}\right)\right)}{I_m(\alpha)} & 0 \le R \le a \\ 0 & R > a \end{cases}$$

Note that $R$ is the module of the frequency vector $\bar{R}$ and $I_m$ is the modified Bessel function of the 1st type and order $m$. The inverse Fourier transform of the kernel is

$$k(r) = \begin{cases} \dfrac{(2\pi)^{\frac{n}{2}} a^n \alpha^m I_{\frac{n}{2}+m}\left(\sqrt{\alpha^2 - (2\pi a r)^2}\right)}{I_m(\alpha)\left(\sqrt{\alpha^2 - (2\pi a r)^2}\right)^{\frac{n}{2}+m}} & 0 \le 2\pi a r \le \alpha \\ \dfrac{(2\pi)^{\frac{n}{2}} a^n \alpha^m J_{\frac{n}{2}+m}\left(\sqrt{\alpha^2 - (2\pi a r)^2}\right)}{I_m(\alpha)\left(\sqrt{\alpha^2 - (2\pi a r)^2}\right)^{\frac{n}{2}+m}} & 2\pi a r > \alpha \end{cases}$$

where $J_m$ is the Bessel function of order $m$.

## 2.4. Compensation for the interpolation in Fourier space

It is important to note that we used trilinear interpolation to estimate $w(\bar{R}_i)$ from $w(\bar{R})$ in frequency space, which is equivalent to the convolution in 3D Fourier space by a trilinear kernel. Compensating for this interpolation kernel is extremely important because the 3D reconstruction is otherwise masked in real space

by the inverse Fourier transform of the interpolation kernel, resulting in a loss of intensity as the diameter of the macromolecule being reconstructed increases. None of the previously published works on Fourier gridding that performed any type of interpolation of the weights in Fourier space (i.e., Scheres [12]) mention this effect. To the best of our knowledge, the trilinear kernel does not have a known inverse Fourier transform, but can be approximated by a spherically symmetric kernel [17] whose shape in Fourier space is defined as

$$K_i(R) = \begin{cases} 1 - R & 0 \le R \le 1 \\ 0 & R > 1. \end{cases}$$

This convolution in Fourier space must be corrected in real space by dividing by the function

$$k_i(r) = \left( \frac{\sin(\pi r)}{\pi r} \right)^2.$$

### 2.5. Parallelization of the algorithm

The processing time of the algorithm depends on the number and size of the projections, which may amount to hundreds of thousands of projections with several hundred pixels per side for SPA. However, most of the processing steps can be accomplished independently on each projection image, which makes the algorithm naturally suitable for parallelization to achieve higher processing speeds. Parallelization of the algorithm has been performed at two levels: at the level of multiple processors distributed over several computing nodes using the Message Passing Interface (MPI) [18] and at the level of threads executing in the same node using the POSIX threads [19].

Three major computations are discernible in our proposed reconstruction algorithm: FT of projection images, FT of the 3D object on a regular grid, and calculation of the weighting function. The rest of this section explains our approach for parallelizing each of the aforementioned steps.

Considering a multicomputer and multithreaded scenario, our algorithm acts as follows. First, a master process is created at one of the cluster nodes (computer). This process (master) spawns other processes (slaves) at different nodes in the cluster using MPI tools. The master process is in charge of assigning tasks to the slaves and coordinating them. It is also responsible for gathering results from the slaves and building up the final, complete 3D reconstruction. Once the slaves are spawned, they initialize their local memory structures to hold 3D grids and FTs. Upon initialization, they spawn on their own a series of threads that will be used locally to make use of multiple processors and/or cores. Once these initialization steps are completed, the slave informs the master that it is ready to process. Then, the master communicates back to the slave the range of projections it will have to process. Finally, each slave will proceed independently to build a partial 3D regular grid based on this set of projections using threads to speed up the process.

The computation of the FT of each projection image can be performed independently without any interference, with an expected speedup for this step growing almost linearly with the number of available processor/cores. When executing in multi-threaded mode, each thread computes the FT of a projection image and retains it in its memory. Note that at this step, the threads do not continue with another projection image, but instead share their local FTs with the other threads to jointly compute the effect of their local irregular samples on the final interpolation of the 3D regular grid. Thus, each thread is responsible for processing a number of rows of the image and uses the MKB window function to estimate the impact of each sample in these rows on each

regular sample. After finishing with the local FTs for all threads, new projection images are loaded by threads to continue with the same process until no more projection images are available. It is worth mentioning that the number of rows to process can be reduced by avoiding exploration at certain resolutions. To ensure the highest performance, row processing is dynamically balanced. Usually, many more rows than threads will exist, but not all rows will necessarily represent the same workload (see previous sentence). If the number of rows is equally distributed to the existing threads, some threads could finish their assigned tasks while others are still running. Thus, rows to be processed are dynamically assigned as threads finish processing previously assigned rows. This mechanism must be arbitrated to avoid the same rows being processed by different threads. To allow this, synchronization is accomplished by defining a shared bit array in which each bit shows whether a row has already been processed.

Finally, the master will ask for the interpolated values of the 3D grid estimated by each node and sum them up to obtain the final interpolated 3D grid. This grid must now be weighted before proceeding to perform the inverse FT and obtain the final 3D reconstruction.

To obtain weights, after computing the 3D regular grid we use the same parallelization strategy to compute the weights on a regular grid. The only difference here is that now all of the nodes should retain the total weights at the position of the 3D regular grid to compute the convolution in Eq. (4). For this mechanism, each node computes the weights at the position of the regular grid points by mean of its projection images and waits until the rest of the nodes have also completed this task. At this point, all nodes combine their values orchestrated by the master node. Then, in the next iteration of weight refinement these corrected weights are used by the processors to compute Eq. (5).

Finally, the master node carries both the interpolated 3D regular grid and weighting function. Then, the final 3D reconstruction is obtained by applying the weighting function and an inverse FFT. The schema of our parallel framework for computing the interpolated 3D regular grid is presented in Fig. 1.

## 3. Results

We compared the proposed reconstruction method with two other well-known methods: (1) *NN direct* inversion implemented in SPARX [13] and (2) RELION DFR [12]. These methods were compared in terms of accuracy, memory complexity, and the time required for execution. Two asymmetric test objects were used to achieve this goal: first, a 70S ribosomal subunit as an instance of a small complex and second, a DNA-origami object as a good representative of a large complex. Different data sets (set of projections) with specific properties (i.e., noise in projections, noise in angular assignment, and CTF-affected) were made for each test object to examine the robustness of each method.

To evaluate the accuracy of each reconstruction method, we used the Fourier Shell Correlation (FSC) criterion calculated with respect to the ground truth, which is the standard quality measurement in single particle analysis. All experiments were run on a cluster with 28 nodes, each with two Intel Xeon E5405 running at 2 GHz and with 16 GB of RAM memory.

### 3.1. 3D reconstruction of the 70S ribosome

In this experiment, we compared the reconstruction results of the proposed method with the results from RELION and SPARX using different data sets that were generated from simulated projections with specific characteristics corresponding to the 70S ribosomal subunit (PDB ID: 3V2D) represented on a cube of size
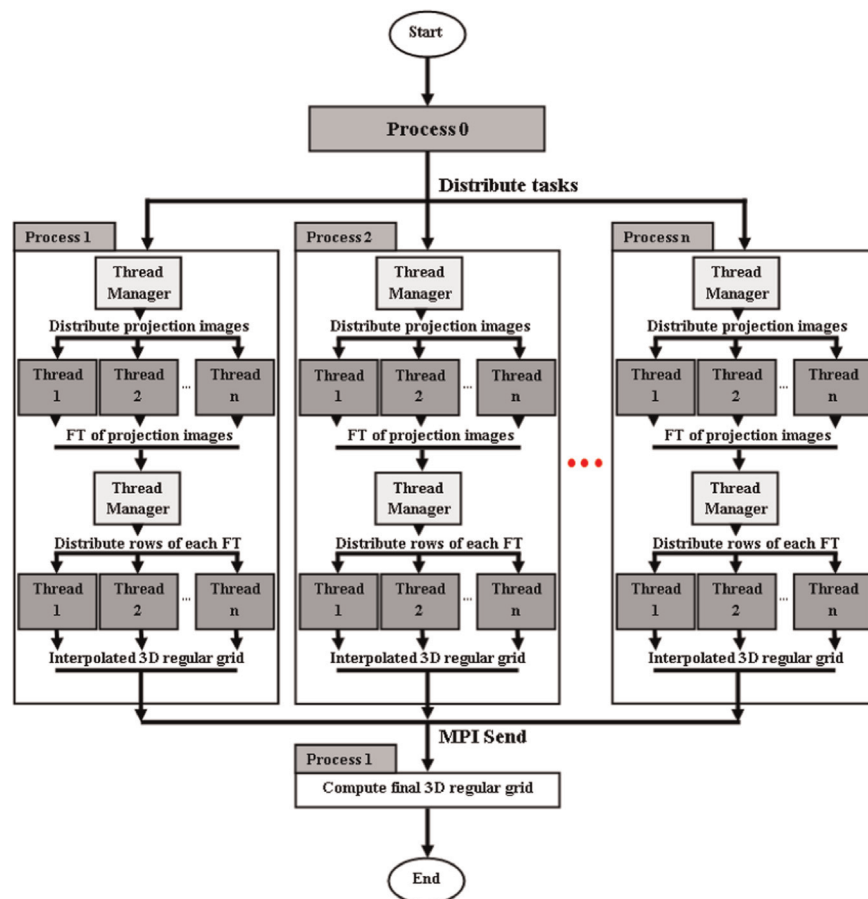
**Fig. 1.** Parallelization strategies. The two levels of parallelization of the proposed algorithm, including multiple threads and multiple processors are shown. The algorithm can be executed on several nodes of a cluster (multi-CPU parallelism) with several threads (multi-thread parallelism).

$174 \times 174 \times 174$ voxels (the sampling rate was 1.5 Å/pixel).

We applied the three reconstruction algorithms (with the exception of the first experiment) to five sets of simulated projections. The first data set is a set of noise-free projections of the test object, the second and third sets comprise projections contaminated by noise in angles and noise in pixel values, respectively, the fourth set contains a set of projections for coarser angular sampling (i.e., when the Crowther frequency is lower than the Nyquist frequency), and the fifth set contains CTF-affected projections.

We used default parameters for each method in these experiments. All methods use a padding factor of 2, and in our tests we set $\varepsilon = 0.01$.

*3.1.1. Projections without noise*

For this experiment, we applied five reconstruction algorithms to a set of $N=10,000$ free of noise projections: our proposed algorithm, RELION method [12], SPARX method [13], ART, and WBP. In Fig. 2, we show the FSC curves for the five reconstructions compared to the known ground truth (the 70S ribosome subunit at atomic resolution with a sampling rate of 1.5 Å/pixel). In this figure, the horizontal axis shows the spatial frequency and the vertical axis shows the FSC values. This figure shows that the RELION method reaches higher FSC values than the SPARX method, while our proposed gridding-based DFR method is slightly more accurate at all frequencies than either method. Although the difference is not large, it proves the validity of our approach. It is obvious from this figure that the RELION, SPARX and XMIPP methods are able to yield larger FSC values for higher resolution than the ART and WBP methods.
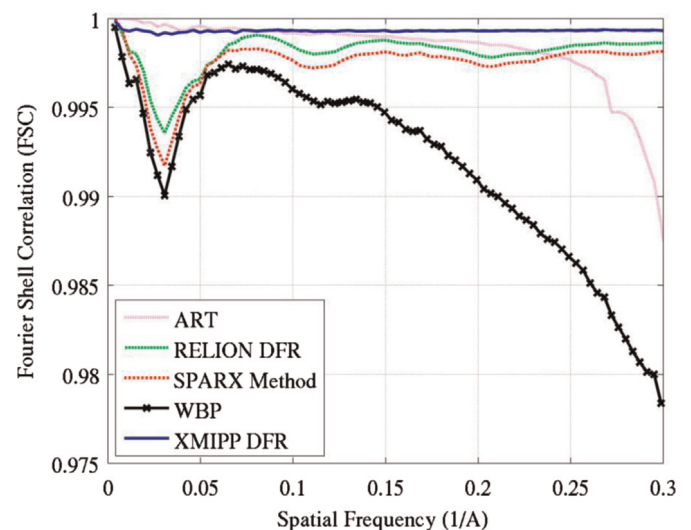


**Fig. 2.** Comparison of the reconstruction methods using noise-free projections of the 70S ribosome. FSC curves for the 3D reconstructions of the 70S ribosome using a set of 10,000 noise-free projections applying the following reconstruction algorithms: the proposed method (XMIPP), RELION method [12], SPARX method [13], ART and WBP.

For this reconstruction, the serial implementation of our proposed method requires 13 min and 40 s, while it takes 5 min and 37 s for RELION and 20 min and 44 s for SPARX using one core of the CPU. In Fig. 3, we show how increasing the number of threads can decrease the required computational time for our proposed method to one-fourth of the SPARX and half of the RELION
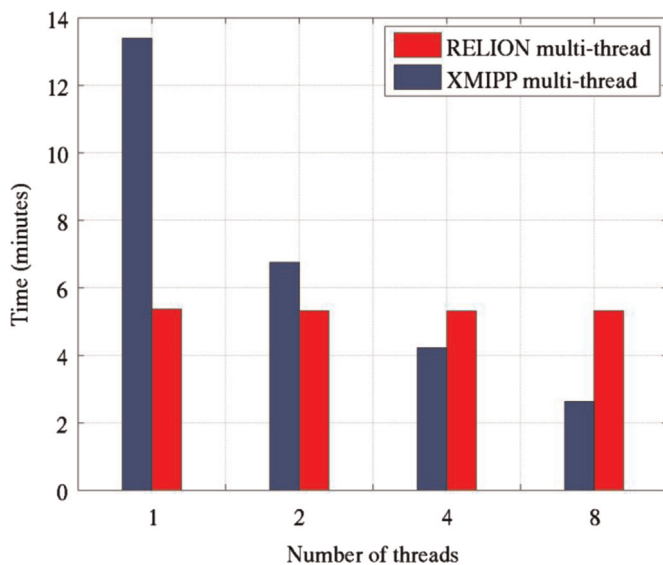
**Fig. 3.** Multi-thread implementation comparisons. Computational time required for the 3D reconstruction of the 70S ribosome from a set of $N=10,000$ noise-free projections applying the proposed (XMIPP) and RELION reconstructions using different numbers of threads.

reconstruction methods (2 min and 63 s using 8 CPUS). Note that although RELION supports threads, its parallel implementation does not improve the execution time significantly in this case (the threads simply compute the Fourier transform of the projections). As shown in this figure, increasing the number of threads reduces the computational time; however, more than eight threads make no or little improvement due to the required time for the initialization and synchronization of the threads.

In Fig. 4, we show the average execution times of 10 independent runs for the proposed method and the SPARX method (note that RELION implementation does not support multi-CPU parallelization) using different numbers of nodes (cluster computing parallelism). From this figure, it is evident that increasing the number of nodes improves the computational time for both methods. Indeed, increasing the number of nodes by a factor of two increases the speedup by a factor of two (which is ideal). It is obvious from this figure that the parallel implementation of our algorithm is faster than the parallel implementation of SPARX
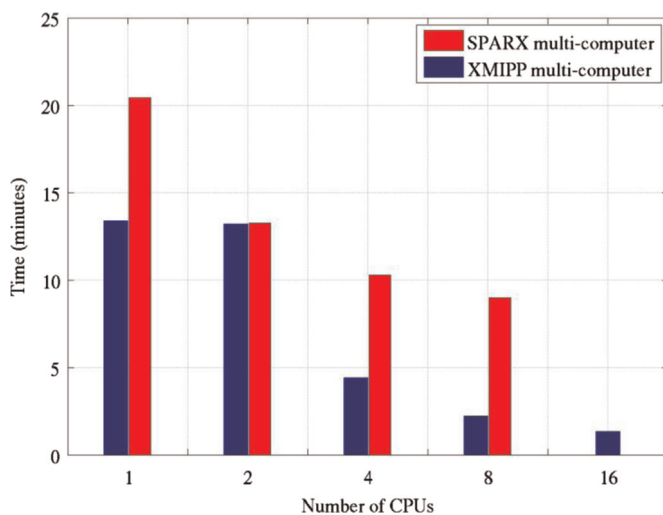


**Fig. 4.** Multi-computer implementation comparisons. The required computational time for the parallel execution of the proposed reconstruction method and the SPARX method for reconstructing the 70S ribosome using a set of 10,000 noise-free projections utilizing different numbers of CPUs.
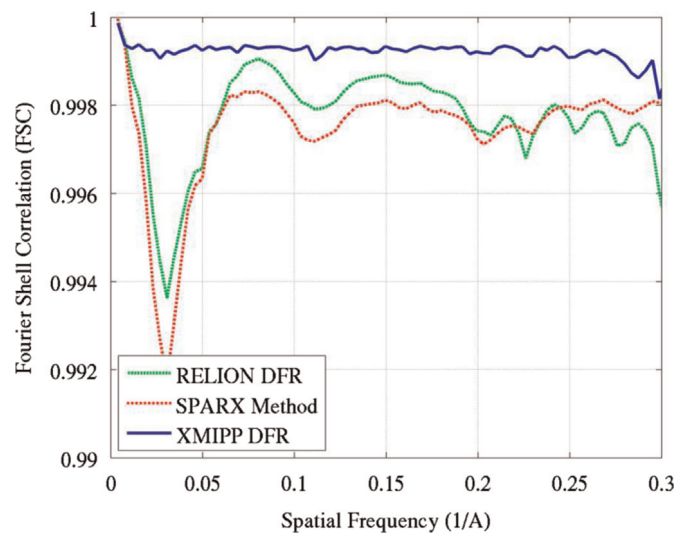


**Fig. 5.** Comparison of the reconstruction methods using CTF-affected projections of the 70S ribosome. FSC curves for the reconstructions of the 70S ribosome generated from $N=10,000$ projection images affected by 10 different sets of CTF parameters using our proposed reconstruction method (XMIPP), RELION method [12], and SPARX method [13]. The CTF parameters include: physical pixel size of 3.54 Å, microscope voltage of 300 kV, spherical aberration of 2.0 mm, amplitude contrast of 0.1, and different defocus values.

(approximately twice as fast for more than 2 CPUs because our implementation uses one node as a coordinator). Additionally our parallel implementation is scalable, while this does not apply to the parallel implementation of SPARX (increasing the number of CPUs from 4 to 8 provides approximately the same execution time).

### 3.1.2. CTF applied projections

In this experiment, we modified the entire $N=10,000$ simulated projections using 10 different sets of CTF parameters (each consequent 1000 particles have the same CTF parameters) with a physical pixel size of 3.54 Å, microscope voltage 300 kV, spherical aberration of 2.0 mm, amplitude contrast of 0.1, and different defocus values. We calculated the 3D maps for the three algorithms by enabling the CTF correction. The results for our method (XMIPP), RELION [12] and SPARX [13] with CTF correction are shown in Fig. 5. As seen in this figure, our algorithm can properly correct for the CTF, thereby maintaining the previously demonstrated pattern of providing slightly better results than the other two methods.

### 3.1.3. Projections with noise

In this test, which is closer to an experimental case, we introduced two different types of noise into the simulated projections to generate two sets of noise-corrupted projections. First, a Gaussian noise with a signal-to-noise ratio of approximately 0.8 was added to the pixels of the noise-free projections. Second, a Gaussian noise with zero mean and variance of $0.5°$ was applied to each projection direction to take into account angular inaccuracies. For each data set, $N=10,000$ projections were generated.

Similar to the previous experiment, we performed the reconstruction for each data set using our method (XMIPP), RELION [12] and SPARX [13] to evaluate the resilience of each reconstruction algorithm to noise. In Figs. 6 and 7, we compare the reconstruction results of these three methods in two situations: using projection images contaminated by noise in pixel intensities and using projection images contaminated by noise in angular assignments. From this figure, it is clear that there is a region of low frequency at which the FSC values are larger for the proposed
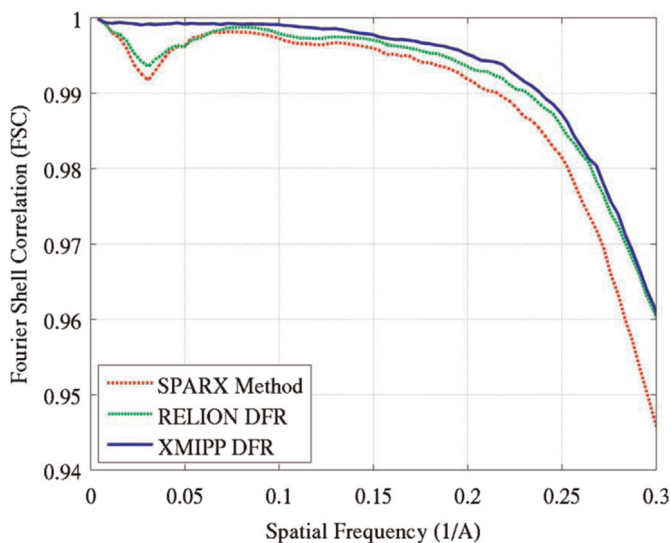
**Fig. 6.** Comparison of the reconstruction methods using projections of the 70S ribosome contaminated by pixel noise. FSC curves for the reconstructions of the 70S ribosome generated from $N=10{,}000$ projection images contaminated by pixel noise using our proposed reconstruction method (XMIPP), RELION method [12], and SPARX method [13].
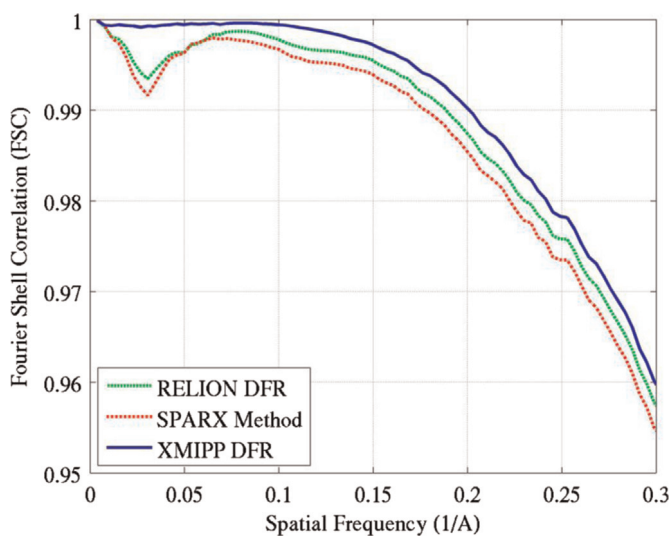


**Fig. 7.** Comparison of the reconstruction methods using projections of the 70S ribosome contaminated by angular noise. FSC curves for the reconstructions of the 70S ribosome generated from $N=10{,}000$ projection images contaminated by angular noise using our proposed reconstruction method (XMIPP), RELION method [12], and SPARX method [13].

method; at the same time, the proposed method is significantly faster than the SPARX method and faster than RELION when more than two threads are used.

### 3.1.4. Projections with coarser angular sampling

In this experiment, we used coarser angular sampling to compare the proposed method with two other methods when larger gaps are introduced into Fourier space. First, we generated $N=500$ evenly distributed projections (according to the Crowther frequency). For simplicity in the analysis, these projections were generated by fixing the azimuthal and in-plane rotation angles and simply changing the tilt angle. Three more datasets with $N=250$, $N=100$, and $N=50$ projections were generated in the same manner, and three reconstructions methods were applied to each data set as in the previous experiments. The obtained FSC curves resulting from each reconstruction method for these three data sets

are shown in Fig. 8. The FSC curves for the data set with 500 projections (Fig. 8a) are close to the curves shown in Fig. 2 (with the FSC valley at a low frequency for SPARX and RELION). However, when we decrease the number of projections to 250 (Fig. 8b), all of the methods show slightly lower values for the higher frequency, although the effect is more noticeable for the SPARX method. Finally, by further decreasing the number of projections to 100 and 50 (as shown in Fig. 8c and d, respectively), we notice that our reconstruction method shows remarkably higher FSC values (especially for higher frequencies) compared with the other two methods (note that RELION uses the same default values for gridding interpolation as the proposed method). These experiments show that our algorithm is more robust to gaps in the Fourier space.

### 3.2. 3D reconstruction of a 3D DNA-origami object

The purpose of this experiment is to determine how each reconstruction algorithm can manage memory for a large complex. For this purpose, the model of a discrete DNA object [20] (PDB ID: 2YMF) of size $400 \times 400$ (sampling rate 1.2 Å/pixel) was considered.

As in the previous experiment, we simulated $N=10{,}000$ projections at random orientations of this complex and used the proposed method (XMIPP), RELION method [12], and SPARX method [13] for the reconstruction. Due to the large size of this complex, neither the SPARX nor RELION reconstruction algorithms (with one CPU core) nor the parallel implementation of the proposed algorithm can be executed on our machine (with 16 gigabytes of RAM memory). However, we can take advantage of the multi-thread implementation of our algorithm to use the process resources efficiently without a lack of memory.

To accomplish this reconstruction, our method requires 1 h and 29 min with only one CPU core. Using our multi-threaded implementation, the execution times for different number of threads were computed as: 42 min and 72 s for two threads, 24 min and 83 s for four threads, and 15 min and 48 s for eight threads. Each execution time was obtained using the average of ten independent runs.

### 4. Conclusions

In this paper, we introduced a weighting function for gridding-based DFR in SPA based on the method published for PET by Matej and Lewitt [11] to discard the artifact connected with the non-uniformity of samples in the Fourier space. This accurate and fast weighting function corrects the values of frequency samples on a 3D regular grid to compensate for the uneven distribution of projection images before applying the final inverse FFT to recover the original 3D object. The CTF correction is performed during the reconstruction to correct for the effects of the CTF. The implementation of the proposed algorithm supports two parallelization frameworks: multi-thread parallelization and multi-computer parallelization. These parallelization frameworks allow our approach to efficiently take advantage of all cores of a node and nodes of a cluster to more quickly produce the final 3D reconstruction. When multi-computers run out of memory due to the large size of the complex (this happens when multiple MPI processes attempt to run on the same node because memory needs are multiplied by the number of processes), multi-thread implementation can still be employed to take advantage of multiple processing resources (here, the memory requirements, regardless of the number of threads, are the same as the requirements for a single MPI process). We compared our gridding-based DFR with two well-known methods in the field (the method of
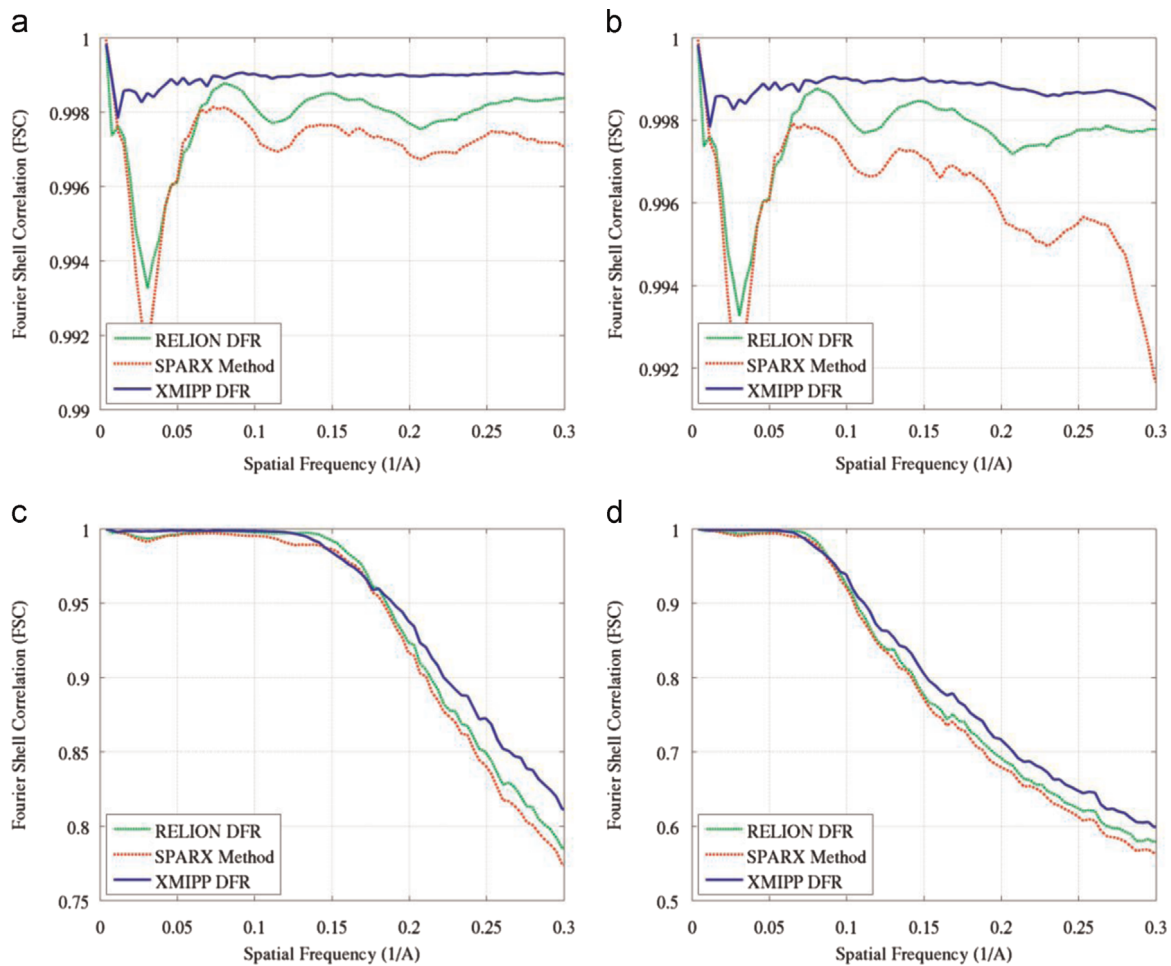
**Fig. 8.** Comparison of the reconstruction methods using a set of projections with larger angular sampling of the 70S ribosome. FSC curves for the reconstructions of the 70S ribosome generated from $N=500$ (a), $N=250$ (b), $N=100$ (c), and $N=50$ (d) projection images generated by an even angular sampling of the object (fixing the azimuthal and in-plane rotation angles for simplicity) using the proposed method (XMIPP), RELION method [12], and SPARX method [13].

Zhang et al. [13] implemented in SPARX and the RELION reconstruction method [12]) using data sets (sets of projections) of two test objects: the 70S ribosome as a small complex and a DNA-origami object as a large complex. The FSC to the known ground truth was used to compare the reconstruction results of these methods. The proposed method achieved slightly higher FSC values compared with the two other approaches for both the noise-free and noise-contaminated projections. However, the proposed method showed significantly higher values (especially for higher frequencies) compared to the other two methods for data sets with a coarser angular sampling (i.e., when the Crowther frequency was lower than the Nyquist frequency), which confirmed that the proposed approach could achieve more accurate reconstructions when there was a large gap in the Fourier space; this property should be further analyzed for its use in Tomography. Comparing the execution time of each algorithm for reconstructing the 70S ribosome shows that our algorithm is approximately twice as fast as RELION and SPARX when more than two CPUs are used. For the 3D DNA-origami object with a large size, the SPARX and RELION reconstruction methods ran out of memory in our machine with 16 GB of memory, while our multi-threaded implementation could improve the reconstruction speed by efficiently using processing resources. This lower memory requirement allows the proposed algorithm to work with larger volumes, which is in demand in the field.

This algorithm is included in Xmipp 3.1 under the name reconstruct_fourier and is downloadable from http://xmipp.cnb.csic.es.

The program is accessible through the protocols described by [21].

### References

[1] P.A. Penczek, Chapter one-fundamentals of three-dimensional reconstruction from projections, in: J.J. Grant (Ed.), Methods in Enzymology, Academic Press, 2010, pp. 1–33.

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

[3] P. Gilbert, Iterative methods for the three-dimensional reconstruction of an object from projections, J. Theor. Biol. 36 (1972) 105–117.

[4] C.O.S. Sorzano, J.A. Velázquez-Muriel, R. Marabini, G.T. Herman, J.M. Carazo, Volumetric restrictions in single particle 3DEM reconstruction, Pattern Recognit. 41 (2008) 616–626.

[5] R. Marabini, G.T. Herman, J.M. Carazo, 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs), Ultramicroscopy 72 (1998) 53–65.

[6] G. Harauz, M. van Heel, Exact filters for general geometry three dimensional reconstruction, in: Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, 1986, pp. 146–156.

[7] M. Radermacher, Weighted Back-projection methods, in: J. Frank (Ed.), Electron Tomography, Plenum, New York, 1992, pp. 91–115.

[8] C.O.S. Sorzano, R. Marabini, N. Boisset, E. Rietzel, R. Schröder, G.T. Herman, J.M. Carazo, The effect of overabundant projection directions on 3D reconstruction algorithms, J. Struct. Biol. 133 (2001) 108–118.

[9] P.A. Penczek, R. Renka, H. Schomberg, Gridding-based direct Fourier inversion of the three-dimensional ray transform, J. Opt. Soc. Am. A 21 (2004) 499–509.

[10] J.I. Jackson, C.H. Meyer, D.G. Nishimura, A. Macovski, Selection of a convolution function for Fourier inversion using gridding [computerised tomography application], IEEE Trans. Med. Imaging 10 (1991) 473–478.

[11] S. Matej, R.M. Lewitt, 3d-FRP: direct Fourier reconstruction with Fourier reprojection for fully 3-D PET, IEEE Trans. Nucl. Sci. 48 (2001) 1378–1385.

[12] S.H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination, J. Struct. Biol. 180 (2012) 519–530.

[13] W. Zhang, M. Kimmel, C.M.T. Spahn, P.A. Penczek, Heterogeneity of large macromolecular complexes revealed by 3D Cryo-EM variance analysis, Structure 16 (2008) 1770–1776.

[14] R.N. Bracewell, A. Riddle, Inversion of fan-beam scans in radio astronomy, Astrophys. J. 150 (1967) 427.

[15] H. Schomberg, J. Timmer, The gridding method for image reconstruction by Fourier transformation, IEEE Trans. Med. Imaging 14 (1995) 596–607.

[16] S. Matej, R.M. Lewitt, Efficient 3D grids for image reconstruction using spherically-symmetric volume elements, in: Nuclear Science Symposium and Medical Imaging Conference, 1994 IEEE Conference Record, vol. 1173, 1994, pp. 1177–1181.

[17] E.H. Meijering, W.J. Niessen, M.A. Viergever, Quantitative evaluation of convolution-based methods for medical image interpolation, Med. Image Anal. 5 (2001) 111–126.

[18] E. Gabriel, G. Fagg, G. Bosilca, T. Angskun, J. Dongarra, J. Squyres, V. Sahay, P. Kambadur, B. Barrett, A. Lumsdaine, R. Castain, D. Daniel, R. Graham, T. Woodall, Open {MPI}: Goals, Concept, and Design of a Next Generation {MPI} Implementation, in: Proceedings of the 11th European PVM/MPI Users' Group Meeting, 2004, pp. 97–104.

[19] F. Mueller, Pthreads Library Interface, in, 1994.

[20] X.-c. Bai, T.G. Martin, S.H. Scheres, H. Dietz, Cryo-EM structure of a 3D DNA-origami object, Proc. Natl. Acad. Sci. 109 (2012) 20012–20017.

[21] J.M. de la Rosa-Trevín, J. Otón, R. Marabini, A. Zaldívar, J. Vargas, J.M. Carazo, C.O.S. Sorzano, Xmipp 3.0: An improved software suite for image processing in electron microscopy, J. Struct. Biol. 184 (2013) 321–328.

# Appendix B

# Supplementary Data

**Supplementary Data for: A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs**

Supplementary data related to this article can be found in the following link:

http://bioinformatics.oxfordjournals.org/content/suppl/2013/07/24/btt429.DC1/supplementry.pdf

**Supplementary Data for: Alignment of direct detection device micrographs using a robust Optical Flow approach**

Supplementary data related to this article can be found in the following link:

http://www.sciencedirect.com/science/article/pii/S1047847715000313