LOCAL QUALITY ASSESSMENT OF CRYOEM RECONSTRUCTIONS AND ITS APPLICATIONS

JOSÉ LUIS VILAS PRIETO

SUPERVISORS:
PhD. Carlos Óscar Sorzano
PhD. Javier Vargas Balbuena

Facultad de Ciencias. Universidad Autónoma de Madrid
Centro Nacional de Biotecnología
BioComputing Unit

March 2019

José Luis Vilas Prieto: *Local quality assessment of cryoem reconstructions and its applications,* From local resolution to local sharpening, © March 2019

SUPERVISORS:
PhD. Carlos Óscar Sorzano
PhD. Javier Vargas Balbuena
LOCATION: Calle Darwin, 3, DATE: March 2019

Wherever you are, when it be, we'll be brothers for life, family forever.

— Jorge Maister

ABSTRACT

The cryo-Electron Microscopy (cryoem) addressed to the elucidation of macromolecular complexes has suffered a revolution in the last years. Single Particle Analysis SPA and electron tomography are the two main branches responsible of this revolution. In both cases, the result is a 3D structure of the specimen. However, it is critical to determine the degree of reliability of the reconstructed structure, problem which is solved with resolution measurements. Resolution can be determined as a global parameter of the structure, or as a local property that can spatially vary.

The main objective of the present thesis is the study of the local resolution and its applications for the validation of the reconstructed 3D structure and for local sharpening. For that, a mathematical method named *MonoRes* was developed, which measures the local resolution of reconstructed density maps and its algorithm represents the cornerstone of this thesis. Hence, *MonoDir* was designed for identifying reconstructions problems like the existence of preferred directions, or alignment errors. Note that the current state of the art does not provide an immediate and simultaneous response to both questions, in particular to the second one, which can be due to a systematic error of reconstruction. The proposed method does it by analyzing local-resolution anisotropy using the information of only the reconstructed map.

Moreover, the measurement and analysis of the local resolution for map sharpening can be also used to enhance the visualization of the protein structure keeping its structure factor. For that, method *LocalDeBlur* was developed, which modifies the local amplitudes of the different frequencies according to the local resolution values.

Apart from that, other methods were developed using slight variations of *MonoRes* algorithmic core. One of these methods is *MonoTomo*, which has been developed as the first method of local resolution in Electron Tomography. The estimation of the local resolution in electron tomograms involves work with spatially dependent noise and very large maps, so the core of *MonoRes* was extended in that direction, resulting in method *MonoTomo*, which solves those drawbacks.

These other *MonoRes* core-derived methods extend the concept of resolution in a local manner. However, resolution is not only a local parameter, but it must be considered a tensor and, therefore, it depends on the position (voxel) and the direction. Thus, the concept of local resolution is then enriched with the notion of directionality.

PUBLICATIONS

This thesis has been the origin of next publication list as first author

[1] E. Ramirez-Aportela, J.L. Vilas, R. Melero, J.M. Carazo, and C.O.S. Sorzano. "Sharpening electron density maps from its local resolution map." In: *submitted* ().

[2] J.L. Vilas et al. "Fast and automatic identification of particle tilt pairs based on Delaunay triangulation." In: *Journal of Structural Biology* 196.3 (2016), pp. 525–533.

[3] J.L. Vilas, N. Tabassum, J. Mota, D. Maluenda, A. Jiménez-Moreno, T. Majtner, J.M. Carazo, S.T. Acton, and C.O.S. Sorzano. "Advances in image processing for single-particle analysis by electron cryomicroscopy and challenges ahead." In: *Current Opinion in Structural Biology* 52 (2018), pp. 127–145.

[4] J.L. Vilas, J. Gomez-Blanco, P. Conesa, R. Melero, J.M. de la Rosa-Trevin, J. Cuenca, R. Marabini, J.M. Carazo, J. Vargas, and C.O.S. Sorzano. "MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps." In: *Structure* 26.2 (2018), pp. 337–344.

[5] J.L. Vilas et al. "Measurement of local resolution in electron tomography." In: *submitted* (2019).

[6] J.L. Vilas, H. Tagare, J. Vargas, J.M. Carazo, and C.O.S. Sorzano. "Measuring local-directional resolution and local anisotropy in cryoEM maps." In: *submitted* (2019).

[7] J.L. Vilas et al. "Re-examining the spectra of macromolecules. Limits of the Guinier approximation and of the current practice of spectral quasi B-factor flattening." In: *submitted* x.x (2019), pp. xx–xx.

For and to the family, the blood and the chosen one

ACKNOWLEDGMENTS

De alguna forma uno considera ser conciente de lo mucho que le han ayudado, pero cuando llega el momento de ejercer ese agradecimiento escribiendo estas líneas, uno se da cuenta de su inconsciencia. Del mismo modo que se necesitan todas las letras de una palabra para formar la palabra, yo os he necesitado a todos para elaborar esta memoría. Trataré de mencionaros sin dejarme a nadie.
Primero expresar mi gratitud a mis directores con especial mención a su calidad humana:
El Dr. Carlos Oscar Sorzano ha ido un verdadero mentor, una auténtica inspiración y un privilegio ser discípulo suyo. También me alegra poder decir que después de estos años es un buen amigo. Gracias por ser como eres. El Dr. Javier Vargas Balbuena, tenemos un origen común, el departamento de óptica de la UCM. Quizá esto creó un vínculo inicial que se ha mantenido hasta ahora. De tí he aprendido a sintetizar y ser pragmático, aunque no lo parezca en ciertos capítulos. Siempre me has apoyado, ayudado y has estado ahí, incluso desde Canadá, tanto nivel personal cómo profesional.
El Dr. José María Carazo, aunque oficialmente no ha podido figurar como director para mí lo es, siempre lo ha sido. Tu nombre debe estar entre los directores. Gracias por aceptarme en tu grupo, confiar en mí como doctorando y darme la oportunidad de iniciar una carrera investigadora.
En segundo lugar a toda la unidad de BioComputación del CNB/CSIC, fastásticos compañeros, profesionales y humanos. Joaquin Otón, Blanca Benítez, Jordi Burguet, José Miguel de la Rosa-Trevín, Josué Gómez-Blanco, Roberto Melero, Jesús Cuenca, Roberto Marabini, Erney Ramírez, Pablo Conesa, Juan Navas, Laura del Caño, Joan Segura, Marta Martínez, Mohsen Kazemi, Vahid Abrishami, Peter Horvath, Yaiza Rancel, David Střelák, Tomás Majtner, David Maluenda, Ruben Sánchez, Javier Mota, Yunior Fonseca, Yaiza Ranchel y Amaya Jimenez. Siempre habeis estado dispuestos a prestarme ayuda.
En tercer lugar a quien fue mi mentor durante la estancia en la universidad de Yale, el Dr. Hemant Tagare. Fue una experiencia fastástica trabajar contigo. También tengo especial recuerdo de Frederic Bonnet y Francisca Tibo con quienes compartí buenos momentos durante el tiempo de estancia.
La tesis nunca habría sido posible sin la experiencia previa adquirida en el departamento de óptica de la UCM de la mano del profesor Eusebio Bernabeu y mis compañeros: José María Herrera, José Luis III, José Antonio Sánchez Martín, Maite Irigoyen, José Augusto Rodrigo, Óscar Martínez Matos, Tatiana Alieva, Rosa Weigand, María Luisa Calvo, Gema Piquero, Alfredo Luis, Pablo de la Hoz, Francisco Torcal, Luis Miguel Sanchez Brea. Con especial agradecimiento a Juan Antonio Quiroga y Francisco Javier Salgado-Remacha por quienes conocí el grupo de Jose María Carazo y esta tesis.
La familia, la cosanguínea y la elegida, ellos lo son todo. De la cosanguínea, a mis

padres, José Luis y Marisol, les debo mucho: la vida, la ética, la moral, la eduación, su amor, cariño, y ahora también la tesis. También sois autores, todo lo que he hecho hasta ahora es vuestro. A mis tíos y tías Carmencita, Rosi, Pepe, Raquel, Tere, Pablo y primos, Pablo, Pablito, Irene, Rosa María, Miguel, José Manuel, Elena, Sara y Adrián. De la elegida a mis hermanos, Jaime Ramos y Jorge Jiménez, no llevaremos la misma sangre pero realmente somos familia, orgulloso de ser vuestro hermano. También a la nueva con sus miembros, a Velichka (baba vili), Penka, Lazar, Lina, Miguel e Iván. Ellos me han aceptado y querido como uno más desde el primer día, somos familia. Finalmente a ti Alex (Alexandra Lazarova) haces que cada instante sea especial, ya son muchos instantes. Para bien o para mal, no me has dejado indiferente al cruzarte en mi vida y quiero una vida a tu lado. Gracias por aguantarme dentro y fuera del CNB.
También estoy muy agradecido a Bonifacio de Andrés y Toro buen profesor, buen amigo, y mejor ajedrecista. Y a Antonio Hernando y Asunción García por su interés y apoyo durante estos años.

CONTENTS

LIST OF TABLES

ACRONYMS

3DFSC 3D Fourier Shell Correlation

ADC Analog to Digital Converter

ADR Average Directional Resolution

AFM Atomic Force Microscope

BIM Beam Induce Movement

cryoEM Cryo-Electron Microscopy

cryoET Cryo-Electron Tomography

CCD Charge Coupled Device

CERN Conseil Europeen pour la Recherche Nucleaire

CMOS Complementary Metal Oxide Semiconductor

CNN Convolutional Neural Network

CTF Contrast Transfer Function

DDD Direct Detector Device

DED Direct Electron Detector

DQE Detective Quantum Efficiency

EM Electron Microscopy

EMAN Electron Micrograph ANalysis (Software)

EMDB Electron Microscopy Data Bank

EMPIAR Electron Microscopy Public Image Archive

FSC Fourier Shell Correlation

FSCe/o Fourier Shell Correlation eve odd

GPU Graphics Processing Unit

HPD Hybrid Pixel Detector

MAPS Monolythic Active Pixels Sensors

MRC Multi-reference classification

MSA Multivariate Statistical Analysis

MTF Modulation Transfer Function

NEQ Noise Equivalent Quanta

NLOO Noise-compensated Leave One Out

NMR Nuclear Magnetic Resonance

OTF Optical Transfer Function

PCA Principal Components Analysis

PDB Protein Data Bank

PDE Partial Differential Equation

PSD Power Spectral Density

PSF Point Spread Function

RCT Random Conical Tilt

RELION REgularised LIkelihood OptimisatioN (Software)

RMS Root Mean Square

SNR Signal to Noise Ratio

SPA Single Particle Analysis

SSNR Spectral Signal to Noise Ratio

STM Scanning Tunneling Microscope

TEM Transmission Electron Microscope

Xmipp X-windows based microscopy image processing package (Software)

Part I

ON THIS THESIS

1

INTRODUCTION ON THIS THESIS

Structural biology has as objective the elucidation and analysis of the 3D structure of macromolecular complexes to then understand their action mechanism and function inside cells. Many conclusions and applications about the complexes can be obtained from the structure, for instance, their biological behaviour, action mechanisms, drugs development, or design of new structural complexes, among other. For that purpose, structural biology makes use of different imaging techniques, for example: the nuclear magnetic resonance (Nuclear Magnetic Resonance (NMR)), x-ray crystallography, or electron microscopy (Electron Microscopy (EM)) among others. The first one measures the changes in the orientation of the spin angular momentum experienced by atomic nucleus as response to intense magnetic fields, recovering the structure from it. The x-ray crystallography uses the x-ray scattering, due to the light-sample interaction to get diffraction patterns and to infer the structure from them. Finally, electron microscopy considers the use of electrons as lighting source to get the image projections of the specimen. This last technique, applied to the reconstruction of biological complexes via image processing techniques, will be the topic of this thesis.

The history of electron microscopy begins with the development of the first microscope by Ernst Ruska in 1925. For that milestone he was awarded with the Nobel prize in 1986[137]. The imaging capabilities of this new kind of microscope were quickly adopted in the study of biological specimens [103]. However, leaving out some works, structural biology suffered a first revolution with the work of Aaron Klug [7] who combined the x-ray diffraction and electron microscopy to introduce the electron crystallography as a technique to elucidate the 3D structure of macromolecular complexes. For this technique, he was awarded with the Nobel prize in 1982 [88]. Although, x-ray and electron crystallography requires the growing of a crystal that contains the protein in a periodic arrangement, which represents an important drawback. Note that some complexes can be difficult or even impossible to be crystallized, and many others lack of high symmetries and periodicity.
The capability of distinguishing details in an image is given by the image contrast, in particular, in a transmission electron microscopy (Transmission Electron Microscope (TEM)), imaging is obtained by phase contrast which is increased with the difference of the atomic number of the atoms that compose the sample. Unfortunately, most of the elements that constitute the biological complexes present low atomic number (H, C, N, O,...), moreover, they are usually diluted in their native state in aqueous solution. As a consequence, the contrast of biological elements will be poor. To solve that issue, the exposure time can be increased, but the prize to pay is the risk of burning the sample [49]. As possible solution, R.M. Glaeser proposed the average of multiple images [49], this concept is one of the roots of

the current workflow in Single Particle Analysis, i.e. noise suppression by multiple averaging. Thus, a few years later and following this idea and making use of the electron crystallography, N. Unwin and R. Henderson (Nobel awarded in 2017) were able to elucidate the structure of the bacteriorhodopsin at a resolution of 7 Å[182]. Two decades later that structure was refined achieving an atomic resolution of 3 Å[73].

It was also observed that if the crystal is frozen, then the radiation damage is alleviated [175]. Hence, the use of cryo-samples was introduced in electron microscopy [176] and also adopted in x-ray crystallography. However, it took a few years more to get one of the most considerable step forward: the vitreous ice samples. The idea was the plunge-freezing method. Since the water is frozen in a thin film so quickly that it has no time to form crystals [5, 31]. For this development Jacques Dubochet was awarded with the Nobel prize in 2017, and Cryo-Electron Microscopy (cryoEM) was named as the field that came in from the cold [34]. It is specially interesting, even in nowadays, the article of J. Dubochet [32] about the state of the art of electron microscopy in the 80's decade, with special incidence on vitreous ice sample preparation. Hence, the acquisition of isolated particles became possible. Unfortunately, the contrast in cryo-samples represents an important limitation. As an alternative, even though they were developed a few decades before, heavy-metal-stained samples were simultaneously used. They provided a good contrast and allowed to analyze non-symmetrical samples (symmetry can be exploited for multiple averaging), but with the disadvantage of dehydrating and, as a consequence, shrinking the sample.

The current workflow of Single Particle Analysis (SPA) began to be established with the first softwares, SPIDER [43] and IMAGIC [65, 68] thanks to Joachim Frank (Nobel prize in 2017) and Marin van Heel respectively in 80's decade. They were focused in increasing the signal to noise ratio by 2D averaging and classification [44, 67] as well as the measurement of resolution. In this last topic, three methods should be highlighted, Q-factor [64], differential phase residual [45], and the current standard, the Fourier Shell correlation (Fourier Shell Correlation (FSC)) [183]. Moreover, in those years it was raised the random conical tilt method (Random Conical Tilt (RCT)), that helps in the reconstruction process by providing prior information about the geometry [128]. Also, it was introduced the idea of reassigning/correcting particle orientation once the first 3D reconstruction is determined [66] can be considered as the first iterative refinement method in 3D-reconstruction. However, for the taking off of SPA and to become what it is now, the field had to wait for the proper combination of three pillars [42]: good sample preparation (which has already been explained), better microscopes, and enough computational power along with mathematical methods. It took a few decades (around 2012) to achieve the convergence of these three subjects.

From the microscope point of view (first pillar), the instrument was improved since the Ruska design, and many manufacturing companies (FEI, JEOL, Gatan, Phylips, Hitachi or Carl Zeiss among many others) worked on this device providing more improvements. Despite that the microscope is the sum of its parts, the main responsible of the imaging improvements were [191], the column stability, Field Emission Electron Guns, new detectors, and microscope automation. The col-

umn stability was enhanced thanks to better magnetic lenses, and vacuum and isolation systems. The introduction of Field Emission Guns along with the use of higher acceleration voltages (100-300 keV), represented a critical step. The new kind of guns made possible to improve the brightness, the coherence, chromaticity of the electron beam, and to avoid the evaporation of cathode material[63]. Detectors played also a special role, at the beginning film plates with silver halides were used as sensor. However, the evolution of technology brought a new sort of sensors: the Charge Coupled Device (CCD) (Charge Coupled devices, William Boyle and George Smith were awarded with the Nobel prize in 2009). These sensors speeded up the data acquisition in 90's decade, and cast digitized data for a later computational process. The disadvantage is that the high energy of electron damages the sensor, shortening its life time. As countermeasure, an intermedia phosphorous layer or a scintillator was placed to protect the detector [36, 170]. The final boosting that determined the push forward of microscope hardware since 2010 was the use of Complementary Metal Oxide Semiconductor (CMOS) technology with the direct detectors. The speed of CMOS joined to a separation between the detection of electrons and the reading of the produced signal shown a considerable improvement in the Detective quantum efficiency (DQE) [108], and a new kind of detectors were developed, direct detectors (Direct Detector Device (DDD)) [107]. In particular, the high acquisition speed changed the acquisition process allowing to record movies instead of just getting a single image. It confirmed that, as a consequence of the interaction radiation-matter, the particles embedded in the ice present a motion frame by frame named beam induce a movement [51, 97]. The movie acquisition increased the demand of high performance computation, note that the raw data of a current microscope session can perfectly reach 1 Tb/day.
The development of computational methods and a robust theoretical background behind imaging were the last responsible of the named resolution revolution [92] and they constitute the third pillar defined by J. Frank. Meanwhile the theory of imaging was devised time ago with the optical microscopy, it needed to be slightly adapted to the electron lighting due to the interaction electron-sample. Thus, assuming a set of considerations like thin sample, inelastic scattering, and weak interaction electron-matter, the so-called weak phase object approximation is the proper model to describe the image formation [63]. However, these conditions make the sample to be almost transparent under the electron illumination. The microscope is characterized by a transfer function, the named Contrast Transfer Function (CTF), that convoluted with the object results in the measured image. Remark that aberrations and defocus are considered in the CTF. The CTF was deeply analyzed up to be standardized to correct aberrations and achieve high resolution reconstructions. However, a question emerges about the lighter macromolecule that can be elucidated with electron microscopy when image acquisition is perfect and the unique limiting factor is radiation damage [71]. Thus, using first principles and only 12000 particle resolution of 33 Å was achieved. It is in this 90's decade when the reconstruction method of projection matching [123] and maximum likelihood [155] were established. They have currently become the core of many SPA methods. A few years later, in an amazing article from P. Rosenthal and R. Henderson [135] three concepts were introduced: the idea of map sharpening by B-factor correction,

the threshold of 0.143 or 0.5 for computing the FSC resolution, and the determination of particle orientation (that was the preambule for a later publication about tilt-pair validation [74]). Although it was in 80s decade when the first EM-softwares appeared, the real revolution in methods began around the year 2000 up to nowadays. Perhaps due to informatics experimented a new age with the boosting of the computational power as predicted by the Moore's law, the software design became a common practice in companies, as so as the implementation of distributing systems and parallel computation, and more recently the usage of Graphics Processing Unit (GPU). Thus, new EM-image processing packages were released by different laboratories such as Electron Micrograph ANalysis (Software) (EMAN) [98, 174], Frealign [58], X-windows based microscopy image processing package (Software) (Xmipp) [134, 162], Bsoft, [76, 77] and more recently REgularised LIkelihood OptimisatioN (Software) (RELION) [84, 205, 206], or Cryosparc [126]. They only work with the acquired images and via image processing attempt to obtain the 3D structure of the macromolecule measured. To do that those packages include many algorithms able to carry out the full pipeline of SPA, for instance: boxer-EMAN for particle picking, CL2D-Xmipp for 2D Classification, AutoRefine-Relion to refine reconstructions. The large number of packages caused a need of integrating and unifying all of them in a single framework. Hence, integration softwares as *Appion* [46] and *Scipion* [207] were developed. However, there also exist independent developers that have released some algorithms like MotionCor for correcting the blurring appeared as a consequence of the beam induce a motion [97, 202], *ResMap*, for calculating the local resolution in 3D EM-maps [90], cryolo for particle picking using a deep learning approach [194], gautomatch also for particle picking but using a reference [80], ctffind for CTF estimation [110, 133] or *LocScale* for map sharpening [81] among many others.

cryoEM has progressed a lot since its origin, a proof of that can be seen in Fig. 1 where the number of structures deposited in databases is shown. A complement analysis is the temporal evolution of the resolution of the deposited maps, currently, to achieve a resolution close to the atomic ones is almost common, see Fig. 1. Despite this progress it remains a long way to walk, the three pillars described by J. Frank. In the pillar of the microscope hardware, perhaps the most important issue is to reduce the high cost of electron microscopes, which represent a non-affordable barrier for many laboratories and institutions. Note that an electron microscope can cost 5M€ leaving out the maintenance cost close to 200k-300 k€ per year. The development of better detectors with a higher DQE have represented and still represents a a critical step in hardware development. Thanks to new and better detectors, the signal to noise ratio might be decreased, and the associated uncertain, that complicates the alignment and therefore the elucidation of the macromolecule may be minimized. It also helps in the beam induce motion corrections to reduce the blurring effect. The improvement of phase plates and its phase stability, may also help. Regarding with the microscope and image acquisition, the automation of microscope session; although the state of the art has been considerably enhanced with software like serialEM [104, 105], leginon [23] or EPU from the company Thermo Fisher Scientific; the acquisition requires a previous analysis of the sample and the grid region to be imaged is an issue that should be

improved.
In the pillar of sample preparation, cryoEM started to be increased with better biochemistry purification method. Many of them came in from crystallography. Nowadays, the efforts on microscope theory and acquisition are focused on the development of grid samples preparation to avoid the charging the sample [140], the existence of preferred directions [111, 173], or the minimization of the interface effect in the interaction macromolecular-interface [141, 142], among others.
In the third column, image processing, the field of cryoEM points to speed up the computational time to elucidate the macromolecule structure in less than a day. For that, two main ideas are on the table, the use of GPU computing and streaming approaches. The second one considers the processing of next the steps of the SPA workflow before the end of the previous ones, for instance, the 2D-classifications starts while the particle picking is on going. This approach is the first attempt towards the high throughput in the microscope session, thus some automatized workflows based on streaming were already proposed [53]. Other hotspot to be solved from the image processing block are the measurement of high order aberration as can be the trefoil or the tilt correction that are currently neglected. The beam induced movement correction was critical to increase the obtained resolution. Thus, the use of frames information by particle polishing could be a promising tool. Another hot topic is the use of map sharpening for refining the obtained 3D-structure, concerning that, this thesis proposes a solution. Finally, the need of new approaches to solve the heterogeneity problem is also an important requirement in the field, in particular, when the heterogeneity is conformational and there exists a continuous number of states.

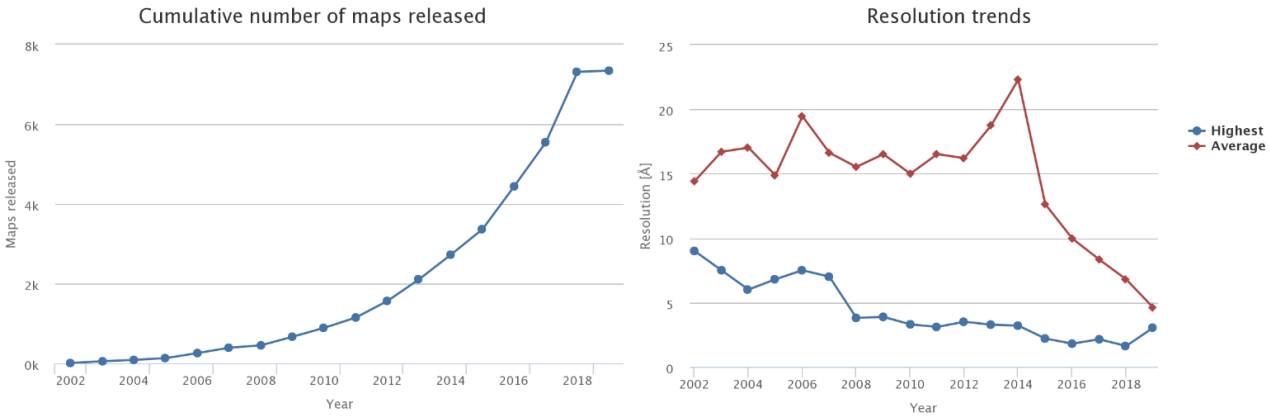

Figure 1: (left) Number of deposited density maps in the Electron Microscopy Data Base (EMDB) obtained by electron microscopy. (right) Temporal evolution of the highest and average resolution of the deposited electron microscopy maps.

1.1 MOTIVATION OF THIS THESIS

This thesis is devoted to the third pillar of the resolution revolution, i.e. mathematical/computational methods, in particular, on the quality analysis in a local manner of the reconstructions, i.e. the local resolution analysis and its applications

on macromolecular complexes reconstructed with SPA techniques.
Resolution has traditionally understood as the size of smallest detail, that can be elucidated in the reconstructed map, it is given by a number. In recent times, this concept was extended in a local manner by considering that the quality of a map is spatially variant. Thus, *blocres* [22] was born as the first local resolution method, calculating a local FSC using a mobile window with two half maps. Curiously, local resolution measures were standardized as part of the SPA workflow with the publication of other method, *ResMap*. [90]. Despite both methods work properly, they present drawbacks that we wanted to solve, as they are: first, to achieve a fully automatic method that avoid the user intervention and second, to design an algorithm computational enough fast to deal with large maps in short computational times. To do that, the use of monogenic signal was considered. This kind of signals, recently introduced [39], extend the concept of analytic signal to several dimensions allowing a local decomposition in phase and amplitude. The local energy of the reconstructed map can be then obtained by determining the monogenic amplitude. Thus, using this local approach a simple and computationally easy algorithm of local resolution was design. The method, that was named *MonoRes*, performs a frequency sweep from low to high, and determines if the local monogenic amplitude is statistically higher than the amplitude of noise (at that frequency). It allows to assign resolution values when the energy of signal is lesser than the energy of noise.
In electron tomography, the measurement of resolution as a global parameter to define the quality is not widely used. Despite that, there exist method to defined it, for instance: the FSC even-odd and the Noise-compensated Leave One Out (Noise-compensated Leave One Out (NLOO)) [21]. The first one is conceptually simple and keeps the standard global resolution measurement of SPA to electron tomography. It is done by splitting the set of images in two sets (odd and even) to then reconstruct two independent tomograms that allows to compute the FSC as two half maps (in SPA). Unfortunately, it remains its cons as it is the mask dependency. The second one is more accurate but computationally hard. However, there no exist any method for estimating the local resolution in electron tomography. Thus, *MonoTomo* was developed to cover this need of the field. This was done following the *MonoRes* algorithm, but adapting it to deal with the problems that tomography presents, as they are the spatially variant noise, and the large size of the tomograms.
Recently, it was To finish with the local resolution measurements, it rests to analyze if resolution presents a dependence on the direction. Recently, a few works were addressed to cast light on this topic [111, 172] but in a global manner. In other words, they analyze how the global resolution varies with the direction with the aim of study the existence of preferred directions. Thus, by means of *MonoRes* core, and adding directional filter, it is possible to determine a local-directional resolution measurement. When all possible directions are analyzed, then an analysis on local resolution anisotropy can be carried out. To do this an algorithm, named *MonoDir* was developed. Up to our knowledge, it is the first local-directional resolution method in cryoEM. Surprisingly, this analysis provides very rich information as validation quality map tool. It informs about: the existence of preferred direc-

tions, angular assignment errors, good coverage of the projection sphere, as well as of course, the local resolution anisotropy.
Since local resolution was introduced in the field, it was applied to understand the behaviour of macromolecules, for instance, flexibility, heterogeneity are usually identified as low resolution areas in the density map. However, and leaving out local resolution filters, local resolution has no find more applications. Understanding local resolution as a measurement of Signal to Noise Ratio (SNR), a very useful application is found in sharpening. Currently, the sharpening techniques considers a global transformation that boosts the high frequencies that map presents weighting them by a merit function. The most representative technique is the b-factor correction, that recover the high frequencies by making flat the structure factor (spectrum) of the macromolecule, being the merit function the FSC. In other words, the global SNR is considered. Thus, we explore the theory behind this transformation under the physics scope, in particular the scattering theory. The result is that despite the good results in the visualization of the sharpened map, this transformation is not supported by physics. Taking this into account, a sharpening technique that restores the high frequency components weighing by local values of SNR (local resolution) was designed. The algorithm, called *LocalDeBlur*, keeps the structure factors compatible with scattering theory and considerably helps in the understanding of the macromolecule for a later atomic modelling.

1.2 STRUCTURE OF CONTENTS

This thesis is structured in two main parts:
Part I, named *Fundamentals of electron microscopy and Single particle analysis* should be understood as an introductory block to the field of SPA. It contains the background in Physics to understand the image formation in the electron microscope and the image processing methods of SPA, which allow to elucidate the 3D structure of the study complex. Finally, the last chapter is addressed to introduce the subject of this thesis, i.e. resolution measurements, showing the traditional concept of resolution. However, despite this chapter is an introduction, it also contains many comments and original reasoning introduced in this thesis. In particular, to be more specific the contents are briefly described:

- *In chapter 2* the fundamentals of the electron microscope are introduced. The chapter provides a quick and consistent explanation but enough detailed about the physics of the microscope and its use in imaging. Hence, it starts by identifying the imaging needs and how the use of electrons pushes the size limit of the imaged elements. Then, the electron microscope is described piece by pieces starting with the physics of the electron guns based on the Fermi-Dirac statistic to derive the Richardson law and the emission field. Next, the mechanism to focus electron are described and how the aberration affects the image. Then, the electron-specimen interaction is explained, and the weak phase object approximation is introduced. Finally, the physics of detectors is also introduced, with special incidence on the direct electron detectors, which have been responsible of the resolution revolution [92] at the field of cryoEM.

- *In chapter 3* the key concepts of SPA are carefully explained, i.e, the hypothesis that allows the reconstruction as they are the projection assumption and that particles are identical copies of the same macromolecule. Then, the central slice theorem, which is the main responsible of most of the reconstruction methods. Once the basis of SPA is established, the general workflow of SPA and the image processing techniques involved are described. Thus, all steps of the workflow are briefly explained: the movies, the frame alignment, the gain correction, CTF estimation, particle picking, 2D classes, initial volume, 3D classes, refinement methods, and validation and analysis methods.

- *In chapter 4* a brief review of the global resolution concept in cryoEM is introduced. This chapter attempt to complete the introduction to SPA by providing specific information about the topic of this thesis, resolution, and how it is currently treated in the SPA workflow. The aim is to provide the background required to understand the part II, in which the developed methods about resolution are described.

Part II, named *Developments and results*, contains the novelty that this thesis introduces in the field of cryoEM, in particular in the local resolution measurement and its applications. Hence, following the previous structure, the contents are briefly summarized.

- *In chapter 5* the concept of local resolution and its origin is introduced. Then, a fully automatic method for computing the local resolution of macromolecular complexes, named *MonoRes* is proposed. It might be considered the central chapter of this thesis, due to the core of many other methods, that will be explained later, is *MonoRes* algorithm. Despite this thesis is focused on SPA, *MonoRes* was also adapted to be able to work in other cryoEM techniques as it is the case of electron tomography. Hence, a method of local resolution in electron tomography, called *MonoTomo*, was designed. Both methods were tested with experimental datasets to show their capabilities.

- *In chapter 6* the study of local resolution continues but now the concept of directionality is introduced to measure local-directional resolution. The idea is that resolution is not a only single number for each position (local resolution), it is determined by a tensor. The component analysis of local-directional resolution will determine the anisotropy of resolution. This analysis provides really useful information about the quality of the reconstructed macromolecule because it allows to detect if the reconstructed structure owns angular assignment error, the possible existence of preferred directions, or the existence of anisotropy. For that, *MonoDir* algorithm was developed, and the theoretical results are supported with experimental tests.

- *In chapter 7* the B-factor correction is analyzed. To do that, the structure factor is derived from the scattering theory. In particular, the Guinier and Porod laws. Theoretical results point out that B-factor corrections are not supported in the scattering theory. To check that, a set of experiments with atomic models from different macromolecules are carried out. Results are in agreement with the theory, concluding that the b-factor correction should be revisited.

- *In chapter 8* a sharpening method is proposed. The algorithm named *LocalDeBlur*, provides a fully automatic method for local sharpening based on local resolution measurements. Thus, an alternative sharpening method to the traditional B-factor correction (that was studied in the previous chapter) is suggested. Finally, *LocalDeBlur* is applied to several experimental structures, and compared with other sharpening methods which belong to the current state of the art.
- *In chapter 9* the conclusions of this thesis are exposed. They summarized the conceptual essence of each method and development.

1.3 OBJECTIVES

To summarize the objectives of this thesis are:

1. Develop a fully automatic method for measuring local resolution in SPA. This method, *MonoRes*, is described in Chapter 5.
2. Extend the local resolution measurements of SPA to of a method able to measure local resolution in SPA. The method, *MonoTomo* is described in Chapter 5.
3. Develop a local-directional resolution method, *MonoDir*, and how its anisotropy analysis allows to validate the reconstructed structure by identifying angular errors and preferred directions. This objective is explain in chapter 6.
4. Analyze the B-factor correction and its use for sharpening from the physics point of view making use of the scattering theory. This analysis is carried out in Chapter 7.
5. Develop a local sharpening method, *LocalDeBlur*, based on local resolution measurements avoiding B-factor correction. This method is explain in chapter 8.
6. Provide a good and concise but complete introduction to the physics of the electron microscope and SPA workflow. These tasks are shown in chapter 2, 3 and 4.

Part II

FUNDAMENTALS OF ELECTRON MICROSCOPY AND SINGLE PARTICLE ANALYSIS

2 THE ELECTRON MICROSCOPE

The beginning of electron microscopy is fuzzy with its origin between 1920-1935 when many advances were carried out in electron optics. The most spread date is the year 1931 when the first electron microscope was built by Erst Ruska [138] awarded for it with Nobel prize in 1986. However, its birth was only possible thanks to the development of magnetic lenses a few years early by Hans Busch in 1926. Since then, and despite a beginning without funding in which E. Ruska was working for free [137], many technological improvements were addressed to achieve higher resolutions solving details beyond than the optical microscope allows [30]. Thus, electron microscopy raised as a new field of study in physics with special applications in the study of materials or biological sciences among many other.
The objective of this chapter is to provide the overall concepts and elements which describe the transmission electron microscope (TEM) and its capabilities for imaging. The starting point will be a definition of the microscope describing its limitations. The use of electrons as illuminating source represents the main differences between optical an electron microscopes, hence the beginning is to understand the advantages that it involves. However, to substitute light by electrons implies many experimental drawbacks. Thus, electron microscope will be analyzed element by element.

*The sections marked with an asterisk, *, in title are out of the scope of the thesis introduction. They were included to complete physics of the microscope, therefore, they can be omitted*

2.1 ON THE MICROSCOPE AND ITS DESCRIPTION *

A radiation microscope (electromagnetic or particles) can be coarsely defined as every device capable of getting a magnified image of an object/sample. The term radiation ought to be highlighted. There are other kinds of microscopes as the Atomic Force Microscope (Atomic Force Microscope (AFM)) or the Scanning Tunneling Microscope (Scanning Tunneling Microscope (STM)) that work without a source of radiation; the radiation term makes reference to every kind of energy which fulfills the duality wave-particle. In particular, this thesis is focused on the transmission electron microscope.
Modern microscopes are constituted by many optical elements. First an illumination system is required, i.e. source of radiation and a collimator or condenser lens are responsible of lighting the sample. Next, the radiation will interact with the specimen to finally reach an imaging system composed by an objective and an ocular that will capture the scattered radiation to record an image. The objective essentially is the microscope and the main lens. It is the closest element to the ob-

ject/sample and its function is to collect the light or radiation and get a magnified projected imaged for in a second step to be more magnified by the ocular system, see Fig. 2.

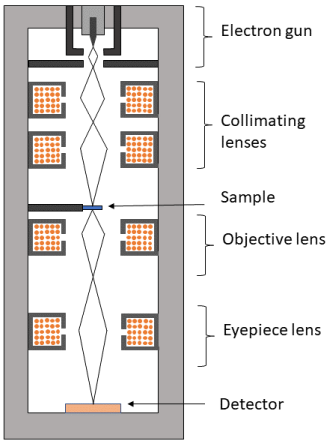

Figure 2: Scheme of the electron microscope showing all its components: the electron gun, collimating lenses, sample, objective lens, eyepiece and detector.

A suitable theoretical framework to describe the microscope relies on the degree of details that the theory is able to reach accordingly with the experimental needs. The choice of one or another description depends then on two factors: wavelength of the radiation and the interaction radiation-measure equipment. The first approach is the geometrical optics, it is fulfilled when the wavelength and energy of the radiation can be neglected in comparison with the dimensions and energy of the microscope elements. In this sense, geometrical optics represents a corpuscular treatment of the radiation, so that, the image can be obtained as impacts of the particle on the detector. The second approach is the electromagnetic theory or physical optics, in which the wavelength of the radiation are comparable to the elements involved in the systems, but the energy of the radiation is not able to interact due to the small energy of the radiation. These two models represent the framework of most kinds of microscopes, in particular, of the electron microscopy. The last description would be the quantum theory, taking into account the interaction radiation-matter when the radiation energy is comparable with the energy of the system. Quantum theory does not apply to the purpose of this study.

Geometrical optics describes the microscope from a macroscopic point of view. Radiation is treated as an idealization of reality called rays. A ray is a unidimensional line defined by the trajectory of the radiation, the direction of a ray can only change if a property of the propagation medium changes. When electromagnetic radiation is used, that optical property is called refractive index, in the case of an electron microscope the propagation medium is the vacuum and the change of direction is achieved by modifying the electromagnetic fields using magnetic

lenses. Thus, geometrical optics describes the microscope in terms of distances object-image, the explanation about why radiation changes its direction or how interacts with matter requires the electromagnetic theory or, to be more precise, scattering theory of waves and particles.

The performance of the microscopes as imaging systems should be as good as possible. Geometrical optics defines an imaging system as perfect, also called stigmatic, if it fulfills the called *Maxwell conditions* or *Maxwell optical theorems*: [106]

1. There exists a relation one-to-one between object and image, i.e, the image of a single point is a single point.
2. There is a similarity relationship between object and image, being the magnification the similarity ratio.
3. If the object is on a plane, its image lies on another plane.

If an optical system does not fulfill any of these conditions, then the optical system will present aberrations, being an aberration any deviation of perfect imaging.

2.2 USE OF ELECTRONS

Before examining the need of using electrons let us to analyze the limits of any optical system, to finally conclude that the use of electrons will extend the imaging capability for measuring smaller details.

2.2.1 *Limit of the optical systems*

Leaving out external agents, there are two issues responsible of breaking the Maxwell conditions (section 2.1) in all optical systems: the aberrations which depend on the system, and the diffraction which is an intrinsic property of radiation (electrons / light / ions, ...) that interacts with matter. Aberrations are defined as any deviation of a perfect image, they can be corrected by introducing new lenses or if they are enough small, by image processing. In contrast, diffraction as an intrinsic behaviour of the interaction radiation-matter cannot be avoided. This means that even an aberration-free system will present deviation from the perfect image by diffraction effects, i.e. all optical systems are limited by diffraction. It is noteworthy that electrons and light photons interact in a different manner with matter, and therefore, they will differ in their diffractive behaviour. Diffraction breaks the Maxwell conditions doing the image of a point object will be a spot instead of another point, the spot is called Point Spread Function (PSF), and it characterizes the optical system. The smaller the PSF is, the higher quality has the optical system, in particular, the microscope. As will be shown, the use of electrons instead of light reduces the dimensions of the PSF, and as a consequence, higher details can observed. To do that, the PSF shape under diffractive effects ought to be analyzed.

Consider a circular aperture with dimensions around the wavelength of the radiation. The choice of this kind of aperture is because in most of optical systems like tubes, lenses, diaphragms, and other optical elements usually present circular shape. To define the smallest detail of an object that can be recorded, a circular diagram can placed in contact with the sample, such as reducing the diaphragm area, smaller area of the sample can be recorded, see Fig. 3. This diaphragm limits the amount of radiation that passes through the system, and therefore will be the aperture diaphragm.

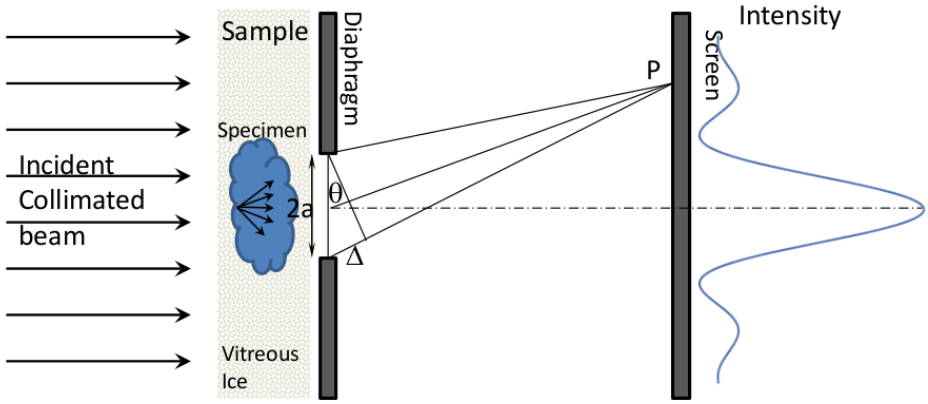

Figure 3: An incident collimated beam reach the sample and is scattered. A circular diaphragm of diameter $2a$ at the exit reduces sample imaged area in a screen. When the diameter of the diaphragm is around the wavelength radiation, and interference pattern called Airy disc is imaged. P represents a far point on the screen and Δ the optical path difference between the marginal scattered rays and the central one.

Under these conditions, a perfectly collimated and coherent beam illuminates the sample and the radiation will be diffracted by the diaphragm casting an interferogram called Airy disc in a far screen. The mathematical shape of the Airy disc can be obtained by means of the Fraunhofer diffraction. Nevertheless, we are interested in the overall shape of this pattern instead of the exact solution, taking a shortcut and avoiding mathematical calculus. First, the intensity pattern will depend on the difference of optical path between rays. Given two marginal rays, the optical path difference, Δ with respect to the central ray, see Fig. 3 will be given by

$$\Delta = k\delta s = \frac{2\pi}{\lambda} a \sin\theta, \tag{1}$$

where $k = 2\pi/\lambda$ is the norm of the wave vector, λ, the wavelength of the incident radiation, $2a$ the aperture diameter. Second, due to the circular aperture, the problem of finding the intensity pattern in a far screen will present cylindrical symmetry. It is well known that Bessel functions commonly appear in problems with such symmetry. Hence, it is plausible to consider that the intensity pattern, I, will be in

somehow proportional to a Bessel function, $J_1(\Delta)$, with dependence on the optical path difference, Δ,

$$I \propto J_1(\Delta). \tag{2}$$

This function presents zero values of intensity for those values, Δ, such as $J_1(\Delta) = 0$, taking the first zero place at $\Delta = 3.8317$, then

$$\frac{2\pi}{\lambda} a \sin\theta = 3.8317 \qquad \Rightarrow \qquad \sin\theta = 1.22\frac{\lambda}{2a}. \tag{3}$$

Note that θ is the angular distance between the center of the screen, called maximum of order 0, and each zero of the Airy disc, in particular the distance to the first zero is named size of the Airy disc. Despite the zeros of Bessel function and the Airy disc are in same position, they are different functions, being the Airy disc given by $J_1(x)/x$, where x is the argument of the function.
Eq. (3) is the starting point for defining the Rayleigh Criterion [17] and the resolution concept of an optical system as will be shown in Chapter 4. However, leaving the issue of resolution out, what Eq. (3) implies is that as the angular separation between the first zero and the center of the Airy disc directly depends on the wavelength and inversely on the size of the object (aperture). Due to the finite size of the aperture, the wavelength determines the separation of the first zero. For this reason, the wavelengths in the visible spectrum (400 - 700 nm) limit the capability of the optical system for solving very small details. To solve this problem requires the use of lower wavelengths, thus, the use of electrons with enough energy (short wavelength) allows to record smaller details. Finally, it is necessary to determine the wavelength of electrons.

2.2.2 *Wavelength of the electrons* *

Louis De Broglie pointed out that any particle with momentum, p, can interact as a wave with associated wavelength

$$\lambda = \frac{h}{p}, \tag{4}$$

where, h is the Planck constant. This asset is known as *De Broglie Hypothesis* [10]. The momentum, p, of the electrons determines the wavelength and it is related with the acceleration voltage, ϕ, of the electron gun. Electrons are then accelerated by a potential difference, acquiring kinetic energy,

$$T = e\phi, \tag{5}$$

where e is the electron charge. The high energy provided, around $200\ keV - 300\ keV$ implies electrons speeds, v as high as the speed of light, c (only if a classical approach is considered). Relativistic corrections are then required to calculated the speed of electrons. To do that the energy conservation law can be used, $(m_e c^2 + T) = mc^2$, where T is the kinetic energy given by

$$T = (\gamma - 1) m_e c^2, \tag{6}$$

with $\gamma = 1/\sqrt{(1 - v^2/c^2)}$, m_e the electron rest mass and m the electron mass. Therefore, by making use of Eq. (5) and Eq. (6),

$$\frac{v}{c} = \sqrt{1 - \frac{m_e^2 c^4}{(m_e c^2 + e\phi)^2}}, \tag{7}$$

gives and idea about speed of electrons in terms of c, in particular, for a potential of $\phi = 300$ kV, the velocity is $v = 0.77c$. Finally taking into account Eq. (4) and explicit expression for the wavelength can be obtained, for convenience and it is better to use its inverse doing $1/\lambda = p/h = pc/hc$, to get the ratio v/c due to the momentum $p = m_e \gamma v$, resulting

$$\frac{1}{\lambda} = \frac{m_e \gamma c}{h} \sqrt{1 - \frac{m_e^2 c^4}{(m_e c^2 + e\phi)^2}}, \tag{8}$$

or removing the γ dependence as

$$\frac{1}{\lambda} = \frac{1}{hc} \sqrt{e\phi(2m_e c^2 + e\phi)}. \tag{9}$$

Again, for an usual potential of 300 kV, electrons will present an associated wavelength of 0.0197pm

2.2.3 *Experimental problems derived of the use of electrons*

The use of electrons pushes the diffraction limit beyond the limit of the optical microscopy. Unfortunately, to use electrons is a challenge from an experimental point of view:

1. There are no natural electron sources, that requires the design of an electron gun that provides enough energy to accelerate electrons toward the sample without interacting.

2. Electrons present a strong interaction with matter, and therefore, the need to work in vacuum. In the case of TEM, the electrons should pass through the sample without too much deviation, meaning a weak interaction electron-matter, which is achieved with high-energy electrons.

3. Radiation can damage the sample. It might ionize the sample, heat it and degrade it, hence, radiation must be as low as possible, compromising the contrast.

4. The challenge of focusing electrons: TEM works in vacuum and imaging electrons requires of devices that act as lenses, electron lenses. Moreover, these lenses must deal with electrons of high energy which makes focusing more difficult.

5. Radiation damages the sensors. The high energy of electrons reduces the life time of sensors or compromises the contrast and quality [49, 71]. A new generation of sensors called direct detectors has opened new possibilities in imaging [107].

2.3 ELECTRON SOURCES: ELECTRON GUNS *

2.3.1 *General considerations* *

All imaging systems require an illuminating source to acquire images, there is no image without a radiation source. The most general definition of electron source is every physical configuration able to release electrons from it. Unfortunately, the imaging needs in electron microscopy impose some constrains, which can be summarized as high energy and stability. First condition is required to allow electrons transmission through the sample under a weak interaction or even without interaction. However, stability is most critical because it involves three experimental challenges: monochromaticity, coherence and collimation.
Monochromaticity means that the energy beam is always the same without suffering fluctuations in energy, or alternatively that the electron wavelength is constant -see Eq. (9); this is allowed by very stable currents.
Coherence makes reference to a known phase relation. The coherence is a necessary condition in TEM because the image is obtained by interference of phase contrast. There are two types of coherence, spatial and temporal, when a beam presents both it is called coherent (for a deep study see [54, 102]). The spatial coherence is defined as a well defined phase difference between two different emitters (usually in the transverse section of the beam). Meanwhile, a beam presents temporal coherence if there is a well defined phase difference between two instants of time. Note that, temporal coherence is related to monochromaticity, a single electron with a specific energy (monochromatic) will present temporal coherence. However, when a beam of many monochromatic electrons is considered, the temporal coherence can be broken despite the monochromaticity.
Collimation: The emitted electrons from the gun should be as parallel as possible to the microscope axis avoiding the divergence of electrons. A beam is named collimated if its wavefront is plane or alternatively if the trajectories of all their electrons are parallel to each other.

2.3.2 *On the electron gun* *

The material that emits electrons should be a metal conductor. The charges or electrons in a conductor lays on its surface, otherwise there would exist an internal electric field in the conductor that generates currents (movements of electrons). It is plausible that those regions of the metal with higher electron density present higher probability to emit electrons under certain conditions. Thus, describing the shape of the emitter, means to analyze the charge density on the surface under different curvature radius. It is known that the electric field E in the surface of a conductor is proportional to the superficial charge density, σ, i.e. $E = \sigma/\epsilon$, with ϵ the dielectric permittivity constant. So, there will be more charges for those regions with higher electric fields. The conductor surface is equipotential (note that the metal and the equi-potential surface have the same shape), and because of relation $\mathbf{E} = -\nabla\phi$, with ϕ the scalar potential, the electric field will be higher in those regions with smaller curvature radius. In other words, the electric field and,

therefore, the electron density will be higher in sharpened regions. This means, that the shape of the electron gun should be as sharpened as possible in order to maximize the charge density and the electric field.
The reason of using sharp shapes of metal as electron guns can also be supported with arguments of spatial coherence. If all electrons are emitted from the end of the tip, the phase difference in frontwave will be close to each other presenting spatial coherence.
Regarding the structure of the electron gun, it is composed by three main components in the following order: the *cathode* or filament or cathode, the *Wehnelt cylinder* or grid, and the *anode*, see Fig. 4. The Wehnelt is a metal hollow cylinder that contains the cathode/tip, and presents a hole in its basis. A high differential voltage is applied between the anode and the cathode in order to guide and allow the electron current, thus, by modifying the potential of the Wehnelt cylinder, the electrons can be focused increasing the directionality.

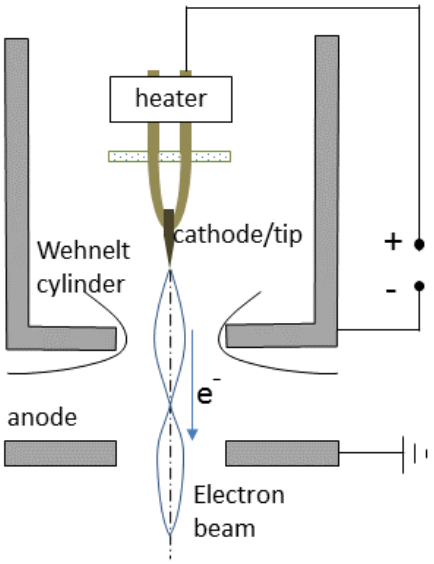

Figure 4: Electron gun structure with all elements.

2.3.3 *Types of electron guns* *

Despite the sharp shape helps to concentrate the charges, electrons ought to break a potential barrier, ϕ_{work} to exit from the metal towards the sample. To overcome this potential two physical mechanisms can be used: thermionic emission and field emission. In the first case, the energy required to leave the tip needs to be greater than the work function, while in the field emission, electrons escape by tunneling effect.

1. *Thermionic Electron Guns*: This kind of device achieves the electron emission by heating the tip of the metal. Thus, by thermal energy the electrons acquire enough energy to overcome the potential barrier. A rigorous explanation of

this kind of effect requires the use of the Sommerfeld-Bethe or Drude model [87]. However a shortcut can be taken via definition of intensity, I as the flux of charges, q per time unit, t. The goal will be then to determine the electron current. If n represents the number of electrons per volume unit $e \cdot n = dq/dV$ and the volume element is decomposed as product $dV = Sdz$ where, z is the electron movement direction and S the area of the transverse surface, then the mean current

$$\bar{I} = \frac{dq}{dt} = en\frac{dx}{dt}S = en\bar{v}, \tag{10}$$

being $\bar{v}$ the mean speed of electrons. Determining $\bar{v}$ requires the knowledge of the density function, for high energies larger than Fermi energy (last occupied energy level) the Boltzmann statistics can be applied [116]. As a consequence, the density function for electrons with energy E is given by $\rho(E) = e^{-\beta E}$, with $\beta = 1/k_B T$, being T the temperature and k_B the Boltzmann constant [1].

$$\bar{I} = enS\bar{v} = enS\int_{v_{min}}^{\infty} ve^{-\beta E}. \tag{11}$$

By considering a classical approach, the electrons with energy greater or equal to the potential barrier can escape from the metal gun. It means that the kinetic energy $K = 1/2m_e\bar{v}^2$, determines the minimum speed of the exit electrons $\bar{v}_{in} \geqslant \sqrt{2e\phi_{work}/m}$, due to $K = e\phi_{work}$. This allows to write an explicit expression for the intensity

$$I = enS\left(\frac{2k_B T}{m_e}\right)^2 e^{-\frac{\Phi_{work}}{k_B T}}, \tag{12}$$

which is known as *Richardson's law* [63]. The key point of this law is the dependence on the temperature, the higher temperature, the higher currents are achieved. By heating the metal, the thermionic guns make use of this fact. However, the metal ought to resist temperatures higher than 1000 K. Only a few material as the Tungsten or lanthanum hexaboride LaB_6 with shape of hairpin can work at temperatures of $2,500 - 3,000$ K and $1,400 - 2,000$ K respectively without melting [130].

2. *Field Electron Emission Guns*: To leave the tip, the electrons ought to overcome the potential barrier, ϕ_{work}. Electrons cannot be localized with precision laying on space region under certain probability given by modulus square their wavefunction, ψ. This gives to electron an opportunity to be out the tip when its energy is smaller than the work function, in other words electrons might leave out the tip by tunnelling effect. Field Electron Emission Guns try to support the tunnelling effect by an electric field. Finally, a classical explanation of this kind of gun can be carried out by considering the time-independent Schrodinger equation for the electrons, thus,

$$\left[\frac{\hbar}{2m_e}\nabla^2 + (E - V)\right]\psi = 0, \tag{13}$$

1 This is just an approximation because a rigorous proof should use Fermi-Dirac approximation

with $\hbar = h/2\pi$, E the electron energy and V the potential. To solve this equation requires the knowledge of the potential function, V which is modelled by a more realistic barrier called Schottky-Norheim[115, 151] given

$$V = e\phi_{work} - \frac{q}{8\pi\epsilon z} - eEz, \tag{14}$$

being z the distance along the microscope column to the electron gun. Note that this barrier presents a singularity at $z = 0$ and a straight behaviour at far distances, see Fig. 5. The wavefunction solution of Eq. (13) will determine the probability of the electron to be in or of the tip.

3. *Schottky Emission Guns*: Modern electrons guns makes use of this effect, which can be considered as a combination of the thermionic and field emission effect. The fundamental idea is to heat the cathode to reduce the potential barrier working out well the emission field. This is the technology employed by FEI Titan Transmission Electron Microscope series.

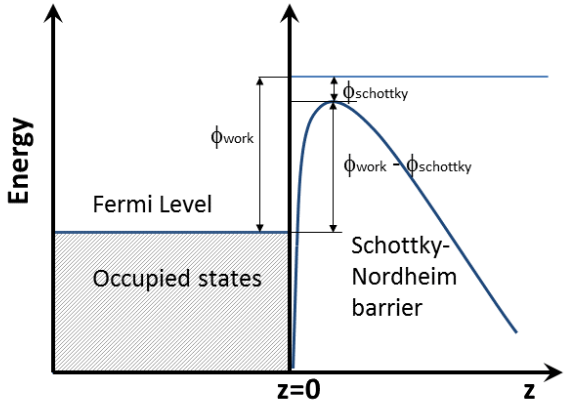

Figure 5: The tip is a metal which electrons occupy the Fermi level. To exit the tip a potential barrier must be overcome, the Schottky-Norheim and the work function are shown.

2.4 ELECTRON LENSES. MAGNETIC LENSES *

A lens is every element that makes a collimated radiation beam converge towards a fixed point. In light optics, the working mechanism is a local variation of the refractive index of the material in which the radiation is propagated. The electron microscope works in vacuum, so that, there is no change of material medium, in which the interaction electron-matter deflects the electrons. As a consequence, the mechanism of focusing electrons must be different. The main difference between light and electrons is mass and electric charge. Lorentz equation shows how a particle with charge, e, and speed $\mathbf{v}$ in an electromagnetic field, $\mathbf{E}$ and $\mathbf{B}$ suffers the action of a force $\mathbf{F}$ given by

$$\mathbf{F} = e\left[\mathbf{E} + \mathbf{v} \times \mathbf{B}\right]. \tag{15}$$

Then, electromagnetic fields can be used to change electron trajectories, which cast two possible mechanisms to get electron lenses: electrostatic and magnetic lenses, being last one the most common and the ones explained in this section. The starting point to define magnetic lenses will be the paraxial approach, and the lenses geometry.

Despite that there are many kind of lenses with different geometries as they can be spheric, aspheric, cylindric, or toric among many other, the most general approach will consider them as rotationally symmetric. This constrain is critical because it imposes a restriction in the shape of the magnetic field, it also must be rotationally symmetric. Moreover, it should mainly depend on the distance to the center of the lens, $B = B(z)$ be confined with a maximum at the center $z = 0$ and a smooth decay as the distance from the center is increased, see Fig. 6. However, the dependence $B = B(z)$, is slightly incorrect. Consider a given plane $z = constant$, an electron travelling along the z axis will pass trough the lens without suffering any deviation, meanwhile, an electron that passes at a given distance r from the z axis will suffer a Lorentz force to be focused that must be higher as higher is the distance to the z axis. It means that the magnetic field must also present a radial component, B_r. Cylindrical coordinates seems to be the most suitable frame reference, thus $\mathbf{B} = (B_r, 0, B_z)$. To determine how strong the radial component is and the relation with the component B_z the flux of magnetic field casts a simple answer [130], see Fig. 6. Maxwell equations impose $\nabla\mathbf{B} = 0$ and, therefore, the flux of the magnetic field through a cylinder of radius r and height Δz, ϕ_l will be $\phi_l = B_z\pi r^2 - B_z(\Delta z)\Delta z + B_r 2\pi r\Delta z = 0$, casting

$$B_r = -\frac{r}{2}\frac{dB_z}{dz}. \tag{16}$$

To obtain this magnetic field coils are used which are almost encapsulated with an iron box except for a small aperture. The encapsulation aims of confine the field increasing its intensity in the aperture area, see Fig. 6. Note that the magnetic field ought to be enough intense to achieve a significative change of the movement direction in a short space.

Regarding to the electron trajectory it is expected a helicoidal movement. The helix pitch is related to the intensity of the magnetic field, so that the shortest pitch will take place in the lens place, getting larger steps as the electron goes away of the center. The paraxial approximation and the high energy of electrons make that the speed along the z-axis, $\mathbf{v} \approx v_z = constant$, and therefore the azimuthal speed must be increased with the magnetic field, being maximum in the plane of the lens. Under all above considerations the Lorentz law Eq. (15) can be applied to determine the behaviour of the electron, in cylindrical coordinates results

$$F_r = m\ddot{r} - mr\dot{\theta}^2 = -eB_z r\dot{\theta}, \tag{17}$$

$$F_\theta = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = eB_z r\dot{r} - erB_r\dot{z}, \tag{18}$$

$$F_z = m\ddot{z} = eB_r r\dot{\theta}. \tag{19}$$

The assumption of paraxial approximation joint to $v_z = cte$ verifies Eq. (19). In addition, this equation system allows to quantitatively describe the electron move-

ment, by substitution of Eq. (16) into Eq. (18), $d/dt(mr^2\dot{\theta}) = d/dt(er^2B_z/2)$, which results in the Larmor frequency [63]

$$\dot{\theta} = \omega_L = \frac{e}{2m}B_z. \tag{20}$$

Note that the frequency reaches the maximum value when $B_z(z)$ is maximum. The helix pitch will be $s = v/\omega$, being minimum when frequency is maximum, as it was predicted. Finally, Eq. (20) can be substituted in Eq. (17) to get

$$m\ddot{r} = -\frac{e^2}{4m}B_z. \tag{21}$$

This equation describe the radial electron movement, the right side is always negative and the left side is a second derivative, therefore, it shows that the radial behaviour should be concave function with maximum in the lens plane. The exact trajectory of the electron can be calculated although it involves a long calculus for dealing and solving with Eq. (21) that is out of the scope of this introduction.

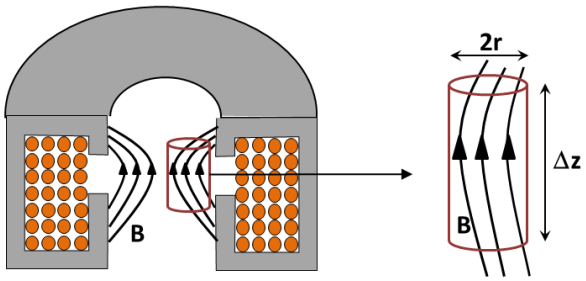

Figure 6: Magnetic lens composed by an iron (gray) encapsulated with a coil inside it (orange circles). The magnetic field, **B** presents cylindrical symmetry. And the flux of magnetic field thought a closed surface, in this case a cylinder of height Δz and radius r is shown.

2.4.1 *Aberrations*

An aberration is every deviation from the ideal behaviour of an optical system. An ideal lens produces spherical wavefront at its exit, the deviations of this sphericity is known as aberration, and the result is that the image of a single point will not be a single image point. The optical design tries to avoid them, unfortunately they remain despite these design efforts, in that case they must be characterized and corrected in the image. In the case of TEM applied to biological samples aberration take critical importance, due to the use of thin specimens and a weak interaction with the sample implies that the sample is coarsely understood as almost *transparent* to the electrons, and imaging transparent object present some drawbacks due to the absence of contrast. There are many solutions to this problem. However, the most spread and the used in the field is to introduced a defocus to increase the contrast, nevertheless the use of aberrations can be also considered, Section 2.6.
It is assumed mono-energetic radiation, therefore all considered aberrations will

take place at a given wavelength. To describe them it is necessary to determine the phase difference δ between the spherical wavefront produced by an ideal lens (paraxial approximation) and real optical system wavefront. The natural way in optics to express phase differences is by means of Zernike polynomial [101]. If the Zernike polynomial is truncated up to forth order and assuming rotational symmetry, then only the even coefficients of the polynomial are non zero. Those terms are the defocus (second order), spherical aberration (forth order), and eventually the phase shift (zero order). Each term of the Zernike polynomial represent a specific type of aberration. In most microscopes the most representative term are defocus and spherical aberration, thus, the introduced phase shift is given by in Fourier space with frequency k, as

$$\xi(k) = \frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4k^4 - \frac{1}{2}\Delta z\lambda^2k^2\right), \tag{22}$$

where C_s and Δz are called spherical aberration and defocus. It is noteworthy that defocus is not properly an aberration because it is produced by a wrong position of the screen or sensor. However, it is considered as an aberration because it implies a phase deviation.

2.5 THE SAMPLE *

The sample is placed on a metallic grid which is introduced in the microscope. Each hole of the grid is also composed by a deposited carbon grating for which each hole will be analyzed by the microscope. After a set of biochemistry processes the complex under study is gotten in pure state in aqueous solution. The main drawback is the incompatibility of the high vacuum of the electron microscope with liquid samples. As an alternative, the water solution is dropped on the grating and a treatment to get a solid sample is applied. This can be carried out by the following methods

1. *Negative stain*: The macromolecules are in aqueous suspension and a small concentration of uranyl acetate is diluted [2]. This solution is settled out on a carbon grid removing the excess of solution to achieve a thin as possible plate. The uranyl effect dries and wraps the sample. Due to the high density difference between the uranyl salt and the macromolecular complex, the contrast in the image is achieved. Note that in the micrograph particles looks white on a black background. However, what is imaged is the interface uranyl-complex, meaning that it only allows to reconstruct the overall shape of the complex and high resolution information is lost. Note that if the uranyl salt does not cover the whole macromolecule, information about the non-covered part will be lost, this is a problem for large particles. Despite this technique has been substituted by vitreous ice techniques, it remains as first check of the sample or when an initial volume (structure at low resolution) is needed due to its high contrast.

2. *Vitreous ice*: It was introduced by Taylor and Dubochet [31, 177] and its concept consists in freezing the sample as quick as possible to get a non-

crystalline structure called vitreous ice. The fast freezing process allows to get the macromolecular complexes in their native state (avoiding drying, flattening effect, guarantying the full particle is embedded and artifacts). In addition to this biological benefit, the use of ice embedded specimens keeps the high resolution information. Note that now the densities of ice and complex are similar which means that micrographs will present low contrast and not only the interface is imaged. Moreover, considering that the probability of scattering is increased with the atomic number, the conditions for the weak phase object approximation that will be shown in next section are more properly fulfilled. Finally, in terms of image, looking the particles in the micrograph look as black on a white background. The use of this kind of samples joint to development of direct detector has represented a revolution of cryo-electron microscopy in last years.

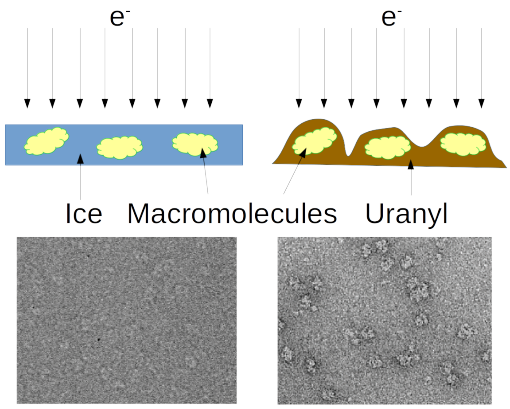

Figure 7: (left) cryo samples and (right) Negative stain, with their corresponding micrographs (down).

2.6 IMAGE FORMATION

Despite electron lenses allows to treat the electron microscope in the framework of the geometrical optics, the imaging process requires to go in depth with electron matter interaction. As starting point consider an object $u(\mathbf{r})$, the objective will be to determine the output image $v(\mathbf{r})$.
To do that, it is known that in real systems Maxwell conditions are broken due to aberrations and diffraction limit. This means that the image of a point-object will not be a point (correspondence one to one object-image), instead of that it will be a spot the called *Point Spread Function - Point Spread Function (PSF)*. In other words, the PSF is the image of a point object under the optical system. Electrons only can impact in a single position, as a consequence, the PSF is the probability distribution to hit the screen at a specific position. The spatial distribution is obtained on the screen when many electrons are used, getting many impacts and mapping the shape of the probability function i.e. the PSF. If an object is a set of points, it is

possible to know its image by applying the PSF to all of them. This suggests the simplest imaging model in which 2D-object an 2D-image are related by a linear relation as

$$v(\mathbf{r}) = h(\mathbf{r}) * u(\mathbf{r}) = PSF * u(\mathbf{r}), \tag{23}$$

being $h(\mathbf{r})$ the PSF of the system and the symbol $*$ denoting the convolution operation. Therefore, the PSF models the behaviour of the microscope. If Eq. (23) is translated to Fourier space it results

$$V(\mathbf{k}) = H(\mathbf{k}) \cdot U(\mathbf{k}) = OTF \cdot U(\mathbf{k}), \tag{24}$$

where $\mathbf{k}$ is the frequency vector and the capital letters are the Fourier transforms, $\mathcal{F}[\cdot]$, of the corresponding lower case functions, $V(\mathbf{k}) = \mathcal{F}[v((\mathbf{r})]$, $U(\mathbf{k}) = \mathcal{F}[u((\mathbf{r})]$, $H(\mathbf{k}) = OTF = \mathcal{F}[h((\mathbf{r})]$. Last function is called *Optical Transfer Function* or Optical Transfer Function (OTF) and has the same information than the PSF, but in this case in frequency terms. This function is a key element in the analysis of optical system, for a deep study of these function see [17, 54, 79]. Unfortunately, because of the OTF is defined in Fourier space, it usually is a complex function which makes difficult its understanding. To avoid this pitfall the modulus of the OTF is defined,

$$OTF = |OTF|e^{i\,arg(OTF)} = MTFe^{i\,arg(OTF)}, \tag{25}$$

it is called *Modulation Transfer Function* or $MTF(\mathbf{k}) = |OTF(\mathbf{k})|$, and arg denotes the argument function. Note that the Modulation Transfer Function (MTF) is a real positive function and gives information about how the microscope or optical system attenuates each frequency. In electron microscopy the OTF function is named as *Contrast Transfer Function*[2] or CTF and since this moment the OTF will be referred as CTF.

Up to now the imaging mechanism is a linear model given by Eqs. (23) and (24), being the microscope reduced to a single function, the PSF or OTF. In this epigraph the underlying physics to describe this function will be exposed, see [86, 193]. To do that it is necessary determine the electron interaction with the sample and its behaviour; Schrodinger equation, gives information about the electron wavefunction, φ, under a potential interaction, V. Hence, if an incoming electron is propagated along z-axis with energy E, then the solutions of

$$\left[\frac{\hbar}{2m_e}\nabla^2 + (E - V)\right]\varphi = 0, \tag{26}$$

will determine the electron wavefunction. To solve it, some assumptions must be done.

1. *Small angle approximation*: the propagation direction of the electrons is always close to the optical axis (paraxial approximation). This condition can be written in terms of the wave vector $\mathbf{k}$ with $k = \frac{2\pi}{\lambda}$, such as $k^2/k_z^2 = 1 + (k_x^2 + k_y^2)/k_z^2 \approx 1$.

2 It should be highlighted that in optics the CTF is function that measures the contrast at different frequencies, but it different than the MTF and OTF.

2. *Classical approach*: relativistic corrections are left out, otherwise Dirac equation should be considered which complicates the solution. This means that the wave-vector will present modulus $k = \frac{2\pi}{\lambda} = \sqrt{2mE}/\hbar$.

3. *Slow varying wavefunction*: The sample should not perturb to much the electron wavefunction, and coherence in their propagation should be desired, thus, it is plausible that the wavefunction will be the product $\varphi = \psi(x, y, z)e^{ikz}$ with $\psi(x, y, z)$ a slow variation function, $|\partial^2\psi/\partial z^2| \ll |k\partial\psi/\partial z|$.

   As a consequence, the Schrodinger equation can be written as

$$\frac{\partial\psi}{\partial z} = \left[\frac{i\lambda}{4\pi}\nabla^2_{x,y} + i\sigma V\right]\psi, \tag{27}$$

   with $\nabla^2_{x,y} = \partial^2/\partial x^2 + \partial^2/\partial y^2$ and $\sigma = 2\pi m\lambda/h^2$.

2.6.1 *Weak phase object approximation and projection assumption. The contrast transfer function*

In addition to the conditions enumerated to solve Eq. (26), two more hypotheses must to be done, that are included in this subsection due to their importance: the weak phase object approximation and projection assumption.

1. *Projection assumption*: It considers that the sample is thin enough to approximate the potential of the sample as uniform. This allows to neglect the kinetic term $\frac{i\lambda}{4\pi}\nabla^2_{x,y}$, leaving $\partial\psi/\partial z = i\sigma\psi$ and the wavefunction as

$$\psi = \psi_0 e^{i\sigma\int V dz} = \psi_0 e^{i\sigma V_z}, \tag{28}$$

   where ψ_0 represents the incoming electron wavefunction. What the projection assumption implies is a huge simplification of Eq. (26) because it neglects the x,y dependence, assuming $\varphi = \varphi(z)$ as a slow varying function.

   Now the wavefunction at the output of the sample is known, however, the effect of aberrations ought to be added. They can be introduced considering that the wavefunction at the output of the microscope, ψ_{out} will be the convolution of the transmitted wavefunction, ψ convolved by the a transfer function, $\psi_{out} = \psi * h$. Introducing aberrations can be carried out taking into account that they only affect on the phase. It means a complex unitary function that modifies the wavefronts, it means that

$$H(\mathbf{k}) = \mathcal{F}[h] = e^{i\xi(\mathbf{k})}. \tag{29}$$

2. *Weak phase object approximation*: It states that the interaction electron-matter is weak. The nature of biological and the small thickness of the specimens support this hypothesis. The result is that some electrons pass through the sample without interacting with it and the rest with a weak interaction. Hence, the image is formed in the sensor by phase contrast between scattered and unscattered electrons. The weak interaction allows to approximate the wavefunction by its Taylor expansion to first order as

$$\psi = \psi_0 e^{i\sigma V_z} \approx \psi_0(1 + i\sigma V_z). \tag{30}$$

which is the called weak phase approximation. Note that the first term, 1, of Eq. (30), represents the unscattered wavefunction, meanwhile the second one is the wavefunction of those scatters that suffer a weak interaction. The explicit form of the wavefunction allows to get the image. Note that the image is an intensity pattern as the probability distribution of the wavefunction, i.e. square modulus of the amplitude,

$$I(x,y) = \|\psi_{out}\| = \|\psi * h\|^2 = \|\psi_0\|^2\|(1 + i\sigma V_z) * h\|^2. \tag{31}$$

Without loss of generality and because it represents a global scale factor it is set $\|\psi_0\|^2 = 1$. Taking into account that the weak phase object approximation implies small values of σV_z in (31) will only accepted term of first order in σV_z, therefore,

$$I(x,y) \approx 1 + 2\sigma V_z * (Im(h)), \tag{32}$$

Finally, if this expression is translated to Fourier space

$$\mathcal{F}[I] \approx \delta(\mathbf{q}) + 2\sigma V_z(\mathbf{q})Im(H). \tag{33}$$

Note that the term 1 in Eq. (32) or $\delta(\mathbf{q})$ in Eq. (33) represents a reference intensity level in the image, so that it can be neglected, and thereby

$$\mathcal{F}[I] \propto 2\sigma V_z(\mathbf{q})\sin(\xi(q)). \tag{34}$$

This expression is the most common equation in electron microscopy, just naming $\sin(\xi(k)) = CTF$ thus, the linear model (23) is verified.

2.7 DETECTORS

Imaging detectors/sensors are responsible of registering the intensity pattern obtained at the output of the microscope. It is remarkable that the obtainment of a good image is a combination of a good optical system (microscope setup) and detector. The sensor by itself lacks of sense without a good optics. Moreover, taking into account that images will present very low contrast due to long exposure times can burn the sample, thus detectors ought to have a high sensitivity along with high resolution, dynamic range, among other.
In the past, detectors were plates sensitives to radiation and had to be developed. Finally, the information was extracted by inspection of the developed and digitized film. Currently the use of sensitive-radiation plates is almost obsolete and they have been substituted by digital sensors. Their introduction, in particular, the CCD-charge couple device was carried out by Roberts and Chapman in 1980 using a simple sensor of 100×100 pixels [132]. However it is more known the work of Spence & Zuo as one of the first uses of CCD sensors for TEM applications, where combining the CCD with an electron scintillator and an optical coupler they avoided the direct measurement [36, 170]. Later the scintillator was replaced by a phosphorus screen. Also CMOS technology (complementary metal oxid semiconductor) sensors started to be used in the same way as CCD in early 90 year but exhibiting a worse performance than CCDs. They had to wait more than one decade to undertake a

revolution in the detection field.
The main advantage of digital sensors in comparison to the old photographic plates is their automatic digitization and computerization of micrographs in real time without the need of plate replacement. In contrast, radiation damage limits the life time of these devices.
The technological advances focused on the improvement of detection capabilities implied to get better the main requirements of a sensor [37, 61] as they are: efficiency (high DQE), resolution, noise level and durability or radiation resistance. The result was a new generation of detector the Direct Detector Devices (DDD) or Direct Electron Detectors (Direct Electron Detector (DED)). In fact, the development of this new kind of sensors is considered one of the mains reasons of the boost of electron microscopy in last years [29]. The DDD success key is the direct measurement of radiation avoiding the conversion of the electron beam into light using a phosphorous layer which produces a blurring of the image. For this reason, it is common to classify the detectors as direct or indirect. Moreover, the use of active pixel sensors based on CMOS technology allows a quick reading of the acquired information, as a consequence, they can acquire movies and observe the movement induced by the electron beam. Thus, in terms of imaging DDD increases the resolution with less magnification, and also allows the tracking and movement correction of particles, the result is a better MTF of the system.

2.7.1 *Detectors characterization**

Sensors are complex devices and their performance is calibrated under specific working conditions, thereby their behaviour depends on many variables external variables, a few examples are: energy of the incident radiation, temperature, tilt beam, exposure time, among other. The most known property of a detector is the resolution, however, resolution does not fit all external variables, thus they are characterized by a set of parameters. Next are only a few are the most important of them

- *Pixel size*: An imaging sensor is a matrix of identical elements called pixels (acronym of picture element). Each pixel is an individual sensor that measures the amount of radiation (light/electrons) that reach its surface, in particular, they measures the irradiance of the beam. Usually pixel geometry is square, the dimension of the edge of a pixel is named pixel size. The detection area of the pixel is smaller than whole pixel area; and the ratio between them is called fill factor. Moreover, manufacturing processes imposes a narrow, but non negligible separation between a pixel and its neighbours, this separation is called pixel pitch, see Fig. 8. Usually the smaller pixel pitch, the higher defined and sharpened the images can be.

- *Resolution*: Coarsely speaking is the capability of a system to discriminate two different points. It is the highest spatial resolution what can be solved. In the field of imaging detectors this discrimination is carried out by the total number of pixels, being the sensor resolution, the number of horizontal and vertical pixel. It is denoted as $N \times M$, where N and M are the number

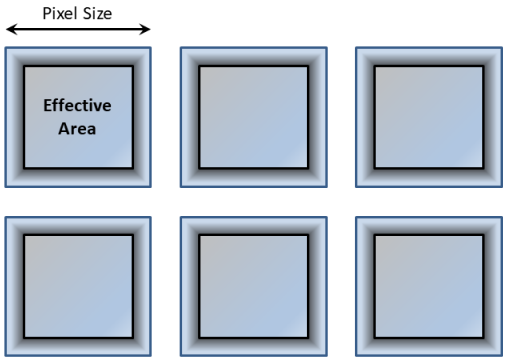

Figure 8: Sensor scheme with resolution 2×3 in which the pixel size, effective pixel area and pixel pitch are shown.

of horizontal and vertical pixel respectively. Thus, the sensor introduces a discretization of the image, with a sampling rate determined by the pixel size.
From an optical point of view resolution is given by the size of the Airy disc. However, the pixel size introduces a second limitation, the pixel size must be smaller than the Airy disc to be limited by diffraction, in other words, the Nyquist sampling limit (twice sampling rate) ought also to be taken into account.

- *Analog to digital converter (Analog to Digital Converter (ADC))*: The intensity registered in each pixels ought to be digitized as discrete values for its later processing. The ADC is responsible of carrying out this task. Its performance is measured in bits; an ADC or a sensor of N bits will convert the intensity into 2^N different gray level. The result of this element is a second discretization in the image. If the sensor matrix introduces a first spatial discretization, the ADC action implies a discretization of intensity levels. The ADC is closely related to the called linearity, a sensor is linear when the read intensity is proportional to the incident intensity on each pixel.

- *Dynamic range*: The ratio measured between the maximum value of signal/intensity, $I(\mathbf{r})$, and the Root Mean Square (Root Mean Square (RMS)) of the noise, N_{rms}. It is measured in decibels

$$DR = \frac{\max I(\mathbf{r})}{N_{rms}} \qquad \text{or} \qquad DR = 20 \log \frac{\max I(\mathbf{r})}{N_{rms}} \text{ dB}, \tag{35}$$

some authors also considers the dynamic range has the distance peak to peak between largest and smallest value of signal and noise respectively [79]. It shows the length of the measurement interval, but it lacks of sense without the ADC value. Usually, the dark noise or also called floor noise is taken as reference value, being the all values measured as the distance from it.

- *Gain*: The simplest explanation, the gain, G, is the ratio between the recorded intensity I_{rec} and theoretical intensity, I_{theo} [168],

$$G = \frac{I_{rec}}{I_{theo}}. \tag{36}$$

A perfect sensor would present a gain equal 1. Reality is different, detectors are composed by many pixels each of them with a slight differences in their gain values, but non neglected if high detailed images are required.

- *Detective Quantum Efficiency (DQE)*: It is measurement of the detector performance in term of how an input signal is preserved in the measurement process. An ideal sensor should keep the signal to Noise ratio of an input signal, however, real detectors might introduce noise and as a consequence of the conversion process. Thus the quality image can be degraded. This degradation occurs at different frequencies. The DQE measure how the performance of the detector can deteriorate the image in frequency terms. In particular, DQE is defined as the amount of noise called *Noise equivalent quanta*, Noise Equivalent Quanta (NEQ) ,to compensate certain amount of quanta (energy), it means

$$\mathrm{DQE} = \frac{\mathrm{NEQ}}{\mathrm{Q}} \tag{37}$$

Note how the closer DQE to 1, the better detector performance. In Fig. 9 a performance comparison of different sensors is shown.

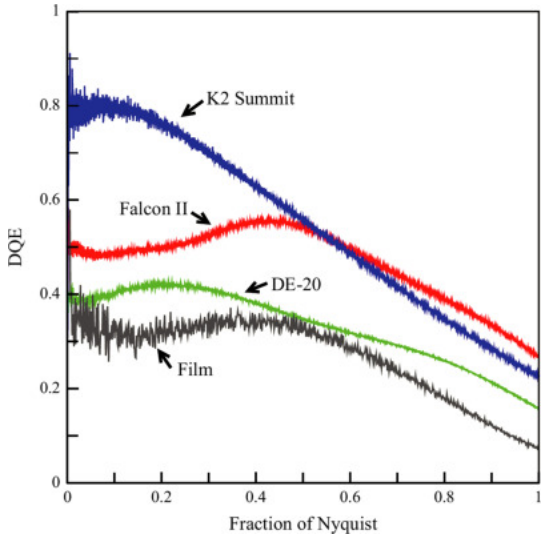

Figure 9: DQE curves for different detectors as they are: a film, and the cameras Falcon II, DE-20, and K2. (Image taken from [107])

2.7.2 *Direct detectors - DED or DDD*

The way of recording image in a detector can coarsely be classified in two ways: indirect, or direct. Classic detectors, film and CCD are representatives of the first ones, which requires the conversion of electron into signal by means of a scintillator or a phosphorus plate, transfer the signal to the sensor, and then detect and read the signal. The second class, the called DED avoids the first two steps, recording

in real time the signal without previous conversion. This is achieved by means of CMOS technology which has suffered a high boost in the last decade. In particular, they present the following advantages with respect to CCD sensors:

1. *Low electric consumption:* They remove the actives charges, it means, that if they are not active they only present parasitic current.
2. *Pixel independence*: Every pixel is autonomous and independent from the other. The CMOS sensor lacks of an external control board. Each pixel has its own control gates. As a consequence of this independence, the blooming effect is removed.
3. *Robustness to noise*: If the digitized input signal is slightly perturbed, then the action of the CMOS over it will recover the original signal.
4. *Reading Speed*: This is the most important item for EM applications. The independence of each pixels with its own electronics, and the absence of converter speed up the reading process allowing a reading in real time. Thus, the limit is given by the shutter method.

These points summarize the main performance differences between CCD and CMOS based sensors, at least in terms of architecture. In particular, robustness to noise and the quick reading time have special impact on the recorded image image. Thus, the image can keep its quality under random perturbations, and second, it allows the acquisition of movies instead of single images. This last point has been critical to reach high resolution. It is due to the high sensitivity and NEQ of the DDDs imply enough contrast in the image under less incident radiation than in a CCD camera. Finally, DDD are more resistant to radiation damage, even without any treatment of radiation hardness, being its life time around a year [38].

2.7.2.1 *Hybrid pixel detectors - DDD **

Hybrid Pixel Detector (HPD) were originally developed in Conseil Europeen pour la Recherche Nucleaire (CERN) laboratories with high succeed and the third generation of them Medipix3 currently is available , (previously Medipix1, and Medipix2). The goal was to minimize radiation damage by means of a separation of the encapsulated form the detection and reading region of the sensor. but keeping the concept of embedded CMOS. In Fig. 10 a scheme of the device is shown. Thus, the first encapsulated device or also called sensor consists in a silicon waffle. It is applied over it a potential difference. When an incident electron goes through the silicon layer a pair electron-gap is generated, the applied voltage produces a charge shift or current of electrons and gaps towards the electrodes of the sensor. This sensor is coupled to a second encapsulated (responsible of reading) by means of a contact soldering which also acts as separation piece sensor-reading. This second encapsulate is the chip or pixel matrix to pick charges from the soldering and determines the position of the incident electrons. The first encapsulated, i.e. the silicon layer is wide enough to protect the pixel matrix from radiation, in other words most electrons are absorbed by the silicon, and only and small amount can reach the pixel matrix, allowing a long durability of the sensor.

2.7.2.2 *Monolythic active pixels sensor - DDD **

Monolythic Active Pixels Sensors (MAPS) were in their initia addressed to spatial applications as a consequence they present an excellent robustness and resistance to radiation damage [16]. Its structure is simpler than HPD, being only composed by a single encapsulated of semiconductor layer, but always keeping the CMOS concept. Again, the thickness of the layer is wide enough to prevent radiation damage. The physical working principle is the liberation of charge in the semiconductor layer when a charge hits on. The use of a single layer requires a pixel matrix on it, but in this case the responsible of picking charges is a diode to finally convert them in voltage using a transistor. (See Fig. 10).

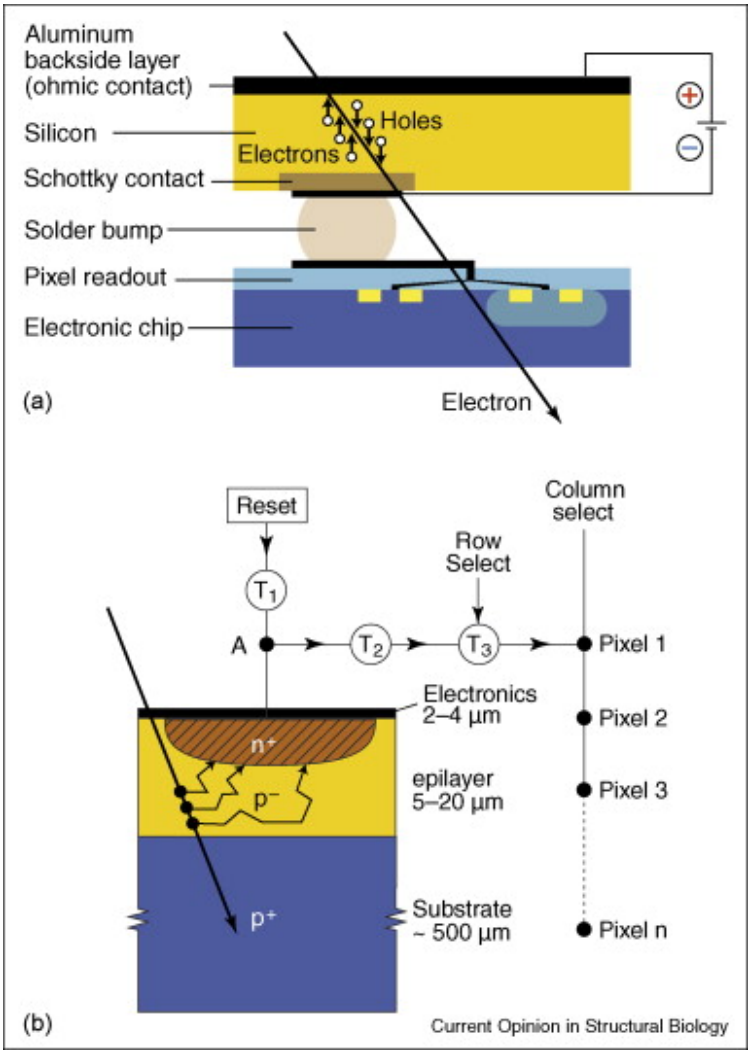

Figure 10: Working mechanism of (a) HPD and (b) MAPS. Image taken from [37]

3

SINGLE PARTICLE ANALYSIS WORKFLOW

The use of electron microscopy to elucidate the structure of macromolecular complexes has become a highlighted technique due to its ability to reconstruct the complexes in their *native state*. The set of image processing techniques that allows to determine the structure of the macromolecule are called *Single Particle Analysis* - SPA. Hence, in this chapter, the standard workflow of SPA is explained step by step. Each step performs a very specific task. However, there are several methods and algorithms that provide solutions to that specific task. Due to the complexity and variety of them, it is out of this thesis an explanation in depth. So that, the aim will be to provide the overall idea of each step avoiding the mathematical details behind them.

3.1 GENERAL CONSIDERATION OF SINGLE PARTICLE ANALYSIS

The objective of SPA is to obtain the 3D structure and if it is possible the atomic model of macromolecular complexes via electron microscopy imaging. To do that, SPA is based on two main hypotheses:

1. *Homogeneous sample*: All or most of the specimens in the sample are identical copies of the same macromolecular complex in the same state but presenting different orientations. When this identity condition (identical copies) is broken it is said that the sample present heterogeneity. Dealing and reconstructing heterogeneous samples is currently a challenge in SPA.

2. *Projection assumption*: The image called *micrograph* can be considered a projection of the sample under a given magnification of the microscope. This hypothesis was mentioned in image formation, see Chapter 2 and Fig. 6, but its importance is critical.

The first one establishes an identity condition while the second is a scale condition. Both hypotheses represent the basis of reconstruction methods. It is important to remark that the identity condition is weaker than the scale condition, because slightly different specimens might cast almost the same projections. In other words, certain heterogeneity is allowed. The heterogeneity effect will be observed as a blurring in the reconstructed map around the heterogeneous area.

3.1.1 *Central Slice theorem*

To reconstruct the structure, it is first necessary to select each projection of the macromolecule, called *particle* from the micrographs and combine all of them properly. The set of all possible directions is named *projection sphere*, and the projection direction of a particle is called *angular orientation*. Assuming that the sample does

not present preferred orientations, a higher number of particles will allow to obtain more information, because the projection sphere will present a better coverage and as it will be shown the *Signal to Noise Ratio* (SNR - ratio between the powers of signal and noise) is increased.
The reconstruction problem is the following: given a set of particles with a determined angular orientation, is it possible to recover a 3D structure of the complex compatible with all particles?. There are many ways to solve this question (see [165]), however the *Central slice theorem* [183, 201] provides a remarkable answer by establishing a relationship between particles and the 3D structure via Fourier Space. It can be enunciated as it follows: *Given a 3D structure defined by a function* $S(x, y, z)$ *and* $P_\alpha(S)$ *a projection (particle) of the structure under certain direction* α*, where P denotes a projection operator along,* α*, then,*

$$D_\alpha[\mathcal{F}(S)] = \mathcal{F}(P_\alpha(S)). \tag{38}$$

where D_α *is the plane defined by the normal vector with direction* α *that passes through the origin of the Fourier Space*. In Fig. 11 the theorem is graphically explained, to reconstruct an object, given a set of particles, is it necessary to determine the Fourier transform of those particles and use them to fill the Fourier space with the projection direction. Finally to recover the structure the Fourier inverse transform is applied. Note that the absence of particles with specific angular assignment, will be reflected as gaps with zero values in Fourier Space. If the lack of angular assignment is large enough to cover a cone or a wedge, the uncertainty is named *missing cone or missing wedges*, respectively, and the structure will be elongated in real space.

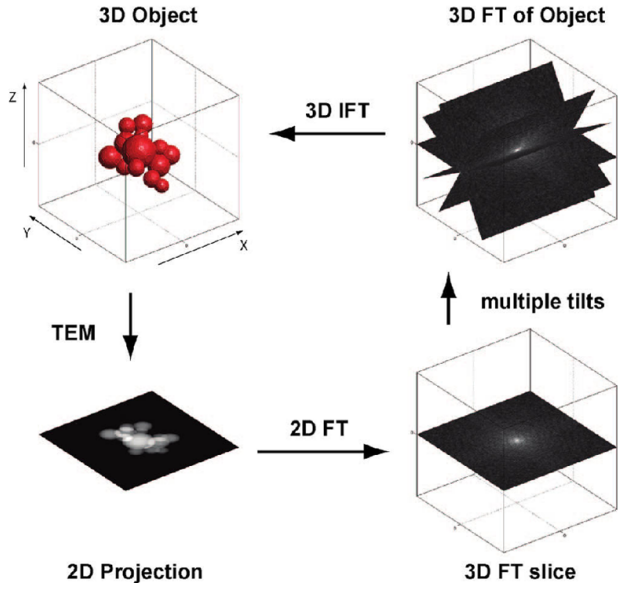

Figure 11: Explanation of the central slice theorem.

3.1.2 *The SPA workflow*

Even though the reconstruction method given by the central slice theorem shows how the structure can be obtained from particles, their orientation is unknown. This is one of the main problems to solve in SPA, but there are still some other remarkable problems which need to be solved too. The set of steps to solve them define the SPA workflow. In Fig. 12 an scheme of these steps is shown.
The workflow starts with the acquired images named micrographs or movies. If they are movies, a *movie alignment* is required. Next, the CTF of the microscope should be estimated to correct possible image aberrations. Nevertheless, the reconstruction requires to select and extract the particles in the micrograph, this process is named *picking* step. Then, the particles are corrected by the CTF estimation, and classified in the called *2D Classes* grouping similar particles. The idea is that all particles of the same class should be projections under close directions of the macromolecule. Thus, by averaging them, the SNR is increased, helping to the determination of an *initial volume* in next step, i.e. a first and coarse estimation of the macromolecule structure. However, it must be remarked that classification can also help as screening step to remove bad particles, and check the quality of the data. Finally, by means of the initial volume, particles can be reclassified by trying to find different conformations of the same volume and refining the map by better angular assignments.

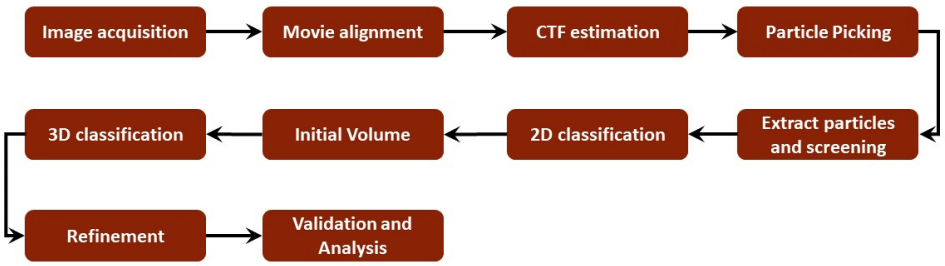

Figure 12: Main steps of the basic workflow of SPA.

3.2 MOVIES AND MICROGRAPHS

The structural information about the macromolecule is given by the contrast between the ice (background) and the complex [55] (signal), the higher contrast the better structural details can be observed. The use of vitreous ice samples limits the contrast due to the small differences in density/atomic number between ice and specimen. Alternatively, electron dose can be incremented, unfortunately, radiation damage limits the applicable electron dose [49, 71]. The development of *Direct Detectors* allowed to alleviate this problem opening new horizons. Thus, instead of acquiring a single image called *micrograph* under an electron dose, new detectors register a set of *frames* creating a *movie* under the same electron dose. Electron dose means number of electrons per surface unit. Note that along the movie record, first and last frames are exposed to different radiation doses, being the first one lowly

radiated and the last one highly radiated. However, the electron dose per frame is globally reduced.

3.2.1 *Beam induced movement. Movie alignment*

The electron beam induced movement (BIM) in the sample is a consequence of the electron-matter interaction. However, the underlying physical mechanism of this movement is currently unknown. Nevertheless, there are some hypothesis like: inelastic scattering, charge sample, relief stresses, or deformation of the ice layer among others [50]. The high acquiring speed of DED allows to track the macromolecule movement in the movie, i.e. *particles* exhibit a displacement with respect to the previous frame.
The contrast in the frames is very poor because they present a low SNR, considering the ice as noise due to its lack of structure. Since this moment noise will make reference to the ice. Hence, instead of working with frames as starting point of the structure reconstruction process, first step consists in increasing the SNR by combining all the frames of the movie by correcting the motion induced by the beam. This task is called *motion correction* or *movie alignment*, and the resulting image presents higher contrast than the frames and by language extension it is named *micrograph*. Note that a micrograph is the single image resulted of a single measurement in the microscope. However, when the microscope acquires movies, the term micrograph makes reference to the result of combining all the movie frames into a single image.

3.2.2 *Movie alignment:*

The blurring that appears as a consequence of beam induced motion has the effect of limiting the resolution or level of detail that can be achieved in the reconstruction [97]. To correct the particle movement due to the induced motion by the beam and increase the contrast, a characterization of the movement is required. The lack of awareness about the physical mechanism of the BIM entails two possibles correction strategies.

1. *Global*: The BIM is produced along the same direction for the whole frame. This approach is used by *MotionCor* [97]. The idea of this algorithm is to estimate the relative shift between two frames using correlations. Once the displacements are known the movie alignment is carried out. Despite *MotionCor* can be considered as the first alignment method, the corresponding measures show that this motion is local, involving a shift and a particle rotation [18].

2. *Local*: The motion induced by the beam is local and different particles may present different movements. The methods of *Optical Flow*[4], *alignframes_lmbfgs* and *alignparts_lmbfgs* [136], *Unblurr* and *Summovie*[55] or *MotionCor2*[202] belong to this group.

For a deeper explanation of all these methods, a movie alignment review can be found in [131]. The common approach to movie alignment is to perform an

alignment of the whole frame with *MotionCor*, and then a local alignment. This is somehow the strategy of *MotionCor2*, in particular, a physical model of sample behaviour is proposed and validated with the alignment. An alternative to the use of these methods is the *particle polishing*[146], in which the macromolecule is reconstructed using the averages of the movies, and then, the 3D structure is used to correct in the frames the inaccuracies introduced by the BIM effect reconstructing again the structure. In Fig. 13 the difference between micrographs, movies and the objective of the alignment is shown.

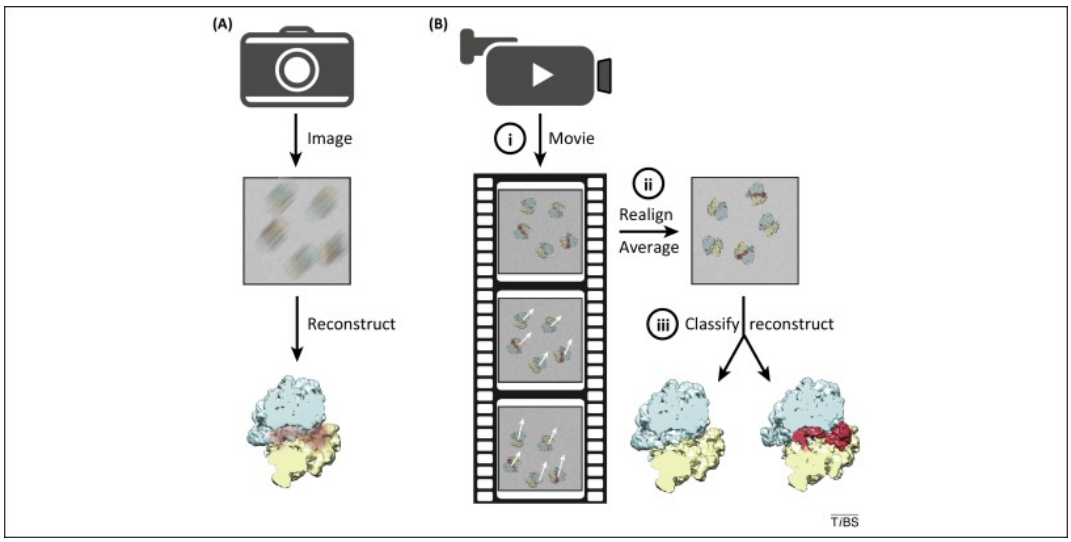

Figure 13: (A) Old measurement, the image/micrograph is obtained. (b) A movie recorder as a set of frames. The BIM and particles are shifted. As a consequence an alignment should be carried out to correct the movement. Image taken from [9].

3.2.3 *Gain correction*

The behaviour of all pixels in the detector might differ presenting slight differences in their radiation sensitivity. The physical origin of these differences may be due to many reasons like: the surface of all pixels is not the same, or presents a slight tilt, thickness, or doping differences in the semiconductor layers, current un-stability, or temperature, among others. The result is a different response under the same input radiation, some pixel will measure stronger signal than others, or even might be dead with lack of response at all. To correct these possible slight but non-negligible differences, a *gain correction* is carried out, where the gain concept models these deviations from the ideal behaviour of the sensor. Rigorously, the gain is defined as the ratio between the recorded image, I_{rec} and the theoretical one I_{theo}, for the pixel $\mathbf{r}$, it means

$$I_{rec}(\mathbf{r}) = G(\mathbf{r}) \cdot I_{theo}(\mathbf{r}). \tag{39}$$

Note that an ideal sensor should present unit gain, $G = 1$, being the acquired image equals to the ideal image, $I_{theo} = I_{rec}$. To give an idea about the deviation due to non-ideal behaviour, the gain error in DED is normally lower than 1%. The gain

correction depends on the application, and normally in commercial applications such as photographic cameras it is not necessary or it is automatically corrected by the own device. Usually, the sensor manufacturer provides information about the sensor gain. In Fig. 14 a gain image, matrix $G(\mathbf{r})$, is shown, note how the gain is not uniform.
In the field of cryoEM, the first correction of these effects was proposed to obtain the same statistics for all pixels [6]. To do that, the whole set of acquired movies is used to determine the average and the variance, performing then an intensity normalization. It was observed how the gain is susceptible to change along a microscope session. Thus, a dynamic measurement of the camera gain was proposed [168]. This measurement also helps as quality control and refinement of an initial gain estimation[168] along the measurement process. The main idea is to estimate the statistical differences in the histograms of gray levels for rows and columns. Note how these distributions are sensitive to gain variations.

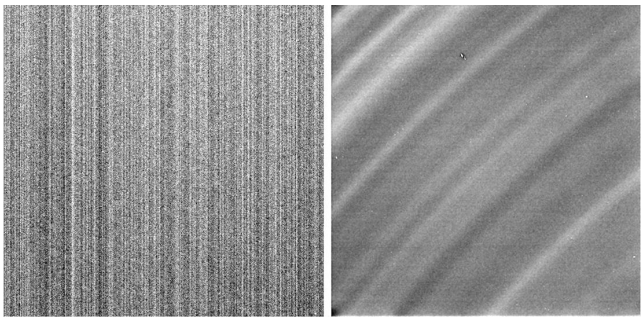

Figure 14: Experimental gain image for (left) the dataset of the EMPIAR entry 10010 using as a detector DE12 and (right) the dataset of the EMPIAR entry 10025 using as a K2 detector. Image taken from [168]

3.3 CTF ESTIMATION

The performance of the microscope in terms of imaging is modelled by the CTF. Despite some information about the microscope is known, acquiring conditions might change so that the CTF must be determined in the image processing step. The objective of CTF estimation is to quantify the effects that deviate the image formation from the ideal image. In other words, the CTF is estimated to correct the aberrations, this should increase the SNR and the quality of the reconstructed macromolecule. In the previous section, it was explained that the recorded image is the convolution product of the Coulomb potential and the CTF function, and therefore, in Fourier Space the Fourier transform of the Coulomb potential is multiplied by a sinusoidal function, the CTF. $\mathcal{F}[I] \propto 2\sigma V_z(\mathbf{q}) \sin(\xi(\mathbf{q}))$. Thus, to determine the CTF it is necessary to explore the Fourier transform of the acquired image. From this point different approaches emerge to determine the expression of $\xi(\mathbf{q})$. However, it is remarkable that the CTF expression is only valid for monochromatic and coherent illumination, under the weak phase object approximation and projection assumption, and as a consequence, the analytic expression of the CTF may vary as

it can be found in many publications, [179]. The most common approach begins by calculating the power spectral density (PSD). i.e. the squared modulus of the image Fourier transform $\text{PSD} = |2\sigma V_z(\mathbf{q})\sin(\xi(\mathbf{q}))|^2$, This function represent a pattern of concentric fringes also known as Thon rings, see Fig. 15. Note how the zeros of the PSD coincide with the zeros of the CTF. This constrain allows to determine the argument of the sinusoidal function, i.e the CTF.
The use of Volta phase plates affects the CTF by introducing a phase shift, δ, whose value is ideally $\delta = \pi/2$ turning the sine function, $\sin(\xi(\mathbf{q}))$, that defines the CTF, into a cosine function, $\sin(\xi(\mathbf{q}) + \pi/2) = \cos(\xi(\mathbf{q}))$. This shift should be considered as an extra parameter in the CTF estimation methods. Thus, the analytic expression of the function $\xi(\mathbf{q})$ given in previous chapter, see section 2.4.1, can be rewritten as

$$\xi(k) = \frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4k^4 - \frac{1}{2}\Delta z\lambda^2k^2 + \delta\right), \tag{40}$$

There are many and different methods to compute the CTF. However, here there will be only highlighted those ones which belong to the current state of art. Perhaps the most known is CTFFIND in its version 3 [110] and 4 [133], which performs a fitting between the coefficients of the theoretical polynomial from the PSD, in particular, spherical, aberration, astigmatism, defocus, and eventually phase shift if a phase plate is used. Other approach, named FASTDEF [163, 184] is carried out by performing a fitting to the Zernike polynomial, which should be able to determine (if exist) higher order aberrations. The current trend of using GPU for image processing computing shows alternative methods, as it is the case of gCTF[199] which also allows a local defocus correction per particle.

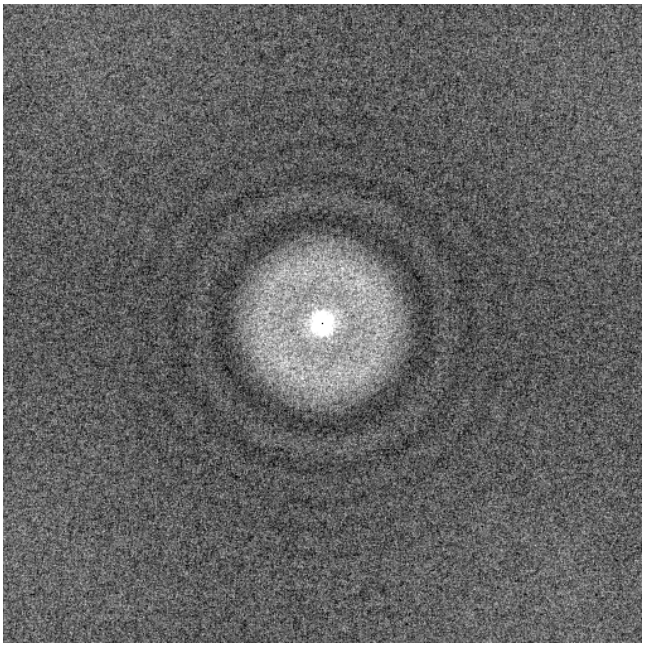

Figure 15: PSD calculated from a micrograph, the image show the Thon rings that are used to estimate the CTF.

3.4 PARTICLE PICKING

The projection assumption considers that all particles in the micrograph are projections under different points of view of the same macromolecule. Particle picking is the step in which all those particles are selected and cropped from the micrograph to undertake the reconstruction process. In other words, particle picking is the step in which particles are identified in micrographs. An example is shown in Fig. 16. This is a critical step in the SPA workflow because it determines the first raw particles to reconstruct the macromolecule. In terms of image processing, the identification/pattern recognition problem has been broadly solved, unfortunately, micrographs present a very low SNR, and therefore, the particle identification is more complicated. Since the beginning of SPA, particle picking has been and still is a critical point in the SPA workflow, so many methods were developed addressed. According to Nicholson et. al. [113] they can be classified, according to their identification algorithm, in: Template matching, edge detection, intensity comparison, texture-based method, or neural networks. However, many methods cannot be strictly assigned to an specific group, due to they make use of several features. As a consequence, it might be convenient to establish a classification in terms of usability. Hence, the methods can be coarsely classify as

1. *Manual methods*: This kind of picking is included only for completeness. The user must identify with his/her bare eyes the particles in the micrograph and select them one by one. It implies long picking times and possible bias as a consequence of user subjectivity.

2. *Semi-Automatic methods*: Previous user interaction is required to guide the picking process. The information can be the particle size, the sensitivity of the method, or a variance value, among many other parameters depending on the algorithm. However, a distinguished place in this group is occupied by template-matching and feature-template methods, in which the goal is to determine an image called *template* which presents similar shape or features with the particles to be picked. Next, by means of correlations (or other similarity metrics) between micrograph and template can be detected. Note that a template is every pattern that allows to identify the particles, thus, simple templates can be a disk or a Gaussian. However, more complex templates can be considered, as it is the case of [3, 174] were the user manually picks a small set of particles and the algorithm learns about the picked particles creating its own template. Alternatively to this template-learning technique, the user can provide one or multiple templates to identify particles in micrographs via cross-correlation [85, 149]. Moreover, GPU computational resources have been used to speed up the process. Thus, other approaches to the picking problem which make use of GPU under the concept of template-matching are available as it is the case of *gEMpicker* [78] or *Gautomatch* [80]. Alternative semi-automatic solutions out the scope of template-matching also exist: generally they work applying a transformation to the micrograph, that allows to extract the particle positions. A representative example of this group *DoG-Picker*, which computes differences of gaussians followed by thresholding [192].

3. *Automatic methods*: This kind of algorithms has been recently introduced and avoid the introduction of a bias as a consequence of human action or the type of selected template. There are some algorithms from the classical image processing methods like *APPLE Picker* [70], in which the template is automatically chosen by the algorithm. However, most of the automatic methods are based on deep learning solutions making use of convolutional neural networks (Convolutional Neural Network (CNN)). These methods, like *DeepPicker* [195] or *DeepEM* [204], define a CNN model as classifier, which is trained with several data sets. Thus, by means of a moving window, the model is able to discriminate particles from noise when a new micrograph is provided. A different solution is proposed by *crYOLO* [194] which makes used of the classification framework "you only look once´´ [129] as alternative to the moving window.

Once particles have been selected on the micrographs and they are extracted. However, picking methods works properly but there are many false positive like: artifacts, wrong picked particles or simply noise picked as particles. To get a true set of particles or at least to minimize the impact of those false positives a *screening* step is usually carried out. In this step, particles can be sorted by a z-score, considering similarity metrics based on morphology or SNR, among other, [185].
In the screening step of the SPA workflow it is also common the usage of denoising and image restoration techniques to get cleaner particles. Hence, *Wiener filters* or more sophisticated techniques like *Covariance Wiener Filtering*[15] are used. Similarly to picking algorithms deep learning approaches are also proposed REF.

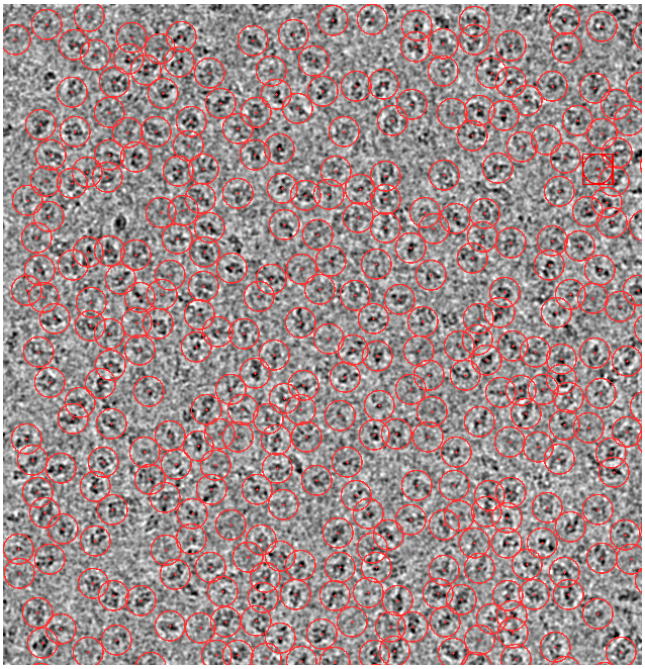

Figure 16: A beta-galactosidase micrograph fully picked. All particles were selected on the micrograph to be extracted in a later step.

3.5 2D CLASSIFICATION

A common SPA project contains from several thousands of particles to a few millions. Remind that the purpose is to recover the 3D structure of the macromolecule from all these particles. Unfortunately, it presents three drawbacks: 1) all picked particles exhibit a very low SNR, which makes it difficult to determine their angular assignment. 2) Moreover, the lack of information about the angular assignment, along with the large number of particles, implies a huge computational effort to elucidate a 3D structure. 3) Taking into account the condition of *Homogeneous sample*, all particles should be copies of the same macromolecule, nevertheless, despite the usage of screening methods to prune false particles in the picking step, many artifact particles are picked (empty particles, contaminants or undesired particles among others), and the homogeneous sample condition is broken as well as the projection assumption. To solve these problems particles are grouped in the so-called *2D-classes*. A 2D class is defined as the subset of particles with similar characteristics under certain error tolerance. Ideally, in SPA workflow the characteristic is the same angular assignment. The mean of all particles inside the same class is named *class average*, *class representative*, or *2D averages*. In Fig. 17 the concept of class is graphically shown.

Grouping particles in classes simplifies the problem, because it allows to work

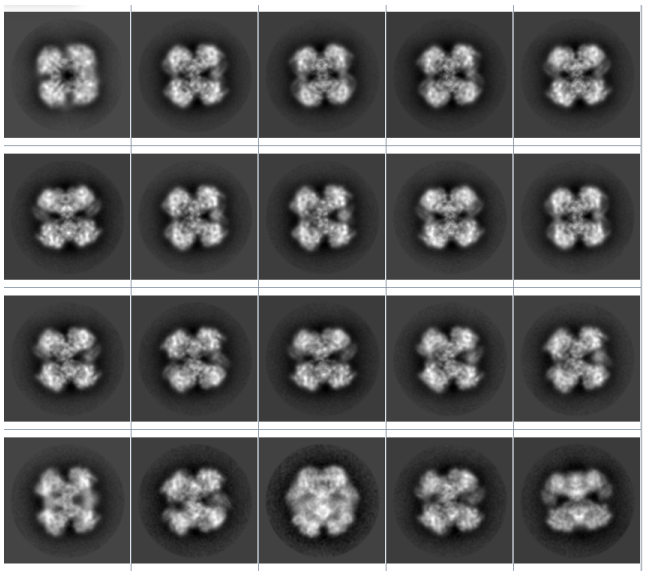

Figure 17: Set of classes obtained with *CL2D* [158], each image is the class representative or class average of a set of particle that define each class.

with a small set of images (class averages) instead of a large number of particles. Moreover, class averages will present higher SNR than isolated particles because the averaging operation reinforces the particle structure increasing its signal and smoothing the noise by reducing its means and increasing the variance. The result is a detailed image taken as the class representative. Commonly to speed up and help in the search of classes particles are *downsampled* by scaling the box size even though their SNR is increased. The downsampling process looses the high frequency components of the particle, but, the goal is to prepare a set of classes for

creating a coarse 3D estimation of the macromolecule structure, that will present very low resolution, so that the loose of high resolution components lacks of importance for the initival volume problem, in contrast with the later refinement step. The use of classes is also useful to undertake a second pruning step. Note that if a class is composed by a set of particles with similar characteristic, many empty particles, artifacts or others will cluster in the same class, and therefore they can be removed from the data set. In addition, other issues as the existence of preferred directions, bad angular cover of the projection sphere or the heterogeneity in the sample can be identified in the classification step. It should be remarked that in addition to the classification task an *alignment* task is also carried out. Note that many particles might present the same orientation (longitude and latitude) but be rotated in the plane. Moreover, they can be slightly shifted because of the picking step. Both effect, rotation in plane and shift are solved with the alignment.
The effect known as *Einstein from noise* [72, 75, 153, 208] can be identified in 2D averages. The high noise present in the micrographs presents might introduce a bias if the picking step is not properly performed. Due to most of the picking methods works with similarity measurement with a template (as cross-correlation), some features of the noise might randomly correlate with the template, and therefore a particle defined just by noise can be picked up. This particle will exhibit features of the template. Thus, when several images are averaged the aligned noise features are reinforced and a bias is introduced. The result will be an wrong reconstructed structure in the subsequent step, and the need of validation tools emerges. In Fig. 18 this effect is shown.
There are many classification algorithms to create sets of classes. The Multivariate statistical analysis (Multivariate Statistical Analysis (MSA)) tries to perform data dimensionality reduction. To do that, it is considered that a particle with $n \times m$ px is represented as a vector $\mathbf{u}$ with mn components in a nm-dimensional space $\mathbb{R}^{nm}$. However, the information contained in many components lacks of sense to the classification task and complicate it. Hence, a dimensionality reduction can be carried out, the idea is to find a vector, $\mathbf{v}$, in a subspace, $\mathbb{R}^p$ of $\mathbb{R}^{nm}$, so that $\mathbf{v}$ will be the best approximation of $\mathbf{u}$ in $\mathbb{R}^p$. In common words, if we have a 3D structure (image), we look for a plane object (dimensionality reduced) such as the plane object is the best approximation to the 3D structure. The shadow of an object is the best approximation, and it is less complex in terms of dimensionality. Thus, the problem is reduced to find similarities between "shadows´´. A deep explanation about these techniques can be found in [69]. Another approach is the Multi-reference classification (Multi-reference classification (MRC)), where several class representative are defined (for instance by means of random subsets of the set of particles), and particles are classified by comparison with those representatives using a similarity metric. The most representative methods of this group are *ML2D* - RELION [148, 150] and *CL2D* - Xmipp [158, 162]. Other methods based on k-means exploits the advantages of stochastic hill climbing to speed up computational time. However, another solution also exists, for instance, *CryoSparc* uses branch and bound solution to avoid falling in local minima and it reduces computational times [126].

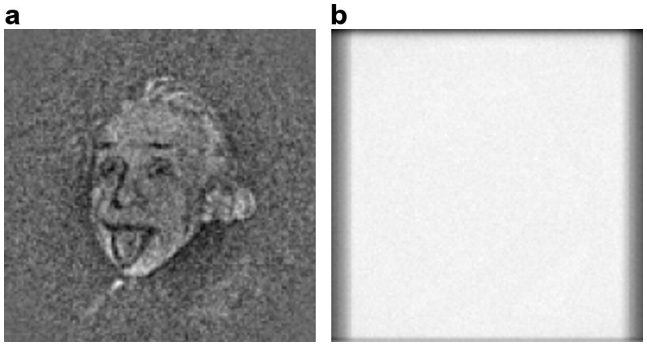

Figure 18: 1000 particles were picked from a white noise image by aligning an image of Einstein chosen as reference. (a) when the 1000 picked particles are averaged, the image of Einstein appears, (b) white noise image used for picking. Images were taken from [153].

3.6 INITIAL VOLUME

As it was mentioned in the last section, the large number of particles, the absence of angular assignment and their low SNR are the main problems to estimate the structure of the macromolecule. However, thanks to the classes, the complexity of the problem is reduced: just a few classes with enough SNR are used to determine a first and coarse estimation of the 3D structure, i.e. a low resolution structure. In the next steps this initial map will be refined up to achieve a detailed structure at high resolution. However, it should be highlighted that, this first estimation is a key-point for the subsequent refinement, because wrong or poor initial maps might introduce bias in the final map, or in the best cases it makes slower to achieve the convergence.

The map reconstruction from projections, or in this case classes, makes use of the Central Slice Theorem, see Section 3.1.1. There are many initial volume methods around this theorem. One framework is to considers pairs of different projections of the macromolecule, their Fourier transforms will intersect in a common line, therefore the problem of determining the angular assignment of the classes is reduced to find the common lines. Next methods will follow this idea [56, 125, 154, 183, 196]. Unfortunately, these kind of methods might detect false common lines. Alternatively, a statistical approach can be used by optimizing the alignment variables i.e. shift and Euler angles (particle orientation). *Significant* [159] and *PRIME*[35] belong to this approach. Normally, they start with the called *ball*, which is a reconstruction considering random orientation of the classes, then an iterative procedure of variable optimization make the ball evolve towards a more realistic map compatible with the classes. In Fig. 19 the evolution of a ball using Significant is shown. The cons of statistical methods are the first estimation of the initial volume (ball) for the iterative process, making it susceptible to get stuck in local minima. Another original way of estimating an initial volume is the *RANSAC* approach [186] which considers random reconstructions and tries to maximize the number of inliers/classes compatible with the reconstruction, then the best reconstructions are refined with a projection matching approach.

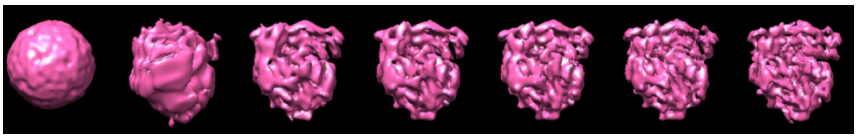

Figure 19: Evolution of the reconstruction obtained by an iterative process of reconstruction (in this case Significant [159]) to get an initial volume.

Recently, a novel approach to the initial problem was introduced [169] by considering as input the whole set of particle and several initial volume solutions obtained from different algorithms. Note that the set of initial volume methods might cast slightly or in some cases totally different solutions to the initial volume problem. All solutions are valid as initial map, but some features of the real map might be contained in some of them and not in the other. By using particle swarm optimization method, this method tries to merge and evolve those initial volume to a better one.

It is not common but in some projects the estimation of an initial volume fails with the current methods. In those cases, an alternative is to turn to negative stain samples and use tilt pair measurements with the aim of performing a reconstruction via Random Conical Tilt (RCT) [127, 160] or Orthogonal Tilt Reconstruction (OTR) [96]. The use of tilt pairs consists in performing two measurements per sample: the first with the sample untilted and the second tilting it. This experimental mechanism introduces geometrical information about the angular orientation of particles, which is squeezed to determine an initial volume. As it was mentioned (see sample preparation section 2.5), the use of negative stain dries the sample producing elongated maps at medium-low-resolution. Currently, the use of this procedure is marginal due to the initial volume problem has been overcome by the methods cited in this section.

3.7 3D CLASSIFICATION

The initial volume provides a first approach to the real structure of the macromolecule, which will be obtained in a later refinement step. To undertake a reconstruction and the SPA workflow two conditions were established, see section 3.1 In particular, it was assumed that all particles are projections under different point of view of the same macromolecular complex. This condition can be broken because of the heterogeneity problem, which is considered as one of the main open problems in cryoEM [114]. According to the heterogeneity origin, it can be classified as *conformational heterogeneity*, which considers that the macromolecular complexes are not rigid and present certain degree of flexibility; or it can be *structural heterogeneity* because, despite the purification efforts, some proteins present slight but non negligible differences in their structure. Note that radiation damage can also be responsible of structural heterogeneity. In both cases, the identical copies assumption is broken, and therefore, if a refinement is going to be carried out, it is necessary to classify the set particles in groups such as all particle that belong to the same group are projections of the same macromolecule, this step is called *3D classification*.

Although the heterogeneity problem is still open, there are many methods that solve the problem under certain conditions. From a physical point of view, the information about vibrational modes of the macromolecule and its dynamic can be used to define a phase space in which all the possible conformations are contained [28, 60, 93]. Alternatively, the use of statistical approaches can be undertaken, by means of Bayesian marginalization algorithms [126], Principal Components Analysis (PCA) (Principal components analysis) [62] or analysis of covariance matrix [164]. However, the most spread solution is the method of *ML3D* a maximum likelihood approach integrated in RELION[147, 156, 206].

3.8 REFINEMENT

Once it is possible to guarantee that the set of particles is composed of a homogeneous population of projections of the same macromolecular complex, the elucidation of a high resolution map can be undertaken. The goal of the *refinement* step consist on using the initial volume or the reconstructions from 3D classes as initial maps to by means of an iterative search converge to a detailed structure of the macromolecule compatible with the particles. Rigorously the problem can be defined as it follows: Given an homogeneous set of N particles (images, I_i with $i = 1, ..., N$) that can be considered as projections of the same structure $\mathbf{V}$ from different points of view, the goal is to determine the Euler angles and shifts, such as the distance between the image and the projection of the structure, $P_\theta V$, which under the direction θ, is minima. P_θ is the projector along the direction θ. Note that the problem is equivalent to solve the linear equation $A\mathbf{x} = \mathbf{b}$, with A the projector, P_θ , $\mathbf{b}$ the particle images, and $\mathbf{x}$ the structure to be determined, i.e. V. For a deeper explanation on reconstruction algorithms see [165]. To solve the problem several approaches can be considered, like maximum likelihood [126, 147, 150, 206], maximum *a posteriori* [206] (this is the RELION approach) or the traditional projection matching [122, 123, 167] . The maximum likelihood problem can be reduced to the optimization problem

$$(\theta_i^*, V^*) = \mathrm{argmin}_{\theta_i, V} \|I_i - P_\theta V\|_W^2, \tag{41}$$

where W is a given weight. This method represented a revolution in the field considering that a single experimental image/particle can be understood as the projection of the structure from many directions each of them with different probability. Thus, the goal is to, by means of an iterative process, collapse the probabilities up to reduce the possible directions. In contrast, maximum *a posteriori* adds information about the structure by introducing an extra term $f(V)$, that penalizes undesired reconstructions

$$(\theta_i^*, V^*) = \mathrm{argmin}_{\theta_i, V} \|I_i - P_\theta V\|_W^2 + f(V), \tag{42}$$

Finally, the projection matching approach splits the maximum likelihood in two simpler problems

$$\theta_i^* = \mathrm{argmin}_{\theta_i} \|I_i - P_\theta V_{k-1}\|_W^2, \tag{43}$$

$$V^* = \mathrm{argmin}_{\theta_i} \|I_i - P_\theta V\|_W^2. \tag{44}$$

The first problem considers an initial map, V_{k-1} and aims to assign angles to the particles, and then in the second problem the assigned angles are used to reconstruct a new structure and measure its compatibility with the experimental images. In Fig. 20 the result of the refinement can be observed.

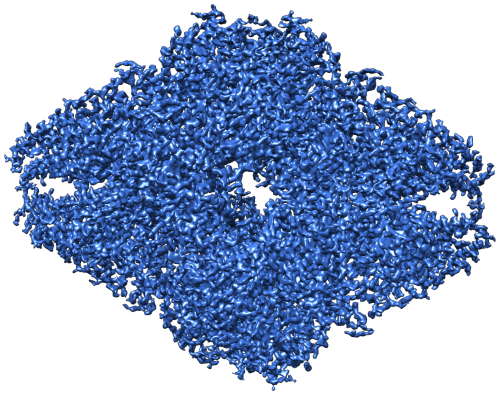

Figure 20: The initial volume or a map from the 3D classes is refined to achieve a high resolution structure. The image shows the refined beta-galactosidase structure.

3.9 VALIDATION AND ANALYSIS

Many steps in the SPA workflow require user decisions, being some of them based on user-subjectivity criteria which might introduced a bias, or bad quality maps. This,along with the dealing with very noisy images (low SNR) makes that, despite the user takes good decisions, methods are susceptible to elucidate wrong structures or fall in local minima. Hence, quantitative tools to validate the reconstructed map and assess its quality are required.

To guarantee the validity of a structure information several approaches can be considered. First, the use of external reconstruction tools from *X-ray crystallography*, NMR, or any other technique that allows to recover the 3D structure of the macromolecule under study. This sort of validation is a kind of blind test, by using two independent tools, however, budgets or resources does not usually allow the use of other techniques. An alternative is to check if similar structures were solved looking up in databases as Protein Data Bank (PDB) [13] or EMDB [95].

Leaving external techniques out as validation tools, the alternative is to seek the compatibility of the elucidated map with the set of particles that comes from it. The first method introduced with this purpose considers the use of tilt pairs [74, 135, 197]. Note that the geometrical constrain introduced by the tilt angle must be kept when the particle tilt pairs are aligned with the reconstructed map. Statistical analysis of all particle tilt pairs will determine the validity of the map [139]. However, the use of tilt pairs is currently in disuse, so that, other validation methods were developed. In particular, many reconstructed structured suffer overfitting. Despite the attempts of avoiding it by splitting the set of particles in two halves, the named gold standard, it might appear in many reconstructions. Therefore, a tool for overfitting detection is needed. To do that, a subset of particles can be replaced

by random noisy particles and then the resolution of the reconstructed map with the original set of particles and the replaced set is analyzed [75]. A quality measurement called FSC that will be introduced later should notice this replacement in presence of overfitting. The last group of methods performs alignability validation [187, 188], by considering the projection assumption each particle should be a projection of the reconstructed map. Therefore, the angular assignment should cast the most probable directions of each particle close to each other, in contrast to a noisy particle.

Once the map has been validated, it should be reported a quality parameter that allows to measure the reliability in spatial term of the reconstructed map. Thus, the resolution concept arises as the smallest detail presented in the map. Despite there are no consensus about an universal definition of resolution, this one is the most spread. Note that *resolution* is the main topic of this thesis, being treated in depth in Chapter 4.

The main problem of cryo-EM images is the high noise present in the data and in the reconstructed volume. Thus, resolution measurements are focused on distinguishing signal and noise at different frequencies. To undertake that task there are different metrics like: the *Fourier Shell Correlation* (FSC) [143], *Differential Phase Residual* (DPR) [45, 123], or the *Spectral Signal to Noise Ratio* (Spectral Signal to Noise Ratio (SSNR)) [119, 181]. For a deep review of these metrics see [166]. Currently, *FSC* can be considered as the standard one in the field. It is defined as the normalized cross correlation between two maps at different frequencies. In this sense, the FSC measures self-consistency instead of being quality measurement. The frequency dependence determines the resolution value as the frequency at which the FSC curve crosses a certain threshold that ought to be properly chosen. For a deep discussion about threshold criteria see [135, 166, 209]. The FSC is defined in Fourier Space, so that concept of locality is lost, and as a consequence, the map background might induces resolution errors. To avoid that the map ought to be masked, enhancing the resolution measurement [8].

The FSC and other resolution measurements are global, reducing to only one number the whole macromolecule information. However, the degree of quality of the reconstructed map can be spatially variant. Thus, methods for computing the local resolution were developed, each of them under its own resolution definition. The origin of this spatially variation of the resolution obeys to many fact as they are: radiation damage, sample heterogeneity, the existence of preferred directions, or a non-uniform coverage of the projection sphere. The first method to undertake a local resolution estimation was *Blocres* [22], which uses two half maps to determine the local FSC resolution values by means of a local moving window. However, the most extended method up to nowadays is *ResMap* [90], that, using a steerable basis, calculates the best detection of local sinusoidal above the noise level. Thus, resolution is defined as the measurement of sinusoidal features above the noise level. Finally, the last resolution method called *MonoRes*, and it is one of the results of this thesis [189]. It computes the local amplitude/energy of the density map at many frequencies determining the highest frequency at which the energy can be measured above the energy of noise.

4

ON THE GLOBAL RESOLUTION CONCEPT AND ITS MEASUREMENT

The main objective of the reconstruction workflow in electron microscopy is to get the best possible density map of the macromolecular complex. It requires a metric or a set of them to quantify what is understood by best possible map. This metric can also be used for guiding the reconstruction process as well as for measuring the quality. Resolution is one of them, unfortunately, this is a controversial concept and the field of EM lacks of consensus about its definition [120]. Despite that, there are different metrics focused on answering how much quality of degree of detail a map has, they share a common point, resolution clearly is a spatial concept.
From the field of Optics/Physics, resolution has a single definition as the capability of an imaging system for distinguishing two different objects/points. This definition makes resolution a property of the imaging system. Note that every imaging system is limited by diffraction and, therefore, resolution will be too. Thus considering circular apertures, resolution be measured by the Airy disk and the Rayleigh criterion [17]. This criterion can be considered as the standard resolution measurement in optics. However, there also exist other criteria, as it is the case of Johnson criterion [83], among others.
In EM the most widespread definition of resolution is the size of the smallest reliable detail in the map. The misunderstanding with this definition is to determine a criterion to assess the smallest reliable detail. Thus many approaches to this problem have been proposed, most important are: Q-factor [64], Fourier Shell Correlation (FSC) [66, 144, 145], the Spectral Signal-to-Noise Ratio (SSNR) [119, 181] or the Differential Phase Residual (DPR) [45, 123]. A review of all of them and their relation each other can be found in [166]. In this chapter a brief review of the resolution concept and its standard measurement is cryo-EM.

4.1 RESOLUTION CONCEPT FROM PHYSICS

Resolution describes the degree of detail that an optical system is able to discriminate, the higher resolution the higher quality and details can be seen in the image. Resolution is then characterized by performance of the imaging system and as a consequence, the natural approach to undertake the study of resolution concept is analyzing the imaging optical systems.
An image of the object is acquired by an optical system, as it is the case of the electron microscope. Maxwell conditions, given in Chapter 2, establish the requirements for getting a perfect optical system, i.e. with infinity resolution. Unfortunately, these systems do not exist in the real world, and resolution is limited due to the existence of aberrations (imperfections of the optical system) or by diffraction. Assuming that the instrument is aberration-free, then, resolution is limited by the wavefunction of electrons (the geometry defines the shape of the wavefunc-

tion), i.e. by diffraction, under this scenario resolution will be analyzed.
Consider a point source and an optical system with circular apertures that produces their image on a screen. Several configurations can be thought, the point-source can be locate in a far point or near the to lenses. In Chapter 2 was demonstrated that when the object is in a far point, the image is the Airy disc. Now it will shown that the Airy disc is obtained with independence of the optical system. In Fig. 21 an scheme of the two possible systems is shown. In (a), the source is enough far from the lenses to be considered at infinity distance, that the image is obtained at the focal plane, F'; in (b) the object, O is relatively close to the lenses. However the imaging lenses of the last scenario can be decomposed as two lenses, as it is shown in (c), where the first image the object at infinity distance and the second one is responsible of map the Airy disc is its focal plane. Hence, the Airy disc, i.e the PSF is always obtained with independence of the system. Remind that the Airy disc was defined as $J_1(x)/x$, and the size of the Airy disc were determined by the first zero of the Bessel function $J_1(x)$ (see chapter 2) casting

$$\sin\theta = 1.22\frac{\lambda}{2a}. \tag{45}$$

Coming back to the resolution problem, consider two point sources separated each other a distance $2a$ and located in the same plane orthogonal to the optical axis of imaging instrument. Their image will be two Airy discs. Because the image of a point is not a point, the Airy discs can overlap and perhaps the two point sources might not be distinguished. Lord Rayleigh established a criterion to solve this issue, *Two point sources imaged by an optical system with circular apertures can be distinguished if the distance between the center of their absolutes intensity peaks is smaller than the radius of the Airy disc [17].* In Fig. 21 the criterion is graphically explained with the overlapping of two Airy discs.

The criterion defines a resolution limit, but this limit is a property of the instrument as well as an angular measurement, note that Eq. (45) relates the angular accuracy of the instrument, which is the term $\sin\theta$, for imaging two sources separated a distance $2a$. It can also be transformed into linear distance by considering paraxial approximation and multiplying Eq. (45) by the distance sources-lenses s,

$$s\sin\theta \approx s\theta = 0.61\frac{s}{a}\lambda, \tag{46}$$

Note that the quotient $\frac{a}{s} \approx \sin\sigma$ in paraxial approximation, where sigma is the semi-acceptance angle of the optical system, i.e. $NA = \sin\sigma$ is the numerical aperture of the instrument. What has been derived is the named Abbe criterion, that related the linear resolution limit, r of the and instrument with its numerical aperture NA,

$$r = 0.61\frac{\lambda}{NA}. \tag{47}$$

According to Eq. (47) there are two ways of increasing the resolution limit, by decreasing the wavelength of the illumination source, or reducing the numerical aperture of the imaging system, but again, resolution is a property of the instrument

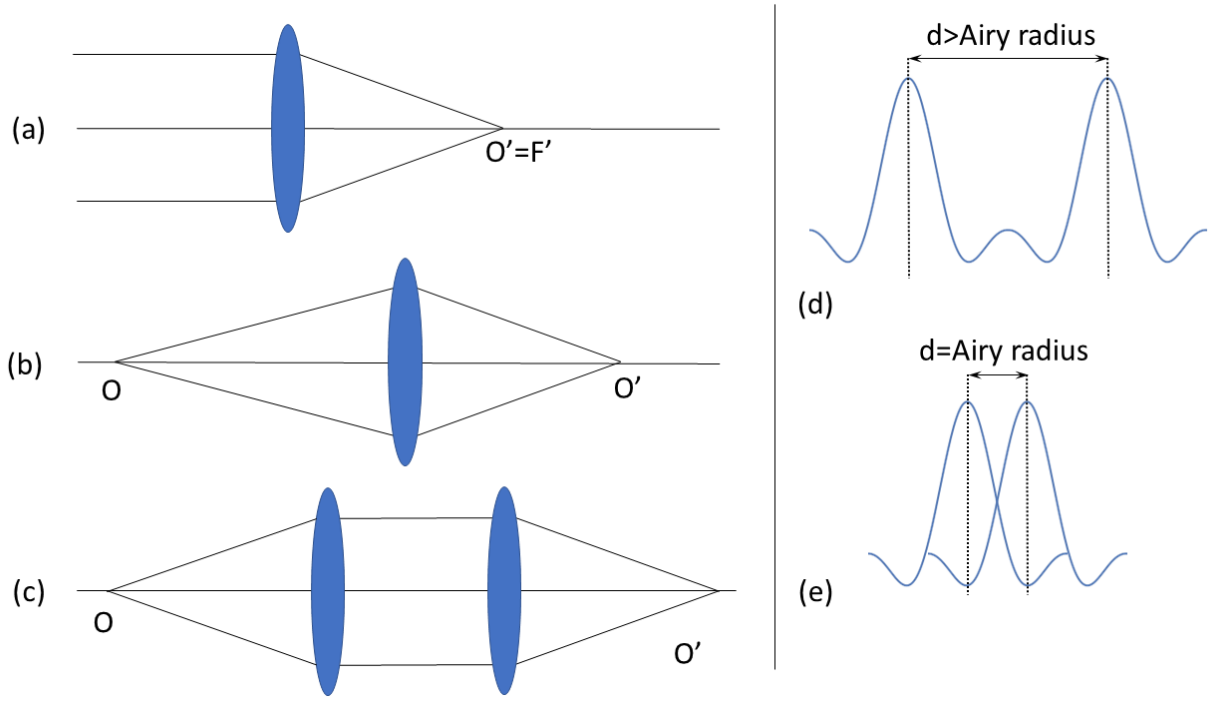

Figure 21: (left) Equivalence between several optical system with the object located at different distances from the optical system (a) at the infinity, (b) close to the lenses. Note that (b) and (c) are equivalent systems. (right) Rayleigh criterion explained, (d) two point sources can be distinguished because the distance between their Airy is greater than the Airy radius, (e) the distance is exactly the Airy radius, i.e resolution limit.

(dependence on NA). Note, that in the back focal plane, the Fourier transform of the object is obtained, thus, by reducing the numerical aperture, the high frequency information in the focal plane is removed, and therefore the resolution gets lower. Let us finally summarize the Rayleigh criterion in a different manner as follows: Consider the image of a point-object by a perfect optical system that is limited by diffraction is a certain probability distribution named PSF; then two different point-objects separated a distance smaller than the PSF size will not be distinguished. This criterion only considers the measured intensity of the PSF, and therefore it is incomplete because it neglects the phase information. If the phase is added assuming that the wavefunction of the two emitter sources presents and entanglement (i.e. they are not fully independent), then the resolution limited given by Rayleigh can be surpassed [117, 118]. However, this new step forward is currently under investigation being a hot topic in quantum optics, in contrast electron microscopy is still far of the Rayleigh criterion.

Alternatively to the analysis of the instrument, it also possible to analyze only the image to report a quality value. This is the idea behind the Johnson criterion [83]. Despite it was born with a military purpose and is not too much widespread it is very illustrative of the resolution concept. Johnson's approach is focused on the recognition of targets in acquired images. To do that, it makes use of a Ronchi test, i.e. a fringe pattern composed by white and black lines (a white line followed

by a black one is named pair of lines). Thus, the Johnson criterion determines the threshold that gives the minimum number of pairs of lines to first detect, second identify and third recognize a given target. In Fig. 22 the Johnson criterion is graphically shown. Thus, in a very simple manner and considering a given distance object-instrument, the resolution that allows to detect/identify is obtained by overlapping a fringe pattern (on the display/screen) along the smallest dimension of the image, and then counting pair lines. Note that the Ronchi test introduces the notion of frequency, and the contrast of the pattern is of course related with the SNR which will be seen it is close related with the standard resolution measurement in electron microscopy.

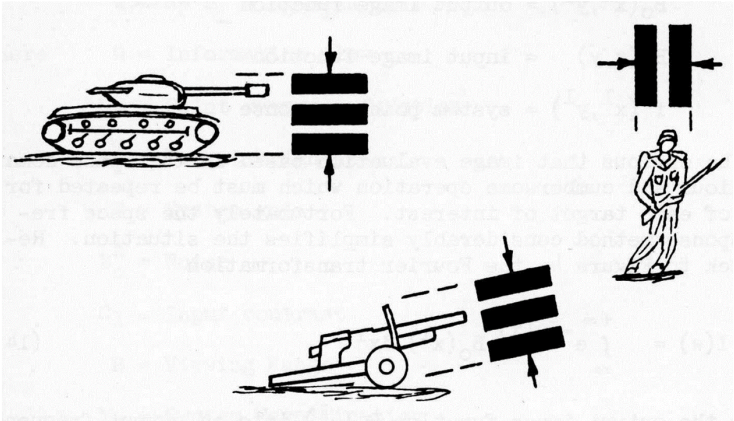

Figure 22: Scheme of the Johnson criteria, a fringe pattern is overlapped with the image in the display, the resolution that allows to detect/identify/recognize is obtained by counting the number of pair lines along the smallest dimension. (Image taken from [83])

4.2 RESOLUTION IN ELECTRON MICROSCOPY

As it was shown in the previous chapter, in electron microscopy applied to structural biology an structure is reconstructed from a set of images. There are many kinds of electron microscopes that can used to acquire images. Note that from an optical point of view the resolution is a property of the instrument and therefore because of this variety the application of optical resolution criteria to EM-maps is unfeasible. Moreover, at the end of the the reconstruction workflow a 3D structure instead of an image is obtained, meanwhile in optics the two points to be distinguished lay on the same plane. However, by establishing an analogy with optics but leaving the instrument out, the resolution of a complex can be defined as the smallest detail/information that can be distinguished in the structure. Thus, it is necessary to provide a metric, independent of the instrument, that reports this smallest dimension. As well as it occurred with optics criterion, the problem is the arbitrariness of this definition. In particular, how it is established a size threshold that determines if a detail of the complex is reliable or not. To cast more light on this issue, consider instead of a complex structure, a single atom, and we wonder about the resolution of a single atom. This is a fuzzy question to ask, physics

lacks of a unique answer; electrons present an uncertain in their position being the orbitals the probability regions. If the bond of atoms is considered hybrid orbital appear, and again probability regions in which electron can be found in are defined, but their limit is undefined. This fact is even worse when the effect of temperature is included, i.e. atomic vibrations of atoms also make fuzzy the atoms limits. Summarizing, from physics there is no criteria to define the limits of atoms or bonds or complexes, and therefore, coming back to electron microscopy, to find a unique resolution definition will be impossible.
In cryoem several resolution metrics were developed but each of them involves different criteria, in this chapter only the most important will be explained, the FSC. For a deep review of resolution measures see [166].

4.3 FOURIER SHELL CORRELATION - FSC

The macromolecule is modelled by a function, $x(\mathbf{r})$, depends on the position $\mathbf{r}$. The macromolecule is embedded in ice, and its reconstruction is susceptible to present certain kind of noise modelled by $n(\mathbf{r})$. Thus, the reconstructed structure is modeled as $f(\mathbf{r}) = x + n$ or equivalently in Fourier Space,

$$F(\boldsymbol{\omega}) = X(\boldsymbol{\omega}) + N(\boldsymbol{\omega}), \tag{48}$$

where ω represents the frequency in Fourier space and the capital letters F, X and N are the Fourier transforms of their respective functions in small letter, f, x and n.
The FSC is defined as the normalized cross-correlation at difference frequencies between two maps, $f_1(\mathbf{r}), f_2(\mathbf{r})$. This metric is perform in Fourier Space as

$$FSC(r, \Delta r) = \frac{\sum_{r\in} F_1(\mathbf{r}) F_2^*(\mathbf{r})}{\sqrt{\left(\sum_{r\in\Delta r} |F_1(\mathbf{r})|^2\right)\left(\sum_{r\in\Delta r} |F_2(\mathbf{r})|^2\right)}}, \tag{49}$$

where F_1 and F_2 are the Fourier transform of the respective maps f_1 and f_2, and Δr is the thickness of a Fourier shell. Note that FSC presents a frequency dependence, the resolution value is determined as the frequency at which the FSC crosses a given threshold that ought to be chosen properly. The most used, as they are 0.5, 0.143 and 1/3 will be briefly discussed taking advantage of mathematical shortcuts. However, for rigorous mathematical proofs see [166, 209].

4.3.1 *On the FSC measurements*

Currently, the *FSC* has become the standard resolution measurement in the field. For this reason its meaning and properties must be analyzed.

1. *Self-consistency measurement:* The FSC is calculated as a cross-correlation, that measures the self-consistency at different frequencies between to functions f_1 and f_2. Thus, the FSC is a self-consistency measurement more than a quality metric. It does not take into account systematic errors in the reconstruction process, on contrary, it rewards them.

2. *Global measurement:* Resolution can locally vary, as a consequence of the existence of heterogeneity, flexibility, angular assignment problems, or radiation damage. The FSC does not consider these effects, it reports one number as resolution of the whole map.

3. *Threshold dependence:* There exist different criteria to establish a resolution limit that defines the global resolution of the structure. Below, several thresholds are discussed.

4. *Mask dependence:* As a cross-correlation measurement it depends on the shape of the functions involved in the cross correlation. When they are masked to remove the background (noise) the cross-correlation is affected and therefore the FSC. Or alternatively, the FSC is defined in Fourier Space, as a consequence, concept of locality is lost and the map background might induce resolution errors. To avoid that the map ought to be masked [8].

5. *Invariant under global linear transformations:* Linear transformations as it is the case of B-factor correction (see, Chapter 7) leaves invariant the FSC. The FSC enhancement that the a sharpening post-processing using a global B-factor it is due to a tight mask.

6. *Measurement of SNR:* Considering the FSC as the normalized cross-correlation between two maps, f_1 and f_2, then it can be proved the following relation [14, 166]

$$\mathrm{SNR} = \frac{\mathrm{NCC}}{1-\mathrm{NCC}} = \frac{\mathrm{FSC}}{1-\mathrm{FSC}}. \tag{50}$$

   In this sense, the FSC is an alternative but equivalent measurement of the SNR.

4.3.2 *FSC-threshold of 0.5*

This threshold appears when it is assumed that there is the same amount of noise and signal, i.e. $\mathrm{SNR} = 1$. Under this assumption by substituting in, Eq. (50) a value $\mathrm{FSC} = 0.5$ is obtained.
The named *Gold Standard* that was introduced in Chapter 3 suggests a method for map validation and prevents the existence of overfitting. It considers two independent reconstructions by splitting the original set of data (set of particles in the case of SPA) in two halves [57, 148]. Of course, both reconstructions should be similar, and ideally the same. Thus, to quantify the degree of self-consistency between them the FSC curve is calculated, and the proper threshold for this scenario is justified to be set as 0.5, i.e. $\mathrm{SNR} = 1$. However, this threshold underestimates the resolution because it only consider the half number of particle to perform the reconstructions.

4.3.3 *FSC-threshold of 1/3*

This threshold considers the whole set of data to properly estimate the resolution. To double the data implies to reduce the amount of noise in a factor two, rewritting Eq. (50) and substituting $SNR \rightarrow SNR/2$,

$$FSC = \frac{SNR}{1+SNR} \qquad \rightarrow \qquad FSC = \frac{SNR}{2+SNR}, \tag{51}$$

And considering again a value of $SNR = 1$ the resolution for the whole data set will correspond with and FSC threshold of $FSC = 1/3$. Unfortunately, despite being a very straight forward derivation this threshold is not commonly used.

4.3.4 *FSC-threshold of 0.143*

In many cases, it is known the atomic model of the structure under study. Hence, it is possible to calculate the FSC between the reconstructed map considering the whole set of particles and the reference, i.e. the atomic model converted into density map. Note that reconstructed map will sum of a perfect map s and noise n, therefore, in Fourier space it can be written as $F_1 = S + N$ where S and N are the Fourier transforms of the perfect map and noise, respectively. In the case of the reference, it is noise-free and thereby $F_2 = S$. Then the FSC with the reference will be

$$FSC_{ref} = \frac{\sum_{r\in\Delta r}(S + N/\sqrt{2})S^*}{\sqrt{\left(\sum_{r\in\Delta r}\left|S + N/\sqrt{2}\right|^2\right)\left(\sum_{r\in\Delta r}|S|^2\right)}} = \tag{52}$$

$$= \sqrt{\frac{\sum_{r\in\Delta r}|S|^2}{\sum_{r\in\Delta r}|S|^2 + |N/2|^2}} = \sqrt{\frac{2FSC}{1+FSC}}. \tag{53}$$

Note that if $FSC = 1/7$, then $FSC_{ref} = 0.5$. In other words, when the of the reconstruction considering the whole set of particles is $SNR = 1$, then, it correspond with and FSC value half-half $FSC = 1/7 = 0.143$.

Despite these three derivations of FSC threshold consider an exact result [135], a rigorous proof must take into account the distribution of noise and signal by calculating the expected value of the FSC as it is done in [166]. However, for sake of simplicity the original derivation is kept in this introduction.

Part III

DEVELOPMENTS AND RESULTS

5

LOCAL RESOLUTION

In electron microscopy, the resolution concept is frequently a global parameter that describes the degree of detail that can be detected in density maps. However, as G. Cardone pointed out in his article entitled *One number does not fit all* [22], resolution presents a spatial dependence. It means that the quality of the volume may vary usually from one region to another, thus the resolution concept ought to be extended as a local measurement. The explanation of these resolution differences lays on the structure shape, existence of heterogeneity and the reconstruction process among others.

In contrast to the overall resolution, where the FSC is the most spread method and currently could be considered the standard one, the measurement of local resolution maps is relatively recent. It was first introduced in 2013, with the article by G. Cardone et. al [22] when the *blocres* method was proposed. The idea is simple, it computes a local FSC by means of two half volumes and a moving window wherein the FSC is calculated. The main pitfall and inconvenience is to determine properly the window size. Moreover, all problems related with FSC are inherited, like the background sensitivity (mask dependency) or its invariance under isotropic transformations. The second method for estimating local resolution is *ResMap* [90], and it is currently the most widely used method. Briefly, it considers that one point can be measured at certain resolution, $1/\lambda$, by checking if a local sinusoid of wavelength λ is statistically detectable above the noise level. The main advantage is its simplicity, requiring only the user interaction for the prewhitening step. Moreover, it can work with a single volume or two half maps which increase the versatility.

Despite *blocres* and *ResMap* provide good results they present drawbacks that we wanted to solve: first, to achieve a fully automatic method that avoid the user intervention and second, to design an algorithm computational enough fast to deal with large maps in short computational times. Thus, in this chapter, two new methods addressed to estimate local resolution in density maps called *MonoRes* and *MonoTomo*, are explained. They share the mathematical algorithm under slight modifications that make it suitable to be applied in SPA or electron tomography reconstructions. To do that, *MonoRes* and *MonoTomo* make use of an extension of the concept of analytic signal, called monogenic signal. Then, making a frequency sweep and calculating the correspondent monogenic signal at each frequency, a decomposition in phase and amplitude is achieved. The resolution is assigned voxel by voxel by means of hypothesis tests. If the amplitude of the voxel is significantly higher than the amplitude of the noise, then that voxel can be measured at the filtering resolution. To our knowledge, *MonoTomo* is the first local resolution method in electron tomography up to nowadays, which may open new horizons and applications in the field.

5.1 ON LOCAL RESOLUTION AND ITS ORIGIN IN SPA

As was pointed out in Chapter 4, it is broadly accepted that the resolution represents the smallest detail that can be distinguished in an object seen by an optical system. In the case of electron microscopy resolution is a property of the reconstructed structure and defines the smallest feature that can be identified. However, when different regions of the same the map are studied, it is noticed that the smallest measurable feature size in each region may be different. This suggests that resolution is a local measurement. It is then useful to quantify these resolution gradients as a resolution map that provides the spatial degree of reliability of the reconstruction. To do that, the spatial variability of the SNR can be used, which allows to reinterpret resolution in a local manner as the highest local frequency above the noise level. It avoids the over-interpretation of the map. Regarding with the origin of local resolution, the spatial dependence seems to be very plausible.

Regarding with the origin of the physical and mathematical mechanisms responsible of different local resolution values. The reconstruction process in SPA ideally considers that all images are projections of the same complex (in the absence of flexibility) and that the angular projection space is well covered. Thus, a perfect reconstruction should present the same degree of detail for any direction and any location. However, reality is different. For example, particles might not be exactly the same since there can be some heterogeneity in the macromolecules being imaged. Radiation damage can also be responsible of this fact, or even the flexibility of the structure. Moreover, the angular orientation and in-plane alignment of the particles used for the reconstruction can be incorrectly identified or, instead, a pattern of uneven orientation distribution may be present in the data [157]. The result is an electron density map in which different regions may significantly differ in their quality. Most methods for determining the resolution of a map are global and do not take into account these local or orientation differences.

5.2 DEALING WITH LOCALITY. ANALYTIC AND MONOGENIC SIGNALS

Images or volumes are composed by signals defined by multiple frequencies. Image features are the direct consequence of this spatial-varying frequencies. The access to local properties, as a decomposition in terms of phase and amplitude, allows to characterize image anomalies. The mathematical background to measure local and directional resolution is based on that decomposition, in concrete, in the amplitude term. This task is carried out by means of the analytic signals which were introduced by the laureate Nobel prize Denis Gabor [47] and have and found many applications in a broad kind of fields [48, 124]. Its extension to images is called monogenic signal [39]. In the case of density maps, the concept needs to be redefined due the extra dimension-3D, thus, for this purpose it was extrapolated. In the appendix at the end of this chapter deeper mathematical details are shown.

5.2.1 *Analytic signal*

Analytic signals constitute a mathematical formalism that allows decomposing a function into an envelope and phase terms. Given a 1D signal, $s(t)$, its analytic signal, $s_a(t)$, is defined as the complex representation given by

$$s_a(t) = s(t) + is_H(t), \tag{54}$$

where i is the complex unity and the function $s_H(t)$ is the Hilbert transform of the signal $s(t)$, defined as the convolution between the original signal and the kernel $h(t) = 1/(\pi t)$ as follows

$$s_H(t) = (s * h)(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau. \tag{55}$$

This expression gets simpler defining the Hilbert transform in Fourier space

$$\hat{s}_H(\omega) = -i\text{sign}(\omega)\hat{s}(\omega) = -i\text{sign}(\omega)\hat{s}(\omega), \tag{56}$$

being ω the frequency variable, $\hat{s}$ the Fourier transform of the signal $s(t)$ and $\text{sign}(\omega)$, the sign of the frequency, ω. This means, that the Hilbert transform applied to the Fourier transform, performs a phase shift of the original signal of $+\pi/2$ to the negatives frequencies and a shift of $-\pi/2$ to the positive frequencies. A very simple example to illustrate the Hilbert transform is a sinusoidal function, $f(t) = \cos(\omega t)$; its Hilbert transform will be the function $H[f(t)] = \sin(\omega t)$, which exactly presents a shift of $\pm\pi/2$ for negative and positive frequencies respectively. So that, considering a function that can be expressed in a Fourier series, the Hilbert transform will be such that the sines are swapped for negative cosines and the cosines term are changed to sines.

The analytic, $s_a(t)$, can be then decomposed as the product of an instantaneous amplitude,

$$A(t) = \sqrt{s^2(t) + s_H^2(t)}, \tag{57}$$

and an exponential phase factor, $\phi(t) = \arg(s_a(t)) = \text{atan}(s_H(t)/s(t))$, such that

$$s_a(t) = A(t)e^{i\phi(t)}. \tag{58}$$

In Fig. 23 an example of the decomposition that allows to perform analytic signals is plotted. The original function was $s(t) = \cos(3t)e^{-(t-8)^2/10}$, thus, Eq. (56) that defines the Hilbert transform introduces a phase shift of $\pi/2$, which can be seen in the figure as a displacement of the maxima/minima. Finally, note that the amplitude calculated with Eq. (57) defines the envelope of $s(t)$ signal and the Hilbert Transform performs a phase shift from the original signal.

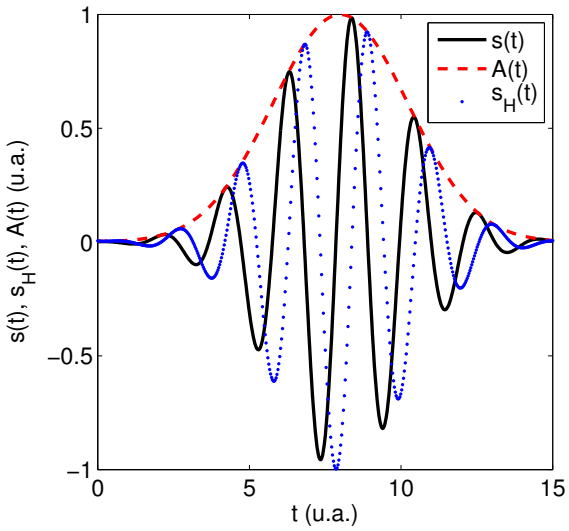

Figure 23: Signal decomposition. A signal $s(t)$ (continuous-black) its amplitude (red-dashed) and its Hilbert transform (blue-dotted).

5.2.2 *Monogenic Signals in 3D:*

It would useful to take advantage of the virtues of the analytic signals for extracting amplitude and phase but applied to 3D-signals. One of the extensions of this kind of signals to 2D, called monogenic signal, is the wide used Spiral Phase Transform [94, 184, 185]. There have been several extensions of the Hilbert transform to dimensions higher than 2, unfortunately, with absence of consensus by the mathematical community. Here, it is used an extension from 2D to 3D via Riesz transform and based on previous 2D-applications [39, 180]. The Riesz transform is defined in the 3D-Fourier domain a

$$\hat{\mathbf{s}}_R(\boldsymbol{\omega}) = -i\frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega}) = -i\left(\frac{\omega_x}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega}), \frac{\omega_y}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega}), \frac{\omega_z}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega})\right), \tag{59}$$

where $\hat{\mathbf{s}}$ denotes the Fourier transform of the signal $\mathbf{s}$, $\boldsymbol{\omega}$ is the 3D-vector frequency variable, and ω_j is its x, y, z component. Note the resemblance between the Fourier definition of the Hilbert transform Eq. (56) and the Fourier definition of the Riesz transform Eq. (59).

The monogenic signal is then defined by next quaternion

$$\mathbf{s}_{MG}(\mathbf{r}) = s(\mathbf{r}) + is_{R,x}(\mathbf{r}) + js_{R,y}(\mathbf{r}) + ks_{R,z}(\mathbf{r}), \tag{60}$$

where i, j, k are the quaternion complex unities. Equivalently, considering the Riesz vector, $\mathbf{s}_R = (s_{R,x}(\mathbf{r}), s_{R,y}(\mathbf{r}), s_{R,z}(\mathbf{r}))$ and $\boldsymbol{\sigma} = (i, j, k)$,

$$s_{MG}(\mathbf{r}) = s(\mathbf{r}) + \boldsymbol{\sigma} \cdot \mathbf{s}_R(\mathbf{r}). \tag{61}$$

The generalization to higher dimension makes use of a Clifford algebra defining N "imaginary units" (complex numbers and quaternions are particular cases of a Clifford algebra for $N = 1$ and $N = 3$, respectively). Note that $i^2 = j^2 = k^2 = -1$,

$ij = k$, $jk = i$, and $ki = j$ (complex numbers and quaternions are particular cases of a Clifford algebra for $N = 1$ and $N = 3$, respectively).

We now calculate the monogenic amplitude as

$$A(\mathbf{r}) = \sqrt{s^2(\mathbf{r}) + s^2_{R,x}(\mathbf{r}) + s^2_{R,y}(\mathbf{r}) + s^2_{R,z}(\mathbf{r})}. \tag{62}$$

As in the 1D case, this variable quantifies the local amplitude/envelope of the input signal (the macromolecule map, in our case) and it can be understood as local energy . For sake of clarifying the concept of monogenic signal a correspondence between 1D and 3D expressions between analytic a monogenic signals is established in Table. 1.

Concept	Unidimensional-Signal	Multidimensional signals
Transform	Hilbert	Riesz
	$\hat{s}_H(\omega) = -i\,sign(\omega)\hat{s}(\omega)$	$\hat{\mathbf{s}}_R(\boldsymbol{\omega}) = -i\frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega})$
Transformed Signal	Analytic Signal, $s_a(x)$	Monogenic Signal, $s_{MG}(\mathbf{r})$
	$s(x) + is_H(x)$	$s(\mathbf{r}) + is_{Rx}(\mathbf{r}) + js_{Ry}(\mathbf{r}) + ks_{Rz}(\mathbf{r})$
Amplitude	$A(x) = \sqrt{s^2(x) + s^2_H(x)}$	$A(\mathbf{r}) = \sqrt{s^2(\mathbf{r}) + \|\mathbf{s}_R(\mathbf{r})\|r}$

Table 1: Comparison between parameter derived from 1D and 3D signals, i.e. between analytic-monogenic signals.

5.3 MONORES ALGORITHM. MEASURING LOCAL RESOLUTION IN SPA

This section explains the details of the proposed algorithm, named *MonoRes*, for computing the local resolution of elucidated structures. The problem to solve is as it follows: Given a density map, the objective is to obtain another map in which the intensity of each voxel represents the local resolution of that voxel in the original density map. This map is named *local resolution map*.

The concept of resolution as a measurement of SNR is considered but in a local sense. The locality is achieved by using monogenic signals, in particular, the monogenic amplitude. However, the key point is the measurement of noise. A mask can be used to establish a frontier between structure and background (noise) and then to analyze them separately. However, it must be highlighted that this separation structure-noise by means of a mask is only an approximation, because the region with structure also contains noise. Note that, macromolecular complexes are embedded in ice, and the projection assumption implies the ice projection. Thus, two approaches can be carried out to characterize the noise, and they will determine the two possible inputs of *MonoRes*:

1. *Single map*: This kind of input assumes that the is no signal outside the mask, i.e. there is only noise outside the mask. The input map will be denoted as $V(\mathbf{r})$.

2. *Two half maps*: This approach allows to determine the noise that is mixed with the structure, i.e, the noise inside the mask. If $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ are both halves, then, two maps are computed, the mean volume,

$$V(\mathbf{r}) = \frac{1}{2}\left(V_1(\mathbf{r}) + V_1(\mathbf{r})\right), \tag{63}$$

and a noisy map, $N(\mathbf{r})$, that can be obtained the difference,

$$N(\mathbf{r}) = \frac{1}{2}\left(V_2(\mathbf{r}) - V_1(\mathbf{r})\right), \tag{64}$$

where, the factor $1/2$ is used to set the variance of the difference map and make it comparable with the mean volume, $V(\mathbf{r})$, as a consequence of the addition of two random variables as it is the case of $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$.

Note that in both cases, the density map to analyze its local resolution with *MonoRes* is denoted by $V(\mathbf{r})$, which in the two half maps will be the mean of them.
Regarding to the details of *MonoRes* algorithm, it is based on two pillars: 1) high-pass filter the volume and compute its local monogenic amplitude, and 2) noise characterization and hypothesis tests at that frequency to determine if the local energy of a voxel is significantly greater than the energy of noise at the filtering frequency. Although, these two pillars coarsely summarize the *MonoRes* algorithm, it is obviously more complex. Thus, a detailed enumeration with all specific steps is given below, moreover, in Fig. 24 a scheme are shown.

1. *High pass filter*: The original map (or the mean map using two halves), $V(\mathbf{r})$, is high pass filtered at a specific frequency, ω_0. The result is a filtered map, $V_{HP,\omega_0}(\mathbf{r})$. It is noteworthy that filtering may cause a ripple known as Gibbs effect. To alleviate it, the high pass filter implements a raised cosine.

2. *Monogenic amplitude*: Using the filtered map, $V_{HP,\omega_0}(\mathbf{r})$, the monogenic amplitude, A_{MG,ω_0}, at the filtering frequency is obtained as (see Appendix)

$$A_{MG} = \sqrt{V_{HP,\omega_0}(\mathbf{r})^2 + \sum_{j=x,y,z} V^{\alpha}_{HPR,\omega_0}(\mathbf{r})}, \tag{65}$$

where $V^{j}_{HPR,\omega_0}(\mathbf{r})$ are the Riesz components of the high pass filtered map

$$V^{j}_{R,\omega_0}(\mathbf{r}) = \mathcal{F}^{-1}\left[-\frac{\omega_j}{||\boldsymbol{\omega}||}\mathcal{F}[V_{HP,\omega_0}(\mathbf{r})]\right] \tag{66}$$

3. *Low pass filter*: The monogenic amplitude is susceptible to show some ripples of high frequency due to the non-linear character of Eq. (65). To avoid them, it is smoothed with a low pass filter at the same frequency ω_0.

4. *Noise and energy threshold*: In this step the noise distribution is characterized. When the input of *MonoRes* is a single volume, the noise distribution is obtained from all voxels outside the mask, as was mentioned above. In contrast, if the input are two half maps, the noise map $N(\mathbf{r}) = \frac{1}{2}\left(V_2(\mathbf{r}) - V_1(\mathbf{r})\right)$ is used to compute the noise distribution inside the mask. In both cases, noise

distribution is characterized by its respective mean and standard deviation, μ_n and σ_n. It allows to establish a threshold (an hypothesis test) to determine if statistically a given voxel can be distinguished from the noise (with significance $1-\alpha$)

$$\epsilon = \text{CDF}^{-1}(1-\alpha), \tag{67}$$

where CDF is the cumulative distribution function of monogenic amplitudes of noise. Note that, in the particular case in which the noise is gaussian, the threshold can be written as

$$\epsilon = \mu_n + \text{CDF}^{-1}_{N(0,1)}(1-\alpha)\sigma_n, \tag{68}$$

beging $\text{CDF}_{N(0,1)}$ the cumulative distribution function of a normal distribution. In *MonoRes* implementation the significance is set to 0.95 (default value).

5. *Determining resolution*: For each voxel of the structure (inside the masked volume), it is checked if the local energy is greater than the calculated threshold, ϵ. If it is the local structure can be measured at the filtering frequency, ω_0. This set of steps is repeated from low to high frequencies, i.e. it is repeated at a new frequency $\omega_1 > \omega_0$ and the resolution of each voxel will be the last measured resolution with local energy greater than the calculated threshold. To avoid the existence of false negatives in the hypothesis test, the resolution is assigned when the hypothesis test fails two consecutive frequencies.

Before showing *MonoRes* results and with the aim of really understanding the method, an unidimensional example is shown. Thus, in Fig. 25 (left-top) a unidimensional signal is composed by 3 frequencies (resolutions) of 1/20, 1/10, and 1/5 a.u (red-line). The signal is perfectly defined in a compact region, that in the case of 3D-maps is defined by the mask. Finally, gaussian noise was added with standard deviation of 0.2 a.u. (blue line). The monogenic amplitude (amplitude of the analytic signal) was calculated, it is represented in Fig. 25 (right-top). The blue, red and green lines are the amplitude of the noise-free signal, noisy signal, and only noise respectively. Note how the amplitude has non-zero value only in the finite region in which the signal is located. When the original signal is high pass filtered (second, third and fourth rows in Fig. 25) the monogenic amplitude is affected, but it still keeps energy in the region with higher frequencies than the frequency filter. Note that the moment for which the amplitude cannot be distinguished from noise, defines the resolution value (leaving out false positive that are not described in this simple explanatory example). Finally, in the last row, the amplitudes of the noise and signal cannot be statistically distinguished. This is the idea behind *MonoRes*, to perform a frequency sweep of high pass filters at different frequencies, determining if the local energy is statistically higher than noise.

5.4 LOCAL FILTERING

Once a structure is reconstructed, it is common to low pass filter the obtained map at its resolution value with the aim of removing noise and enhancing the

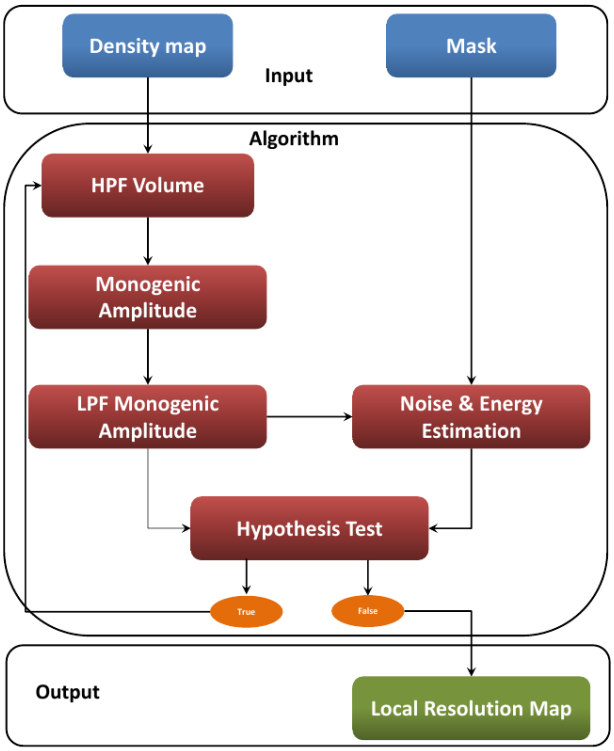

Figure 24: Scheme of MonoRes algorithm. HPF and LPF means high pass and low pass filter respectively.

visualization of the macromolecule. The knowledge of local resolution suggests to do this postprocess filtering in a local manner. Additionally to the local resolution map, *MonoRes* is able to provide a local filtered map, thanks to it implements this filtering taking advantage of the high pass filter. Note that, when the map is high pass filter, a low pass filtered map is also obtained by computing $V_{LP,\omega_0}(\mathbf{r}) = V(\mathbf{r}) - V_{HP,\omega_0}(\mathbf{r})$. Thus, the voxels at resolution ω_0 are taken from the low pass filtered map (at the same frequency) creating the local filtered map, taking advantage of the frequency sweep that MonoRes uses.

5.5 MONORES RESULTS

To analyze *MonoRes*, it was tested with synthetic maps as well as experimental reconstructions showing a good performance in both scenarios. Synthetic data were produced with atomic models taken from the PDB [13] database and, converted then into density map. In contrast, the experimental maps were directly taken from EMDB [95].

5.5.1 *Tests with synthetic maps*

The aim of the tests was to check the capabilities of *MonoRes* in maps with known resolution. Despite the fact that in most of cases the use of synthetic volumes is far from reality, the tests represent necessary conditions for resolution measurement,

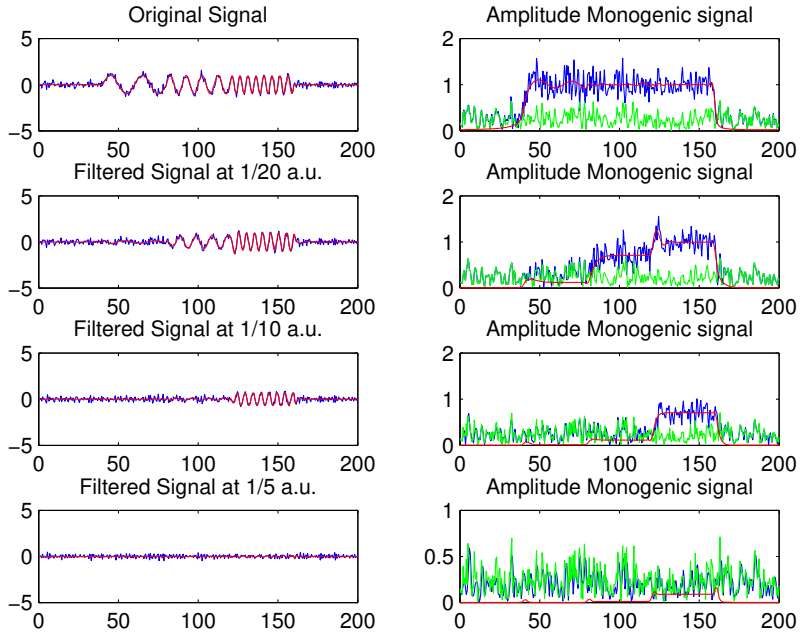

Figure 25: Explanatory unidimensional example of *MonoRes* algorithm. (left) a signal composed by several frequencies is high pass filtered. (right) The monogenic amplitude detects the loos of energy as a consequence of the high pass filter. Original signal and its amplitude in red. Noisy signal and its corresponding monogenic amplitude in blue, the green line defines the monogenic amplitude of the noise.

and therefore, they serve to calibrate the algorithm and determine its accuracy.

5.5.1.1 *Tests with low pass filtered map*

This first test was addressed to: 1) check if *MonoRes* is able to recover the highest frequency of a map when it is low pass filtered at a given resolution; 2) validate the use of two halves and, 3) determine the computational capabilities of *MonoRes* dealing with large maps. To do that, a synthetic map was created from the atomic model of type IVa pilus machine (PDB-3JC9) [24] by converting the atomic model into a density map using `xmipp_volume_from_pdb` [161], which has proved to produce accurate simulations of atomic structures.

1. *Map 1 - Regular size map, filtered at 10Å*: The conversion from the atomic model to the density map was performed specifying a sampling rate of 1 Å/pixel, resulting in a volume of $500 \times 500 \times 500$ voxels. This map was then low pass filtered to 10 Å; meaning that Fourier coefficients are multiplied by 1 up to 10 Å , and then the filter smoothly falls to 0 at a resolution of 8.3 Å. Finally, gaussian noise was added with standard deviation of 0.08.

2. *Half Maps, filtered at 10 Å*: This is essentially the same case as Map 1, but we now wanted to test the *MonoRes* capability for accepting as input to half

maps and correctly estimate the noise map from them. The first half map was Map 1, meanwhile the second half map was created as the Map 1, but making sure that the noise realization was different.

3. *Map 2 - Large size map, filtered at 10 Å*: A large size map was created in order to test how *MonoRes* deals with large volumes. In this case, the synthetic electron density map was calculated specifying a sampling rate of 0.5 Å/pixels, yielding a map of size $1,000 \times 1,000 \times 1,000$ voxels. This map was again filtered at 10 Å with a transition band of 0.02 (normalized frequency units) in width. The filtering process makes that frequencies higher than 7.1 Å should not be part of the filtered map. The noise added was gaussian with standard deviation of 0.08 a.u.

In the three examples the effect of the low pass filter was to remove all resolutions higher than 10 Å. Therefore, the expected value of local resolution should not be higher than the cutoff frequency at 10 Å. The results of local resolution estimation can be seen in Fig. 26 where the local resolution maps and resolution histograms for the Map 1 and the two half maps are shown. The third synthetic map that consider the large map with size $1,000 \times 1,000 \times 1,000$ voxels was not included in the figure, because it casts essentially the same histogram and map that the synthetic Map 1. In all cases, it must be highlighted that as it was expected, *MonoRes* was not able to find resolutions beyond the transition band and most of the resolution are around the cutoff frequency of 10 Å. The resolution histograms are very useful for identifying this fact, see Fig. 26.

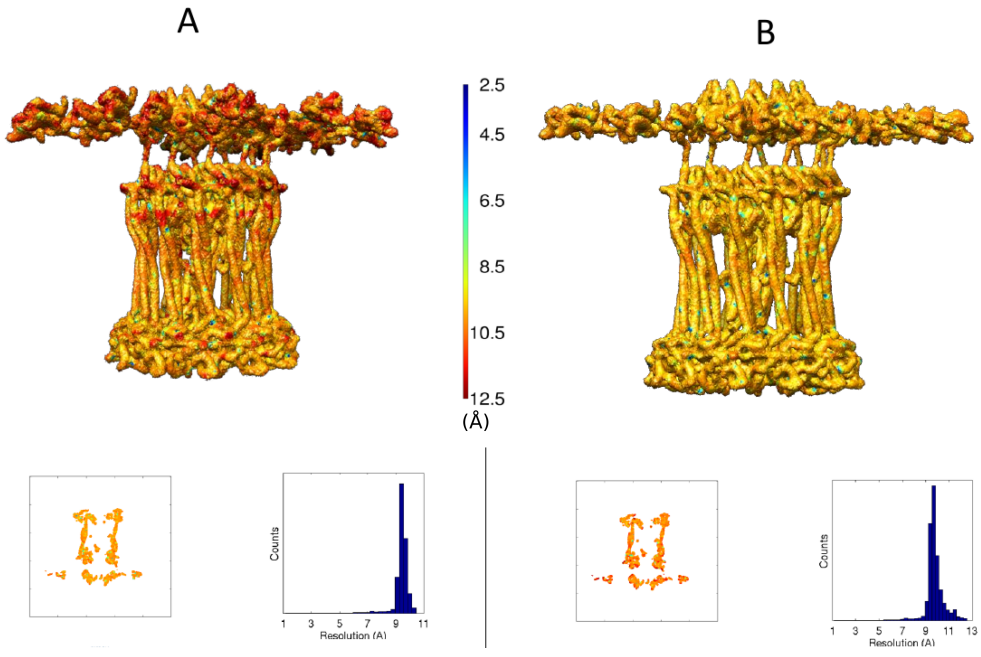

Figure 26: Results of *MonoRes* showing the colored map by local resolution values, a representative slice and the histogram of resolution for (A) Map 1 - Regular size map filtered at 10Åand (B) using half maps, filtered at 10Å.

5.5.2 *Test with multiple low-pass filtered regions*

The previous tests are necessary conditions of a local resolution method, however, density maps present a broad range frequencies. In this test, *Monores* deals with this situation. Hence, the atomic model of the glutamate dehydrogenase (PDB-5k12) [109] was converted into density map using `xmipp_volume_from_pdb` [161]. The converted map was then segmented in the regions that were independently low pass filtered at 2, 4 and 6 Å respectively. Finally, gaussian noise was added with zero mean and standard deviation of 0.08 a.u. The local resolution of this map was estimated with *MonoRes* which was able to correctly determine the filtering frequency of each region as it is shown in Fig. 27.

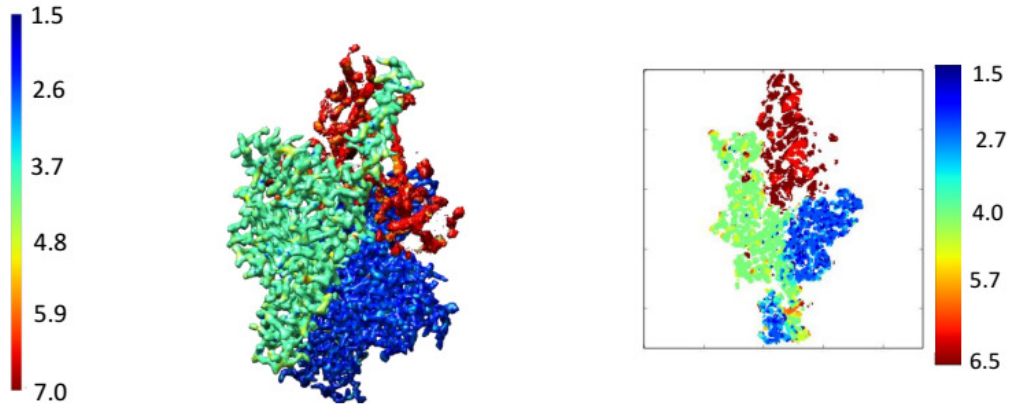

Figure 27: Local resolution map and a representative slice calculated with *MonoRes* for the the map obtained from the atomic model of the glutamate dehydrogenase filtered by segment at frequencies of 2, 4 and 6 Å.

5.5.3 *Tests with experimental maps*

Once the performance of *MonoRes* was checked in controlled scenarios, it was applied to determined the local resolution of experimental maps deposited in EMDB [95]. In particular, the following maps were considered: the β-galactosidase [1] (EMDB-6287), the proteasome 20S (EMDB-6287) [20], capsaicin receptor TRPV1 (EMDB-5778) [99], an aquareovirus (EMDB-5160) [200], and Nucleosome (EMDB-5160) [26]. In these cases, it is not possible to predict the results for the local resolution values. However, it is expected that they will be close to the FSC resolution reported by the authors of these maps. In addition, other local resolution methods, as they are *blocres* and *ResMap* were also used to establish a comparison. Note that they compute the local resolution with different algorithms, and therefore their definition of local resolution is different, that makes that the calculated local resolution maps are susceptible of presenting slight differences.

5.5.3.1 *Experimental Map 1: β-galactosidase*

The experimental map of the β-galactosidase in complex with a cell-permeant inhibitor (EMD-2984)[1] was considered as first experimental example to estimate its local resolution map. It presents a pixel size of 0.637Å/pixel and box dimensions of $292 \times 292 \times 292$ voxels. The reported FSC gold standard at 0.143 was 2.2 Å, so that, it represents a very high resolution map. The EMDB entry 2984 contains the full map as well as the two half maps. Thus, the local resolution map was calculated using both approaches: 1) a mask and the full map or 2) the mask and two half maps. The mask was created preserving the structural information by thresholding the map, and applying a small dilation. In addition, it was provided for computing the local resolution of this experimental example by different methods (*MonoRes*, *ResMap* and *blocres*). The results can be observed in Fig. 28.

1. *A mask and the full map*: The local resolution map was computed using *MonoRes* and *ResMap*. The obtained local resolution values were in the range [2.0,4.8] Å, with median at 3.3 Å and standard deviation of 0.6 Å for *MonoRes* and in the range [2.2,5.0] Å, with median at 2.6 Å and standard deviation of 0.6 Å for *ResMap*. The respectives resolution histograms and a representative slice of the local resolution map are shown in Fig. 28. It is noteworthy the resolution difference between the median that both methods report. To analyze this fact, the histogram of the resolution difference was calculated, they present a difference of 1.1 Å and a standard deviation of 0.6 Å. This fact will be explained below.

2. *A mask and two half maps*: In this case, the local resolution map was computed using *MonoRes*, *ResMap* and *blocres*. The obtained local resolution values were in agreement with the previously obtained for the full map. Thus, *MonoRes* got a median resolution of 3.2 Å with a standard deviation of 0.4 Å, *ResMap* cast a median resolution of 2.7 Å with standard deviation of 0.6 Å and finally *blocres* reported a median resolution of 2.6 Å with standard deviation of 0.1 Å using the criterion of FSC at 0.5. Regarding to the histogram of resolution difference between the three methods, the results are almost identical to the full map. The difference *Monores-blocres* shown a median value of 1 Å with standard deviation of 0.6 Å, and the difference *MonoRes-ResMap* cast a difference median value of 1 Å and a standard deviation of 0.4 Å. This result points out that the resolution measurement is robust using a single or two half maps.

To analyze the resolution disagreement around 1 Å in the median local resolution between *MonoRes* and the other resolution methods, we decided to analyze the map of the β-galactosidase. In particular, we use the definition of median, thus, a median reported by *ResMap* and *blocres* around 2.6-2.7 Å, means that the half of the voxel present resolutions higher than 2.6-2.7 Å. As a consequence, if a high pass filter at 2.6 Å is applied, the half of the structure should be preserved, or at least structural information should be observed. In Fig. 29 this experiment was carried out, thus, it shows the central slice of the high pass filtered map. Note the absence of structural information, this fact is reported by *MonoRes* and is the origin

of the discrepancy between these methods.
It also should be remarked that *MonoRes* uses the mask for locate the points in which the local resolution will be computed. Thus, the *MonoRes* map is very tight in contrast with the reported by *ResMap* that uses another algorithm.

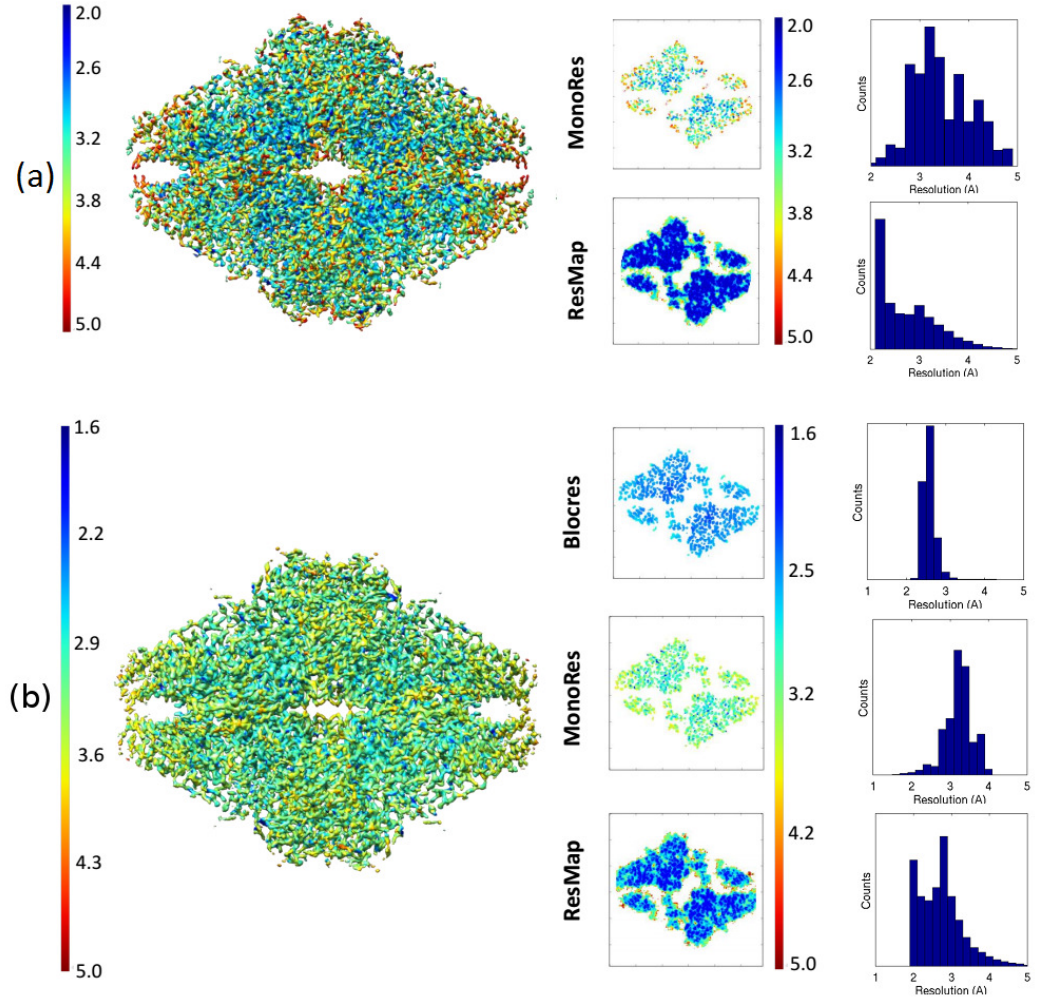

Figure 28: (left) Local resolution map of the β-galactosidase (EMDB entry 2984) obtained with *MonoRes*, (center) a representative slice calculated with *MonoRes* or *ResMap*, and their local resolution (right). (a) Results if a single full map, and (b) considering two half maps.

5.5.3.2 *Experimental Map 2: Proteasome 20S*

The second case of use considers an experimental map of the Proteasome 20S taken from EMDB (EMDB entry-6287)[20]. It presents a pixel size of 0.982 Å/pixel with box dimensions of $300 \times 300 \times 300$ voxels. The reported FSC gold standard at 0.143 was 2.8 Å. As well as it happened with the previous example, this EMDB entry provides the full maps and two half maps. The local resolution map will be then calculated using both, i.e. the full map and a mask, and the two halves and the same mask. Again, the mask was carefully created by thresholding the map avoiding the loose

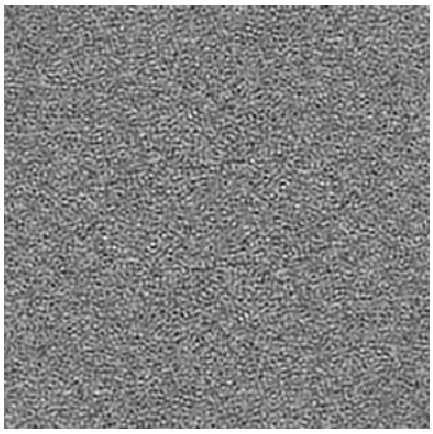

Figure 29: Central slice of the β-galactosidase, when the map is high pass filtered at 2.6 Å.

of structural details. Hence, the local resolution map was calculated with *MonoRes*, *ResMap* and *blocres* (with two halves). The results can be observed in Fig. 30

1. *A mask and the full map*: *MonoRes* and *ResMap* were used to estimate the local resolution of this map. *MonoRes* reported resolution in the interval [2.0, 4.6] Å with median of 2.8 Å and standard deviation of 0.5 Å. *ResMap* got similar results, its resolution range was [2.2, 6.0] with median at 2.7 Å and standard deviation of 1.0 Å. As a consequence, both methods cast essentially the same result. The histogram of local resolution difference between *MonoRes* and *ResMap* supported this result with a median of 0.4 Å and standard deviation of 0.5 Å. The results can be observed in Fig. 30

2. *A mask and two half maps*: In this case, the local resolution estimation was carried out considered *MonoRes*, *ResMap* and *blocres*. The results are summarized in Fig. 30. Hence, *MonoRes* got a median resolution of 2.9 Å with standard deviation, of 0.5 Å, *ResMap* determined a median resolution of 2.9 Å with standard deviation of 1.0 Å, meanwhile *blocres* reported a median of 3.0 Å with a standard deviation of 0.2 Åmeasured with the FSC criterion at 0.5. In essence the three methods got the same results and spectral range. To check that, the histograms of resolution difference were calculated. Thus, the difference between *MonoRes* and *ResMap* exhibits a median of 0.6 Å with standard deviation of 0.6 Å, and the difference between *MonoRes* and *blocres* reports a median of -0.03 Å and with standard deviation of 0.5 Å. This result points out that there is, in this case, a high similarity between *MonoRes* and *blocres*.

5.5.3.3 *Experimental Map 3: Aquareovirus*

To test *MonoRes* with large maps the reconstruction of an Aquareovirus (EMDB entry 5160) [200] was considered. It presents a size of $740 \times 740 \times 740$ voxels with a pixel size of 1.1 Å/pixel. The reported FSC-resolution for this map was 3.2 Å. In the tests involving synthetics maps *MonoRes* was able to deal with larger maps than this one. However, now the real resolution is unknown, but must be close to

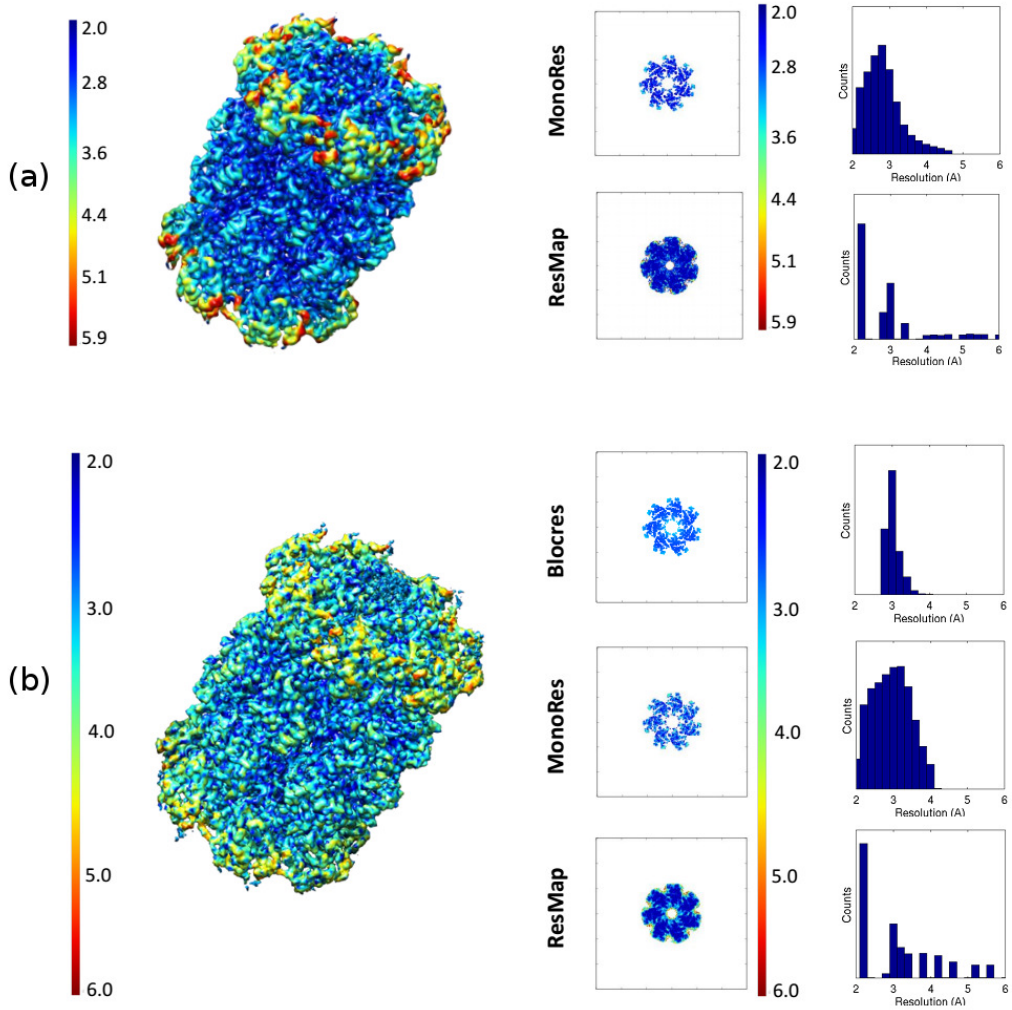

Figure 30: (left) Local resolution map of the Proteasome 20S (EMDB entry 6287) obtained with *MonoRes*, (center) a representative slice calculated with *MonoRes* or *ResMap*, and their local resolution (right). (a) Results if a single full map, and (b) considering two half maps.

the FSC value.

The dimensions of this large maps considerably increase the computational time requires by other algorithms involves to a point of making them impractical. For that reason, only *MonoRes* algorithm was used to compute the local resolution. The results are summarized in Fig. 31. *MonoRes* pointed out a median resolution of 4.2 Å with standard deviation of 0.6 Å, being the resolution range from 3.0 Å to 5.6 Å. It can be observed in Fig. 31 that the inner structure of the virus presents higher resolution than the outer part.

5.5.3.4 *Experimental Map 4: capsaicin receptor TRPV1*

The fourth experimental example considers the capsaicin receptor TRPV1 (EMDB entry 5778) [99]. This map is specially of interest because it is a membrane protein that presents a wide range of resolutions. The map has a dimension of 256 ×

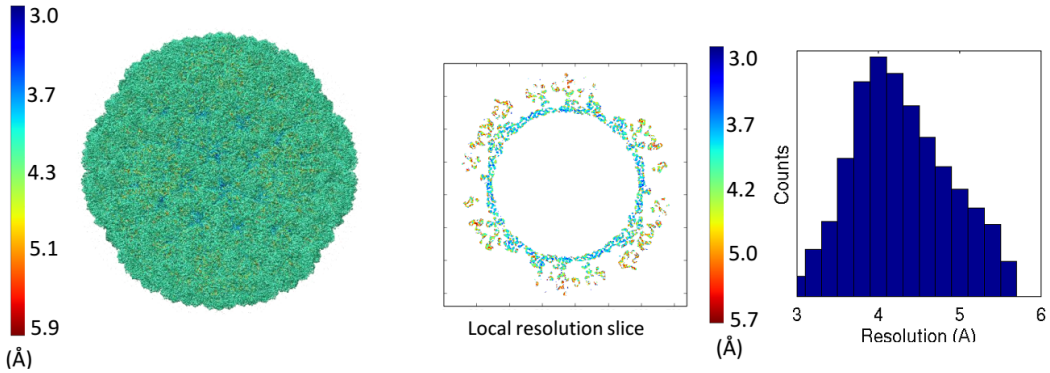

Figure 31: (left) Local resolution map of the Aquareovirus (EMDB entry 5160), (center) a representative slice, and (right) the local resolution histogram.

256 × 256 voxels and a pixel size of 1.1 Å/pixel. The reported resolution was resolution 3.8 Å calculated via FSC at 0.143 with gold standard procedure. The mask was again created by thresholding the map, and then the local resolution was calculated, in this case only considering the full map. *MonoRes* and *ResMap* were applied to estimate the local resolution map. The results, that can be seen in Fig. 32, show a resolution range from 2.0 Å to 8.0 Å. *Monores* obtained a median resolution of 4.2 Å with standard deviation of 1.1 Å. In contrast, *ResMap* got a median local resolution of 4.0 Å with standard deviation of 0.8 Å. These results are very similar, however, *MonoRes* presents a wider dispersion than *ResMap*. It is noteworthy that in both cases, the dispersion is higher than in the previous examples, in agreement with the broad resolution range that this map presents. However, *MonoRes* seems to be slightly more conservative in the resolution measurement than *ResMap*, the histogram of resolution difference support this fact with a median resolution of 0.8 Å and standard deviation of 0.95 Å.
Finally, the use of the local filter that *MonoRes* provides was used. This reconstruction is a good candidate to show this filter, because the broad resolution range that it presents. The results of the local filter are shown in Fig. 32, exhibiting a clear map as a consequence of the local noise suppression.

Experimental Map 5: Nucleosome

This last experimental example will illustrate how *MonoRes* is able to work even at low resolution as it usually occurs when the structure is obtained via subtomogram averaging. Thus, the nucleosome particles decorating chromatin plates from metaphase chromosomes (EMDB-0117) were analyzed [26]. This map presents a pixel size of 4.21 Å/pixel and very small dimensions 50 × 50 × 50 voxel. This structure was reported with a FSC-resolution value (at 0.143) of 30 Å.
The local resolution map was computed with *MonoRes* which casts as result the map shown in Fig. 33. Note that the FSC reported is in the resolution range provided by *MonoRes*. The low resolution of this map complicates to trace an atomic model. However, the local resolution map enriches the reconstruction, the very low resolution regions support the hypothesis that the DNA wraps the core (in red).

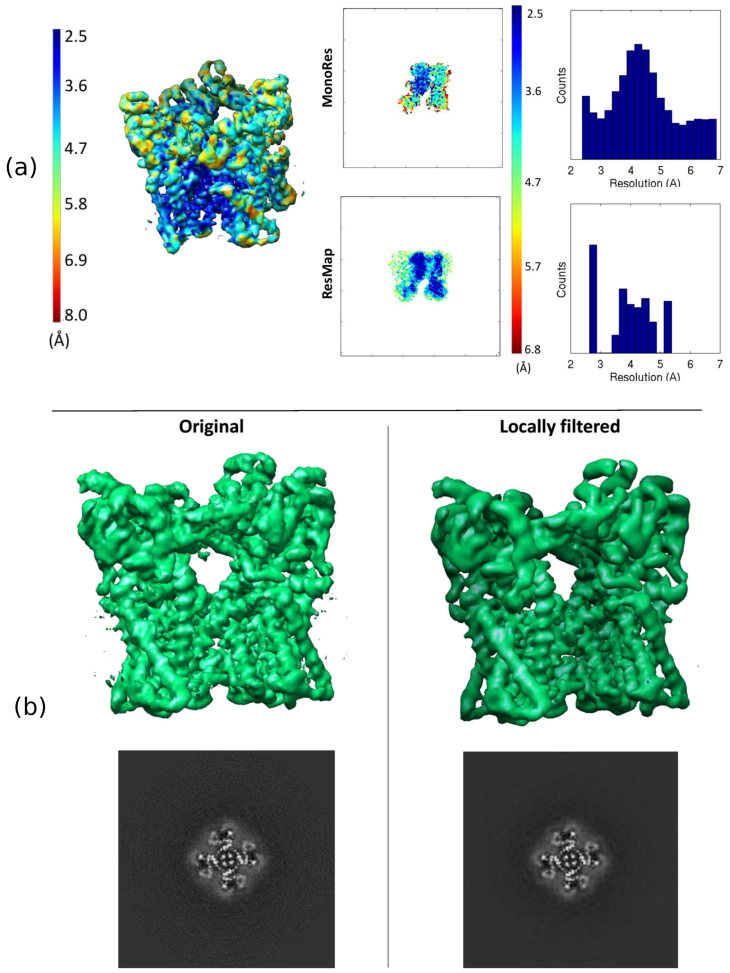

Figure 32: Results of the local resolution analysis with *MonoRes*, *ResMap* and *blocres*. (left) Local resolution map of the capsaicin receptor TRPV1 with *MonoRes* (EMDB entry 5778), (center) a representative slice, and (right) the local resolution histogram.

Moreover, the histones regions present a higher resolution (blue color) than the rest of the map, meanwhile the membrane also exhibits a very low resolution (in red).

Table 2 summarizes the results of *MonoRes*, *ResMap* and *blocres* for the analyzed experimental cases.

5.6 DISCUSSION OF MONORES RESULTS

A new algorithm, named *MonoRes*, for estimating the local resolution of SPA density maps has been proposed. The method is based on a local decomposition in phase and amplitude of the density map at different frequencies, analyzing if the local amplitude of the signal is significantly higher than the amplitude of noise. In this sense, *MonoRes* can be understood as a local measurement of SNR, being the local resolution value as the limit in which noise and signal cannot be distin-

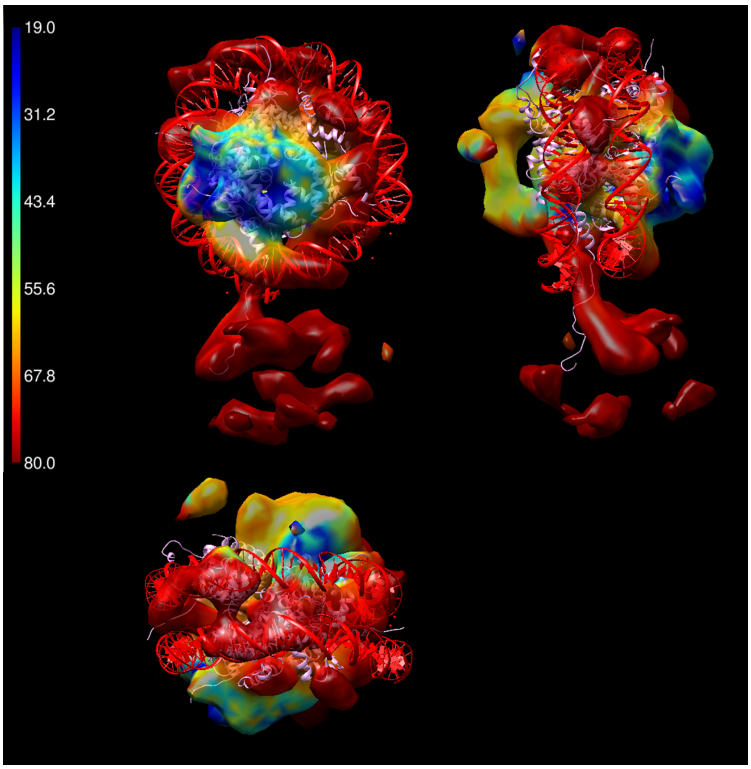

Figure 33: Local resolution map under different points of view of the nucleosome structure obtained via subtomogram averaging from chromatin plates from metaphase chromosomes.

Volume	FSC (Å)	*MonoRes* range (Å)	*ResMap* range (Å)	*blocres* range (Å)
β-Galactosidase	2.2	[2.0, 4.8] (3.3)	[2.2, 5.0] (2.6)	–
(using halves)	2.2	[1.6, 4.0] (3.2)	[2.0, 5.0] (2.7)	[2.2, 4.2] (2.6)
Proteasome	2.8	[2.0, 4.6] (2.8)	[2.2, 6.0] (2.7)	–
(using halves)	2.8	[2.0, 4.1] (2.9)	[2.2, 6.0] (2.9)	[2.8, 4.0] (3.0)
Aquareovirus	3.6	[3.0, 5.6] (4.2)	–	–
Nucleosome	30.0	[19.0, 80.0] (–)	–	–

Table 2: Summary of local resolution for the experimental cases. The number in parenthesis is the local resolution median.

guished.

The set of tests involving synthetic maps provide a very rich understanding of the local resolution concept. When a map is low pass filtered, the higher resolution information beyond the cutoff frequency is completely destroyed. This simple test is, therefore, a proof of concept of the local resolution measurement. *MonoRes* was able to pass these tests with success. Moreover, it can deal with very large maps, in short computational times.

The examples with experimental maps allowed to check the *MonoRes* performance with real data, and to establish a comparison with other methods of the state of the

art, as they are *blocres* and *ResMap*. The comparison showed that *MonoRes* seems to be a more conservative local resolution measurement.

5.7 ON THE RESOLUTION CONCEPT AND THE ORIGIN OF LOCAL RESOLUTION FOR ELECTRON TOMOGRAPHY

Resolution as a global measurement was explained in Chapter 4, in particular, the FSC as the standard metric in SPA. In electron tomography, the resolution measurement is not broadly used, despite that there are a few methods to compute it. The most spread is an extension of the FSC metric. However, it requires two independent tomograms, that are named odd and even because they are reconstructed with the half of the tilt series images, i.e. the odd and even images are used to reconstruct their respective tomograms. Due to this slight difference, this FSC is denoted as $FSC_{e/o}$ [21]. Alternative to this method, the so-called Noise-compensated Leave One Out (NLOO) can be applied [21]. In this approach, the tomogram is reconstructed with all images except one, and then it is re-projected along the orientation of the orientation of the image that was left out to finally compute the FSC between the re-projection and the image. This alternative is more robust than the $FSC_{e/o}$ because it keeps the angular sampling, but less used due to its higher complexity. Other metrics were also proposed as the use of FSC under conical filtering [121]. For a deep review on resolution measures in electron tomography see [40, 91].

The need of a local resolution measurement was discussed in section 5.1. However, the physical and mathematical origins that produce it are different in SPA and electron tomography. It is obvious that the sample geometry and the image quality will determine the local resolution, but there are external aspects that produce resolution loses.

The main difference between SPA and electron tomography is the measurement and reconstruction processes. While in SPA the conditions of homogeneous sample and projection assumptions were considered (see Chapter, 3, section 3.1) in electron tomography the first one does not apply because, there is only one complex to be imaged. Thus, the 3D information is obtained by tilting the specimen and acquiring images at different tilt angles, the set of tilt images is named as a tilt series. In SPA, the homogeneous sample condition can be slightly broken as a consequence of sample flexibility, or conformational changes. Note that these two factors explains the loss of resolution in certain regions of the macromolecule. However, because in electron tomography there is only one complex, local resolution cannot be caused by effect of flexibility or heterogeneity and, it must be produced by other factors.

Electron tomography samples are usually larger and thicker than SPA ones. These two facts along with the tilting process represent sources of local resolution variations in electron tomography. Note that a point close to the tilt axis are better resolved in the tomogram than the further ones. By tilting, the sample suffers a spatially variant defocus in the acquired images. This introduces a blurring that reduces the resolution. The ice thickness is also important, note that the large dimension of the specimen makes that the ice thickness depends on the sample position. These fact imply a spatially variant SNR. Finally, regarding to the recon-

struction process, an image alignment is required (tilting the sample introduces an image misalignment), thus, the quality of the reconstruction depends on the goodness of this alignment, which has also impact on the local resolution.

5.8 MONOTOMO - A LOCAL RESOLUTION METHOD FOR ELECTRON TOMOGRAPHY

The use of *MonoRes* for measuring local resolution has no more limitations than the noise characterization. However, its application to electron tomograms instead of SPA reconstructions imply the following two important problems:

1. *Dealing with very large maps*: In contrast to SPA in which a single macromolecule is reconstructed, electron tomography considers the reconstruction of big structures of assemblies of macromolecular complexes, making that the reconstructed tomogram presents large dimensions. It represents a computational challenge in terms of computation and memory. It implies that *MonoRes* algorithm must be modified in terms of computational speed.

2. *Noise characterization*: This is the main problem that for sake of simplicity it will be split in two:

    a) *Access to a noise model*: The use of a mask gives access to the noise distribution (this happens with a single or two half maps). In the case of electron tomography, the use of a mask is also possible. The very low SNR that tomograms present difficult in many cases the structure distinction, and therefore, the mask creation. Thus, the mask requirement should be avoided, making necessary another method to get a model of noise.

    b) *Spatially variant noise*: One of *MonoRes* assumptions is that noise is spatially homogeneous. This hypothesis is not necessary true in electron tomography. In the previous section, two facts were mentioned, the spatial dependence of the ice thickness and defocus, which might produce a spatially variant noise. Thus, *MonoRes* algorithm must be adapted to deal with this new kind of noise.

A fully automatic algorithm, named *MonoTomo*, was proposed as a *MonoRes* adaptation to overcome these two drawbacks.
The first problem is least important, and it was easily solved by a careful reimplementation of *MonoRes* algorithm. However, the second one completely modifies the way of measuring. The following solutions were proposed:

1. *Solution to the access to a noise model*: It considers the approach introduced in Section 5.3 in which a map of noise is calculated by the difference of two half maps, see Eq. (64). Thus, a tomogram of noise can be obtained by the difference of two tomograms reconstructed with the half of the acquired images. These two half tomograms, named odd and even, can be obtained following one of the next two methods that are graphically explained in Fig. 34:

a) *Splitting the set of micrographs*: This approach was suggested in [21]. The images that define the tilt series (set of micrograph acquired at a different tilt sample angles) is split in two independent subsets of images. They are created by assigning alternatively the images of the tilt series, i.e. the images corresponding to odd and even angles will define the odd and even set of images respectively. Hence, the so called odd and even tomograms are reconstructed. It must be highlighted that if $\Delta\theta$ is the angular step of the whole tilt series, the odd and even set of images will present a step of $2\Delta\theta$.

b) *Splitting the set of frames*: Alternatively to the previous approach, we suggest to split the set of frames that define acquired movie at each tilt angle of the tilt series. Thus, the set of frames is split in two subsets of frame, that by language extension will be also named odd and even. They allow to carry out two independent tomograms again named odd and even. The previous alignment step is carried out with all images, once they have been aligned, the frame spitting is undertaken. It is noteworthy that using this approach, the angular step of the tilt series is kept in the odd and even tomograms. It will be shown that this concept can take advantage specially when the tilt series was acquired with a high angular step.

As it occurred in SPA with half maps, a mean tomogram is calculated, i.e. if $T_o(\mathbf{r})$ and $T_e(\mathbf{r})$ are the odd and even tomograms, then, the mean tomogram is calculated as,

$$T(\mathbf{r}) = \frac{1}{2}\left(T_o(\mathbf{r}) + T_e(\mathbf{r})\right), \tag{69}$$

and the noise as

$$N(\mathbf{r}) = \frac{1}{2}\left(T_o(\mathbf{r}) - T_e(\mathbf{r})\right). \tag{70}$$

Note, that due to the noise model is obtained via difference (odd and even), the angular sampling of the tilt series affects to the model, in particular, when it is high. In this sense, by splitting frames has the advantage of better determination of the noise model because it keeps the angular step of the tilt series. This is the reason why, we observed that when the angular sampling of the tilt series is greater than 2 degrees, the estimation of local resolution by splitting in odd and even micrographs might present drawbacks as a consequence of the $2\Delta\theta$ angular step of the odd and even images, in contrast to the approach of splitting the set of frames. This problem will be explained in the discussion section.

2. *Solution to spatially variant noise*: *MonoTomo* solves it by means of a moving window of size $100 \times 100 \times 100$ pixels. The noise distribution in each window is then calculated, which allows to determine the noise threshold for performing the corresponding hypothesis test. It is reminded that the threshold is calculated from the CDF of the noise distribution with a significance of $1 - \alpha = 0.95$, as was pointed out in Eq. (67)

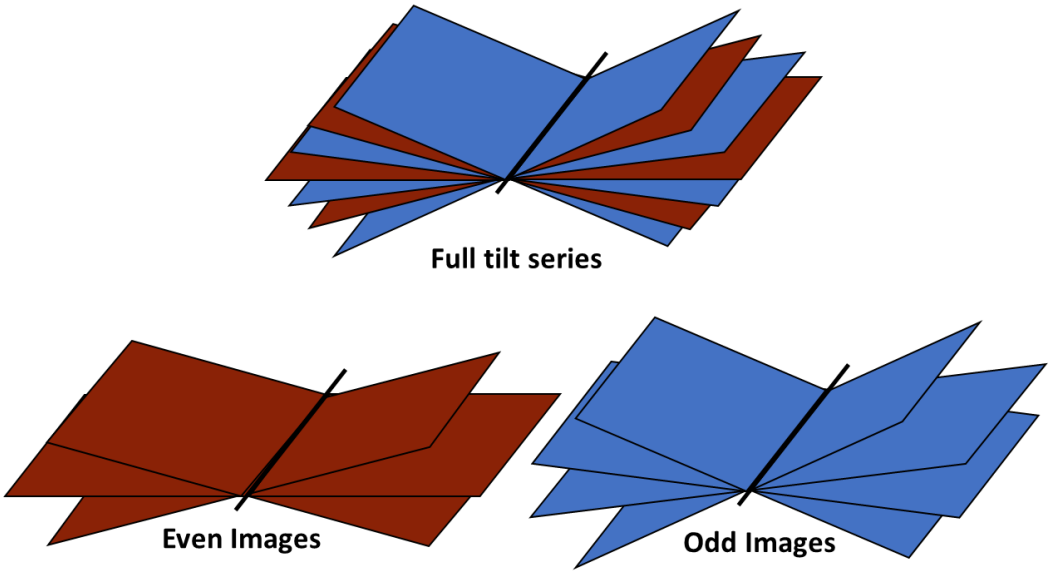

Figure 34: Set of images is split in odd and even images to reconstruct the odd and even tomograms respectively.

Once the noise model is obtained with its spatial dependence, it is possible to carry out the *MonoTomo* algorithm to determine the local resolution of the calculated mean tomogram. The idea is essentially the same as on *MonoRes*, *MonoTomo* will perform a frequency sweep from low to high resolutions, high-pass filtering the tomogram (step 1 - *MonoRes*), $T(\mathbf{r})$, calculating the monogenic amplitude (step 2 - *MonoRes*) and then low-pass filtering it (step 3 - *MonoRes*). Finally, an hypothesis test is performed with the voxels of the filtered monogenic amplitude obtained in the step 3. However, this hypothesis test must be considered in a local manner, due to the noise distribution might be spatially dependent. For this reason, *MonoTomo* calculates the spatially variant thresholds. If the filtered monogenic amplitude of a specific voxel if statistically greater than noise in a local sense (hypothesis test), that voxel can be measured at the filtering frequency. Finally, regarding to the possibility of false positives in the hypothesis test, the *MonoRes* criterion of failing consecutively twice the test, is kept in *MonoTomo*.

5.9 MONOTOMO RESULTS

MonoTomo was tested on experimental Cryo-Electron Tomography (cryoET) data taken from EMPIAR (entries 10110[25], 10115[171], 10027[82] and 10164 [152]). In contrast to *MonoRes* results were the maps can be directly computed once they are taken from the data base, these data sets provide the raw data as a tilt series which implies to carry out the alignment step before reconstructing the tomogram. This process was common to all data sets.
The alignment step was based on fiducials and performed with IMOD [89]. Thus, the splitting step was undertaken. In the cases of EMPIAR entries 10110, 10115, 10027 the raw data set only contains images, and therefore, they were split into odd and even micrograph. In contrast, the EMPIAR entry 10164, that has movies,

was split following the frames approach. Independently of the splitting method, the two subsets odd and even were aligned by considering the parameters of the alignment of the full set. Hence, the odd and even tomograms were reconstructed and these reconstructions are the *MonoTomo* input.
MonoTomo result is a local resolution map in which each voxel represent the local resolution of the same voxel in the mean tomogram computed from the two halves. With the aim of determining the validity of *MonoTomo* results, the $FSC_{e/o}$ was also calculated, with and without a mask to remove the background. Results are summarized in Table. 3.

EMPIAR entry	10110	10115	10027	10164
Tilting Range (^{0})	[-60.00,60.00]	[-64.00,64.00]	[-60.00,60.00]	[-60.00,60.00]
Angular step (^{0})	1.0	1.0	1.5	3.0
Pixel Size (Å)	4.04	9.46	3.3	0.675
Binning (a.u.)	4x Å	2x	4x	2x
Tomogram Size	$928 \times 960 \times 400$	$1016 \times 1096 \times 200$	$960 \times 928 \times 200$	$960 \times 928 \times 400$
Splitting	Micrographs	Micrographs	Micrographs	Frames
MonoTomo (Å)	[40,120] Å	[40,120]	[30,110]	[15,45]
$FSC_{e/o}$ (Å)	121 Å	128 Å	98 Å	32 Å
$FSC_{e/o}$ masked	43 Å	45 Å	31 Å	30Å

Table 3: Experiments summary. The $FSC_{e/o}$ threshold was (0.143).

5.9.0.1 *Experimental Case 1: EMPIAR entry 10110*

The data set was a tilt series taken from the EMPIAR entry 10110 [25], which contains *Vibrio cholerae* cells imaged. The angular step of the tilt series was of 1 degree with tilt angles from -60.00 to +60.00 degrees and a pixel size of 4.04 Å. With the aim of reducing the computational effort, the images were binned by a factor 4 that implied an increment of a pixel size to 16.16 Å/pixel. Once the set of images was aligned (with IMOD), it was split (micrograph splitting) in the odd and even subsets to reconstruct the odd and even tomograms, each of them with $928 \times 960 \times 400$ voxel. *MonoTomo* computed the local resolution. In Fig. 35 slices under different view of the local resolution map as well as the original tomogram are shown. Looking at Fig. 35 *MonoTomo* points out that resolution is achieved in regions with biological information or structures like fiducials or borders, and they are in a range between $40 - 120$ Å. In addition, *MonoTomo* assigns the lowest resolution to the structure free regions, note that also there is a smooth transition in resolution range between structure content and background. Finally, the FSC_e/o (at 0.143) was 121 Å. This value that is in the limit of *MonoTomo* resolution range is explained because of the absence of the mask. However, when a mask is provided, the FSC_e/o (at 0.143) achieves a value of 43Å, in agreement with *MonoTomo* resolution range.

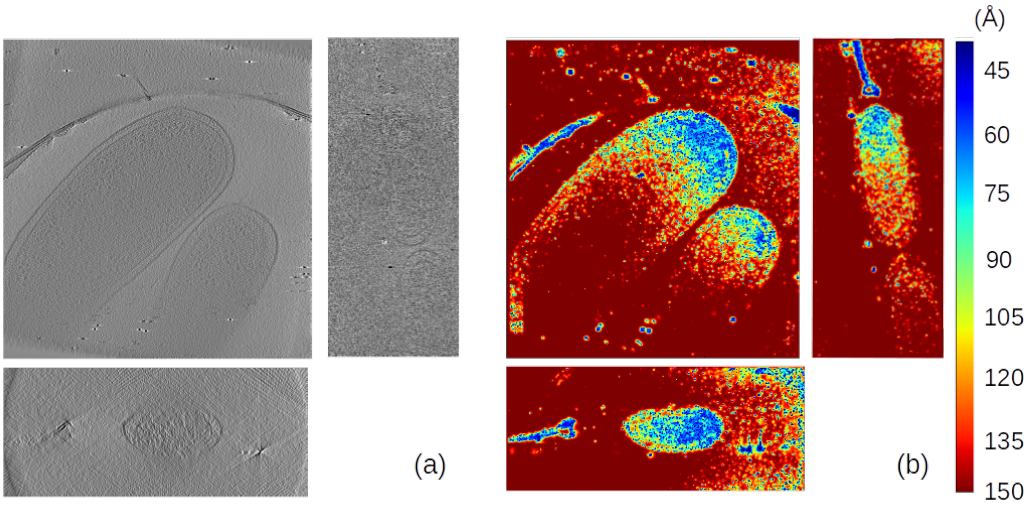

Figure 35: Tomogram (a) of the data set entry from EMPIAR 10110 and (b) its corresponding local resolution slices.

5.9.0.2 *Experimental Case 2: EMPIAR entry 10115*

A tilt series corresponding to imaged *Escherichia coli* cells at tilt angles from -64.00 to +64.00 degrees with an angular sampling of 1 degree were taken from EMPIAR entry 10115 [171]. In this case the pixel size was 9.46 Å/pixel. Again to speed up the computational time and reducing the computational charge, images were binned by a factor 2 which changed the pixel size to 18.92 Å/pixel. The images were aligned with IMOD and split (micrograph splitting) in the odd and even subsets to reconstructed the odd and even tomograms with size $1016 \times 1096 \times 200$ voxels. *MonoTomo* used them to calculate the local resolution map, see Fig. 36. In this case, *MonoTomo* cast resolution in the range from 40-120 Å for region with structural information, meanwhile the FSCe/o was 128 Å(at 0.143). Again, this value in the border of *MonoTomo* resolutions range is increased to 45 Å when a mask is provided.

5.9.0.3 *Experimental Case 3: EMPIAR entry 10027*

The tilt series taken from EMPIAR 10027 [82] contains images of the structure of intact bovine F1Fo ATP synthase in 2D membrane crystals. They were acquired with an angular sampling of 1.5 degrees with a tilt angle from -60.00 to +60.00 and a pixel size was 3.3 Å per pixel. The binning factor was 4 getting turning the pixel size into 13.2 Å per pixel. The alignment step (with IMOD) was carried out and the set of images was split (micrograph splitting) in odd and even sets to reconstruct the odd and even tomograms with dimensions of $960 \times 928 \times 200$ voxels. The local resolution was then calculated with *MonoTomo* showing a range from 30 to 110 Å for regions with structural information, see Fig. 37. Finally the Fourier Shell Correlation eve odd (FSCe/o) reported a value of 98 Å (at 0.143) when it is computed without mask and 31 Å if a mask is used. These values are both in *MonoTomo* resolutions range.

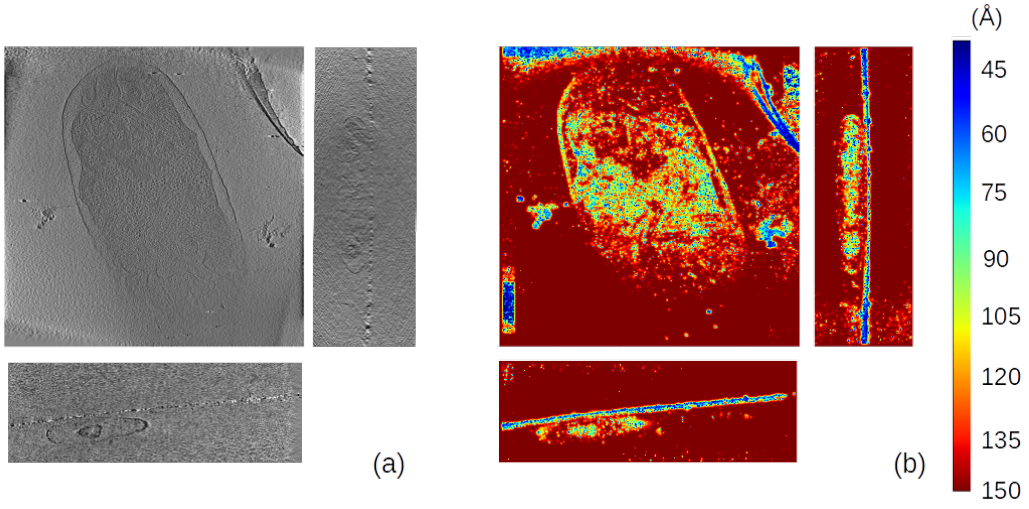

Figure 36: Tomogram (a) of the data set entry from EMPIAR 10115 and (b) its corresponding local resolution slices.

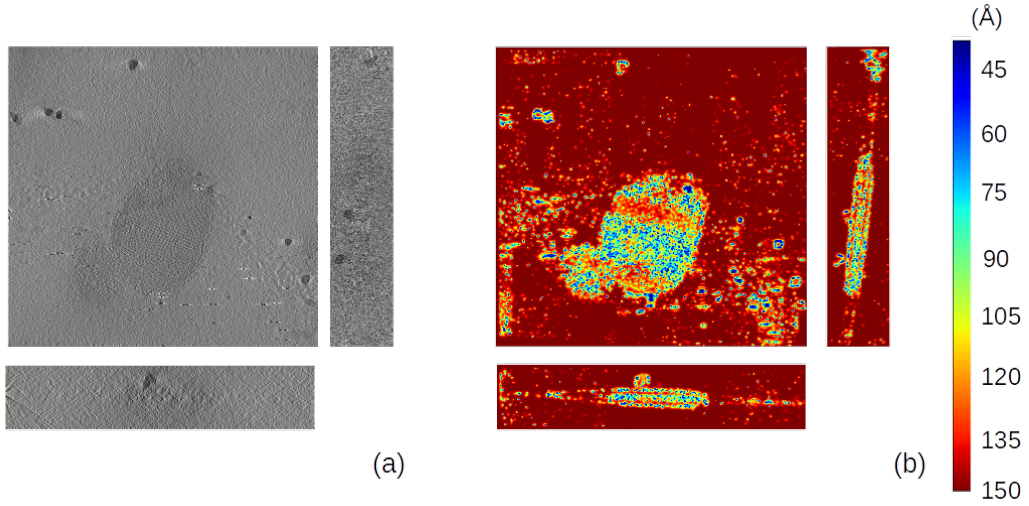

Figure 37: Tomogram (a) of the data set entry from EMPIAR 10027 and (b) its corresponding local resolution slices.

5.9.0.4 *Experimental Case 4: EMPIAR entry 10164*

This experimental example made use of the EMPIAR entry 10164 in which movies of the HIV-1 virus were acquired. It must be highlighted that in this case, the dataset is composed by movies, in contrast to the previous three examples. The angular sampling was 3 degrees from -60.00 to +60.00 degrees of tilt and the pixel size 0.675 Å. The movies of the tilt series were aligned with MotionCor2 [203] to get a tilt series of images. Thus, the set of frames at each tilt angle was split in two subsets (odd and even), and they were aligned using the parameter obtained parameters by MotionCor2. The result was two tilt series that keep the angular step of the original tilt series. Hence, both tilt series (odd and even) were finally

used to reconstruct the odd and even tomograms. Finally, *MonoTomo* was used to computed the local resolution map, which cast values in the range from 15 to 45 Å for regions structural information. The FSCe/o value was also calculated being 32Å (at 0.143) without mask, and 30 Å when a mask is provided. These values are in the *MonoTomo* range.
For this dataset it is remarkable that the angular step is the highest of the four examples shown in this thesis. It implies that if instead of splitting in frames, the splitting is carried out in micrographs, the angular sampling of each half would be 6 degrees, which represent a considerable lack of information causing problems in the detection of local resolution measurement. In fact, that test was carried out, and the result was that the local resolution map could not be determined. Therefore, splitting in frames is a worthy solution to compute the local resolution in tomograms that presents a high angular step.

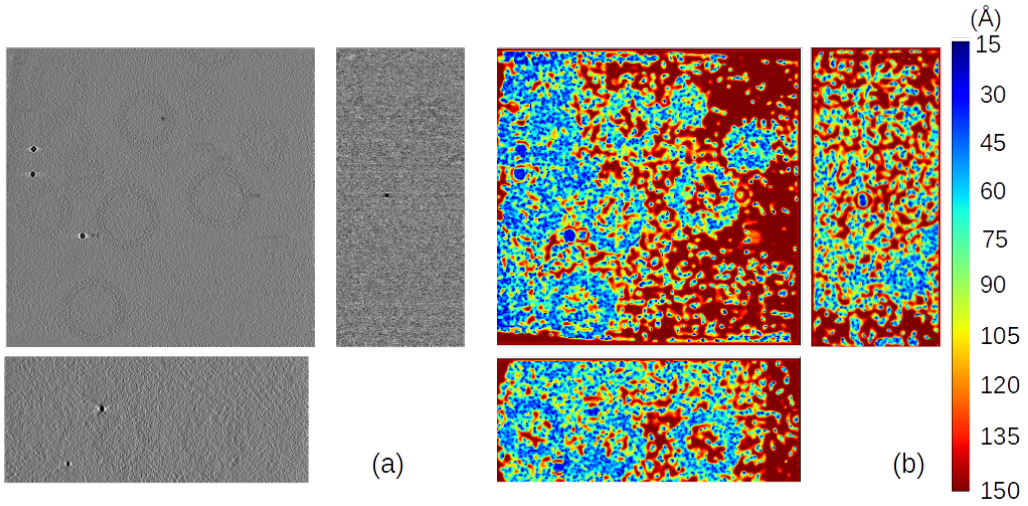

Figure 38: Tomogram (a) of the data set entry from EMPIAR 10164 and (b) its corresponding local resolution slices.

5.10 DISCUSSION OF MONOTOMO RESULTS

A new and fully automatic method, named *MonoTomo*, for estimating the local resolution was proposed. Up to our knowledge it represents the first local resolution method in the field in electron tomography. Its fundamental are the *MonoRes* algorithm with an adaptation to deal is spatially variant noise.
The measurement of resolution in electron tomography is not widely used, the most spread methods are global and they are the FSCe/o and the NLOO [21]. The first one has problems when the angular sampling is low. However, it is the most used. The FSC computes the cross-correlation between two independent reconstructions, however, as a cross-correlation measurement, it is mask dependent. It implies an added drawback for the resolution measurement, because the very low SNR in electron tomography makes the mask creation difficult, and also it may introduce a bias due to the subjectivity of the user when a mask is created. A proof of the

mask dependency can be observed in the results of $FSC_{e/o}$ values with and without mask for the experimental tomograms. *MonoTomo* is mask independent and only report resolution values in those voxels with structural information. As a consequence, it avoids the user bias of mask creation by an accurate statistical analysis of local SNR. Moreover, the reported resolution range that *MonoTomo* provides is in agreement with the $FSC_{e/o}$ with the advantage of being a local result that also informs about the resolution variation along the different tomogram regions.
To estimate the local resolution *MonoTomo* makes use of the monogenic amplitude. In our test we saw that for angular sampling lower than 3 degrees the reliability of *MonoTomo* is compromised, as a consequence of the very low SNR that complicates the proper estimation of the monogenic amplitude and the model of noise. For this reason the splitting frames approach that allows to keep the angular sampling gives good results with a low angular sampling of 3 degrees.
The analysis of local resolution in electron tomography has not been explored yet, being *MonoTomo* the first approach to this need. It obviously opens new horizons with impact in critical steps of electron tomography workflow, as they are the guidance of tilt-series alignment, particle picking for subtomogram averaging, local sharpening and local filtering, or as segmentation tool.

5.11 CONCLUSIONS

1. A new and fully algorithm, *MonoRes*, for computing the local resolution in SPA reconstructions has been proposed.

2. *MonoRes* assigns a resolution value when the local amplitude of the signal cannot be statistically distinguished (hypothesis test) from the amplitude of noise, i.e. it represents a local measurement of SNR.

3. *MonoRes* is invariant under B-factor correction and linear global transformations.

4. Results using synthetic data provides a high accuracy of the resolution estimation.

5. *MonoRes* design and its implementation allows to estimate the local resolution of large maps in short computational times.

6. The proposed method is also able to locally filter map by the local resolution values.

7. Up to our knowledge the first algorithm, *MonoTomo*, for computing the local resolution in electron tomography has been proposed.

8. *MonoTomo* is based on *MonoRes* considering that the noise is spatially dependent.

9. The splitting frames approach is introduced as a new method of splitting datasets.

10. As a first algorithm of local resolution, *MonoTomo* opens new possibilities in the field like: subtomogram averaging, alignment, segmentation, sharpening or local filtering.

5.12 APPENDIX: FUNDAMENTALS OF MONOGENIC SIGNALS

The analytic and monogenic signal in 2D mathematically well defined. Nevertheless, up to our knowledge, the community lacks of a consensus about the generalization of analytic signals to N-dimensions. In this appendix, it is suggested an extension to N-dimensions of the concept of Monogenic signal based on a generalization o the Riesz transform.

5.12.1 *Monogenic Signal in N-dimension*

In this appendix, a justification about the generalization to the advantages of analytic signal for extracting amplitude an phase in 1D signal are extrapolated to N-dimensional euclidean spaces. Unlikely, the Hilbert transform ought to be generalized to higher dimensions. The Riesz transform can be somehow considered that generalization. Following the same way as the Hilbert transform, the Riesz transform is derived in 2D in [39], a generalization to N dimensions is carried out. At the end using this generalization a definition of N-dimensional monogenic signal is proposed.

5.12.1.1 *Riesz Transform*

Considering an orthonormal basis, $\{\mathbf{e}_1, \mathbf{e}_2 ..., \mathbf{e}_N\}$, of $\mathbb{R}^N$, an arbitrary point with coordinates $(x_1, ..., x_N)$ will be represented by a position vector, $\mathbf{r}$. Thus, let us solve the Cauchy problem defined by the partial differential equation under next contour condition

$$\mathbf{g}(\mathbf{r}) = \nabla\rho(\mathbf{r}), \tag{71}$$

$$g_N(\mathbf{r}, 0) = f(\mathbf{r}_{N-1}), \tag{72}$$

$$\nabla \times \mathbf{g}(\mathbf{r}) = \mathbf{0}, \tag{73}$$

$$\nabla \cdot \mathbf{g}(\mathbf{r}) = 0. \tag{74}$$

Where an an harmonic potential $\rho(\mathbf{r})$, implies an harmonic vectorial field $\mathbf{g}$ with coordinates $g_j(\mathbf{r})$ for $j = 1, ..., N$. Conditions given by Eqs. (73) and (74) defines $g(\mathbf{r})$ as a conservative field and the energy is therefore conserved. Due to N is an arbitrary number the vectorial notation $\mathbf{r}_{N-1} = (x_1, ..., x_{N-1})$, has been used in Eq. (72). By taking Fourier transforms in (71), $\mathcal{F}\{g\} = \mathcal{F}\{\nabla\rho\}$, results

$$G(x_1, x_2, ..., u_j, ..., x_N) = 2\pi i u_j P(x_1, x_2, ..., u_j, ..., x_N) \tag{75}$$

where $G(x_1, x_2, ..., u_j, ..., x_N)$ and $P(x_1, x_2, ..., u_j, ..., x_N)$ are the Fourier transforms along dimension j of $g(\mathbf{r})$ and $\rho(\mathbf{r})$ respectively. The condition of harmonic potential, $\mathcal{F}\left\{\nabla^2\rho\right\} = 0$ derives in the new differential equation

$$-4\pi^2 q^2 P(u_1, ..., u_{N-1}, x_N) + \frac{\partial^2}{\partial x_N^2} P(u_1, ..., u_N, x_N) = 0 \tag{76}$$

being $q^2 = \|\mathbf{u}\| = \sum_{i=1}^{N-1} u_i^2$. The solution of Eq. (76) is

$$P = C(\mathbf{u}_{N-1}) e^{2\pi q x_N} \tag{77}$$

where $\mathbf{u}_{N-1} = (u_1, u_2, ..., u_{N-1})$ and $C(\mathbf{u}_{N-1})$ is an arbitrary function defined by initial conditions, that do not concern. Once the function P has been calculated, it will be introduced in the original Partial Differential Equation (PDE), Eq. (71); then by taking Fourier transform, the following set of N equation is obtained

$$G_j(\mathbf{u}_{N-1}, x_N) = 2\pi i u_j P(\mathbf{u}_{N-1}, x_N), \qquad j = 1, ..., N-1. \tag{78}$$

$$G_N(\mathbf{u}_{N-1}, x_N) = 2\pi q P(\mathbf{u}_{N-1}, x_N), \qquad j = N. \tag{79}$$

If Eq. (79) is introduced in (78),

$$G_j(\mathbf{u}_{N-1}, x_N) = i\frac{u_j}{q} G_N(\mathbf{u}_{N-1}, x_N) \tag{80}$$

Making now use of the contour condition Eq. (72), the limit $x_N \to 0$ will be performed, thus, Eqs. (80) is rewritten as it follows

$$G_j(\mathbf{u}_{N-1}, 0) = i\frac{u_j}{q} F(\mathbf{u}_{N-1}) \tag{81}$$

where $F(\mathbf{u}_{N-1}) = \mathcal{F}\{f(\mathbf{r}_{N-1})\}$. The set of Fourier transformations $G_j(\mathbf{u}_{N-1}, 0)$ with $j = 1, ..., N-1$ can be grouped as a

$$\mathbf{F}_R(\mathbf{u}_{N-1}) = (G_1(\mathbf{u}_{N-1}, 0), G_2(\mathbf{u}_{N-1}, 0), ..., G_{N-1}(\mathbf{u}_{N-1}, 0)). \tag{82}$$

The Riesz transform is thus defined in the Fourier domain:

$$\mathbf{F}_R(\mathbf{u}) = i\frac{\mathbf{u}}{\|\mathbf{u}\|} F(\mathbf{u}), \tag{83}$$

which can be expressed in the real space

$$\mathbf{f}_R(\mathbf{u}) = \mathcal{F}^{-1}\left\{i\frac{\mathbf{u}}{\|\mathbf{u}\|}\mathcal{F}\{f(\mathbf{r})\}\right\}. \tag{84}$$

being $\mathbf{f}_R(\mathbf{u}) = \mathcal{F}^{-1}(\mathbf{F}_R(\mathbf{u}))$ the Riesz transformation of a function $f(\mathbf{r})$.

5.12.1.2 *Monogenic Signal*

Using the Riesz transformation a generalization of the concept of analytic signal can be defined by means of a Clifford algebra of M dimensions. An arbitrary element z of this kind of associative algebra is represented as

$$z = z_0 + i_1 z_1 + i_2 z_2 + ... + i_M z_M, \tag{85}$$

where, $z_j \in \mathbb{R}$ for all j, and i_j for $j = 1, ...M$ are called imaginary units and verify next properties: $i_j^2 = -1$, $i_j i_k = -i_k i_j$. The particular cases where $M = 1$ and $M = 3$ derive the fields of complex numbers and quarternions.
Given an integrable real signal $f(\mathbf{r}$ with Riesz transform, $\mathbf{f_R}(\mathbf{r}) = (f_{R1}(\mathbf{r}), ..., f_{RN}(\mathbf{r}))$, the monogenic signal, $f_{MG}(\mathbf{r})$, is defined as

$$f_{MG} = f(\mathbf{r}) - \sum_{\alpha=1}^{N} i_\alpha f_{R\alpha}(\mathbf{r}). \tag{86}$$

This definition allows to define an amplitude $A_{MG}(\mathbf{r})$ as

$$A_{MG}(\mathbf{r}) = \sqrt{f(\mathbf{r})^2 + \sum_{\alpha=1}^{N} f_{R\alpha}(\mathbf{r})^2}. \tag{87}$$

The amplitude of the monogenic signal contains all the information about local energy in a particular point of the space $\mathbf{r}$.

6

LOCAL-DIRECTIONAL RESOLUTION AND ANISOTROPY MEASUREMENT

A complementary analysis to the resolution measurement is to determine the existence of anisotropies. The concept is simple, it aims to elucidate if the obtained structure presents directions in which the quality is better than in others. That is the case of tomograms: they exhibit better resolution in the XY plane than along the Z axis as a consequence of the missing wedge produced in their measurement procedure. The measurement of the resolution anisotropy was recently introduced in electron microscopy, with the simultaneous and independent publications of two methods addressing this issue as they are the 3DFSC [172] and the efficiency [111]. The first one consists on a directional FSC, which is the natural extension to the standard FSC. The difference between them is the use of a cone along the direction under study in the Fourier domain to compute the cross correlation between two half volumes, i.e. it represents a self-consistency measurement. When all the directions in the projection sphere are considered, it is possible to construct a 3DFSC. Using the 3DFSC, Tan et. al, [172] defines the sphericity parameter to determine the degree of anisotropy of a volume. However, this approach is global, leaving unsolved possible local anisotropies. The second method measures a PSF from the fall of the signal power along all the possible directions of the projection sphere. Then, it computes the efficiency parameter to model the shape of the calculated PSF.

Hence, both measurements are global and leave unsolved the existence of local anisotropies. Considering this, a new method named *MonoDir* was developed to measure the local anisotropy and provide local directional resolution maps. The core of this method is *MonoRes* algorithm, in which the high pass filtered is replaced by a high pass directional filter via Fourier cones. The result is a map of local resolutions along the selected direction. When this approach is carried out covering all the directions in the projection sphere, the directional resolution information of each voxel is obtained, and then, the local resolution anisotropy can be determined. In particular, it is reported as the interquartile range of the distribution of directional resolutions for each voxel. Surprisingly, the local anisotropy is less relevant than the information that can be derived from directional local resolution measurements. Hence, it is possible to analyze the quality of the reconstruction via a set of indicators as they are: the Average Directional Resolution map, the local directional interquartile resolution dispersion map, the radial average resolution and the highest resolution angular plot. In this chapter, these indicators will be carefully explained. Briefly and in advance, it is essential to remark that they allow to infer the existence of preferred directions and bad angular coverage of the projection sphere, as well as angular assignment errors. Note that *MonoDir* provides very enriched information about quality reconstruction, without using the set of particles, it means with the only knowledge of the reconstructed map. In addition,

to our knowledge, *MonoDir* is the first local directional resolution method in the field.

6.1 ON THE ANISOTROPY AND RESOLUTION BY DIRECTION

Firstly, the term isotropy has to be defined. The term makes reference to a geometrical property or a transformation which is directionally preserved with independence of the direction. Mathematically, it can be defined as an invariant property or feature under the rotation group ($SO(2)$ for images or $SO(3)$ for volumes). When an image/volume lacks of isotropy, it is called anisotropic, i.e. the studied property changes in at least two different directions.
The chosen property to analyze its anisotropy was the resolution. However, it was shown in the Chapter 5 that it can be measured via two approaches: global (FSC-SSNR) or local (*MonoRes*, *ResMap* or *blocres*). Similarly to the resolution, the 3DFSC [172] and the efficiency[111] are global resolution anisotropy metrics that inform about the coverage of the projection sphere. Here, a local anisotropy measurement is introduced with the *Monodir* algorithm.
Leaving out how to calculate it, resolution by direction means that, it is possible to assign unidimensional sinusoidal waves along the analyzed direction, with wavelength up to the directional resolution value. The set of all resolutions by direction can be represented by the named *resolution surface*, where the distance from a point to the center of the resolution surface is the resolution value along the direction defined by these two points. Because of the symmetry of the projection sphere, this surface must be symmetric and its shape will determine the degree of anisotropy. Resolution will be isotropic if the resolution value is directionally invariant. The anisotropy are then the deviations of this invariance. This means that when a volume is isotropic, the directional resolution surface will be a sphere. Therefore, to measure the anisotropy requires estimating the deviation to this sphere i.e. to characterize the resolution surface shape. Note that this interpretation of anisotropy is valid in a global or a local sense. Thus, the global approaches as the 3DFSC [172] and the efficiency[111] define a surface from directional measurements, and then characterize the shape of the surface as anisotropy measurement. In the local approach that *Monodir* algorithm provides occurs exactly the same, but in a local manner.

6.1.1 *Anisotropy causes*

There are many reasons causing an anisotropic reconstructed volume, however the common motif of them is the lack of information at certain frequencies in a particular direction. A clear example is a random conical tilt volume or a tomogram. They have no information in the missing cone/wedge and the result is an elongated volume along the missing direction and therefore an anisotropy of resolution in such a direction. Other causes are shared with local resolution as it is the case of preferred directions. Even when the angular projection space is well covered some directions can be better resolved due to the quality or the number of projections in

that direction. The last cause is the existence of heterogeneity or mobile element, they introduce a blurring effect, and therefore a low resolution region.

6.2 MEASURING LOCAL DIRECTIONAL RESOLUTION AND RESOLUTION ANISOTROPY - MONODIR

To characterize the local resolution anisotropy it is necessary to analyze the local resolution variation along the all possible directions. The knowledge of this local resolution variation allows to extract local information and from them infer global characteristics about the density map.
The basis of *MonoDir* is *MonoRes*, with the introduction of addional directional filters. As a consequence, the input will be the original electron density map and a mask. The mask establishes a frontier allowing to determine the distributions of noise and signal (particle), which are used in *MonoRes* to assign the resolution values. The algorithm starts uniformly covering the angular projection sphere, and for each direction *MonoRes* is applied with a directional high pass filter along the chosen direction. The result is a map of local resolution by direction. Once all directions have been analyzed, a set of resolution maps as large as the number of directions is obtained. Then, all these maps are analyzed to determine the highest and lowest local-directional maps, the Average Directional Resolution map and its local resolution interquartile range map (local dispersion map), the angular plot and the radial average plot. The meaning of these outputs will be explained below. The details of the algorithm are explained step by step as follows, and in Fig. 39 a scheme is shown.

1. *Generate directions*: The projection sphere is uniformly discretized in a set of N directions. This is a well known problem named point picking. It presents many solutions, the chosen one was considers a dome built from an icosahedron [12].

2. *Directional filter*: For each angular direction, the map is directionally filtered. This kind of filter is defined in the Fourier space as a cone such as only the frequencies inside it are kept, the rest are removed. The cone axis is defined by the filtering direction, and a cone angle 15 degrees, see Fig. 39 (the cone angle was empirically set). When the original density map is directionally filtered, the result will be a fringe pattern oriented along the perpendicular direction to filtering one. Note that this is the concept of Fourier transform, a decomposition of a function as a sum of independent waves with different frequency.

3. *MonoRes - Local directional Map*: In this step *MonoRes* is used to determine the local resolution of the directional filtered map. Unfortunately, this task is not so simple and *MonoRes* must be considerably adapted to deal with two problems that appears in the monogenic amplitude

    a) *Gibbs effect*: This is a well known effect that occurs when the Fourier transform of a function with a non-evitable discontinuity is performed. It is characterized by a high frequency ripple at the extremes of the

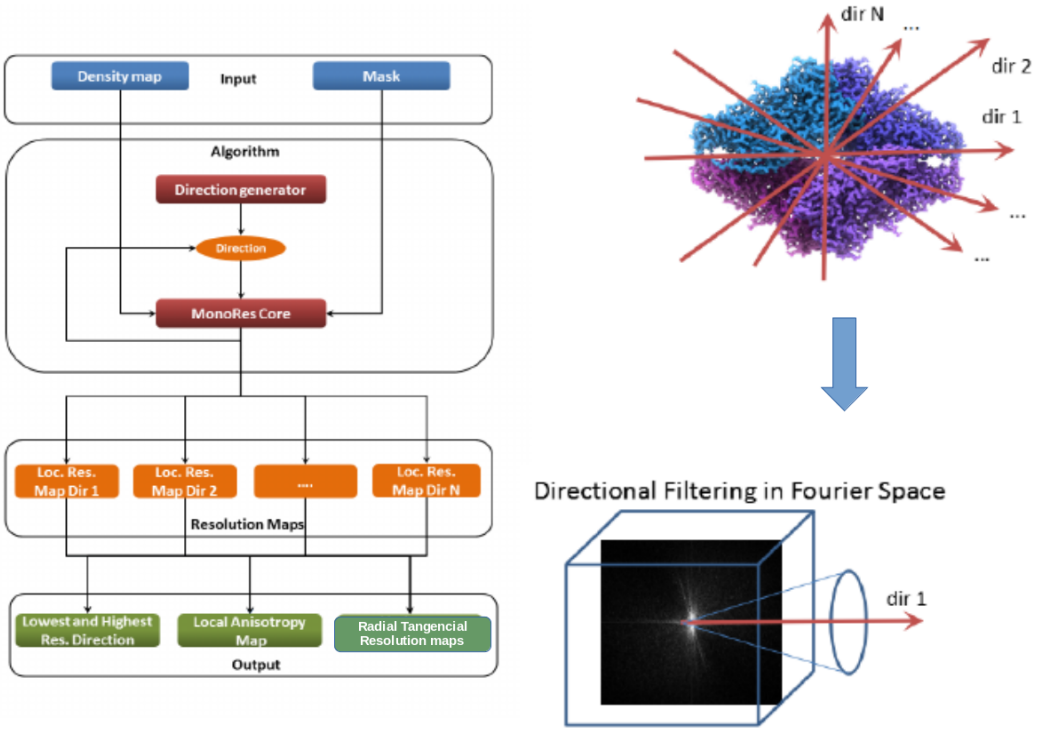

Figure 39: *MonoDir* scheme. A set of directions are generated uniformly covering the projection sphere, the map is then directionally filtered, and by means of *MonoRes*, its local resolution is calculated. When all directions are computed, the result is a set of local resolution maps that allows to measure local resolution anisotropy, angular alignment errors and the existence of preferred directions.

function domain. To avoid it, the monogenic amplitude is masked with a spherical mask smoothed by a raised cosine, this task is performed before applying the low pass filter, see *MonoRes* workflow in Chapter 5.

b) *Noise anisotropy*: The mask established a border between signal and noise. However, the effect of directional filtering might introduce anomalies in the noise region that perturb the noise estimation. To avoid them, the noise is determined in the intersection of a cone with axis along the filtering direction and a spherical crown defined by the largest macromolecular radius and the half of the box size. See Fig. 40.

This fine modification of *MonoRes* to determine local-directional resolution maps is critical to properly estimate the resolution.

Points 2 and 3 are repeated until all the generated directions in the Step 1 are analyzed. The result is a set of local resolution by direction maps as large as the number of directions (Note, that at the end a resolution tensor is calculated). This set contains a lot of information about the original volume, that will be analyzed in next sections.

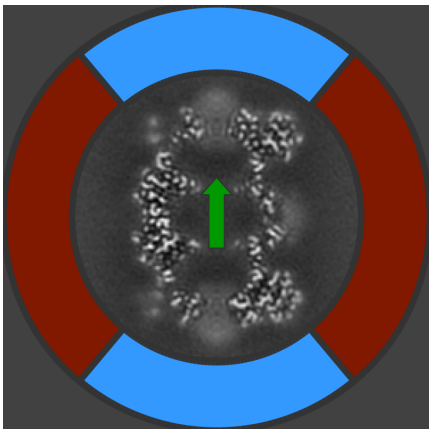

Figure 40: Region considered for noise estimation. The filtering direction is represented by the green arrow. In blue the noise estimation region, it is defined as the the intersection of a cone along the filtering direction with a spherical shell. In red appears the part of the spherical shell that is neglected in the noise estimation.

6.3 LOCAL INFORMATION PROVIDED BY MONODIR

As a consequence of the local-directional analysis a set of local maps are calculated:

1. *Local highest and lowest resolution maps*: For all voxels their lowest and highest resolution can be determined, producing two output volumes: the local highest and lowest resolution maps. These maps establish a resolution bound and loose the directionality, note that given two voxel their highest/lowest resolutions can be achieved along different directions. It must be highlighted that for a given voxel, the highest and lowest local-directional resolution values are determined in an statistically robust manner, it means taking the percentiles 0.05 and 0.95 of the local-directional resolution distribution.

2. *Radial and tangential local resolution maps*: Assuming a reference system in the center of the map, the position of each voxel is defined by three coordinates (i, j, k). If spherical coordinates are considered, then, the vector (i, j, k) determines the radial direction. Thus, for each voxel a radial and tangential decomposition can be carried out. We define the radial resolution of the voxel (i, j, k) as the local-directional resolution of that voxel along the direction (i, j, k). Similarly, we define the tangential local-directional resolution as the average of the direction in the orthogonal plane to the direction (i, j, k). Note that the resolution of these maps must be between the highest and lowest local-directional resolution volumes.

3. *Average Directional Resolution map*: A voxel located in $\mathbf{r}$ will present a directional-resolution distribution as a consequence of measuring its resolution along all possible directions. The Average Directional Resolution map is defined as

$$ADR(\mathbf{r}) = \frac{L_r(\mathbf{r}) + H_r(\mathbf{r})}{2}, \tag{88}$$

where the subindex r specifies the highest and lowest resolution of the voxel $\mathbf{r}$ considering robust statistics, i.e. $H_r(\mathbf{r})$ and $L_r(\mathbf{r})$ are the percentiles 0.05 and 0.95 the directional-resolution distribution. For a perfect reconstruction, $ADR = H_r(\mathbf{r}) = L_r(\mathbf{r})$, and therefore the structure will be isotropic. The existence of anisotropy will increase the ADR value with respect to $L_r(\mathbf{r})$. Ideally, the ADR should be equal to the local resolution estimated by *MonoRes*.

4. *Half Inter-quartile Range Map*: This map provides a complementary information to the ADR map measuring the ADR dispersion. For a given voxel and its directional-resolution distribution the half inter-quartile range is defined as the half distance between the percentiles 0.17 and 0.83. Thus, when the directional-resolution distribution is normal, the half inter-quartile range is equivalent to the standard deviation. A good reconstruction should present a inter-quartile range close to zero, large deviations from zero put the reliability of the volume in question.

6.4 GLOBAL INFORMATION PROVIDED BY MONODIR

The local analysis also allows to infer the existence of global anisotropy. Note that in general, every global measurement can be understood as a collective local measurements in a statistical sense. The set of all local directional resolutions will define global distributions and the expected value usually coincides with the global measurement. This is exactly what *MonoDir* does to predict orientation errors. In particular, the following global properties are can be identified (see result section figures to fully understand the plots)

1. *Orientation errors*: To identify the existence of angular assignment errors, the radial average of the local-directional resolution maps are performed. The loss of resolution as a consequence of angular errors must increase as the radius increase. When the radial averages exhibit a non-zero slope, then angular errors are presented. Nevertheless, if the map is perfectly reconstructed, this loss of resolution is not present, and the radial averages show a flat profile. Finally, if a good angular assignment is performed but there are present translation alignment errors, the whole radial average profile is shifted up resulting in an homogeneous loss of resolution. To do this, any local-directional resolution map can be used. However, since the radial and tangential directions define a simple basis, they were chosen to be analyzed.
   Moreover, looking at the radial average curve it is observed that radial and tangential diverge at some point that we called, divergence radius r_d, defined as follows. If $\bar{n}_{\theta,\phi,r}$ is the average the number of voxels along a direction (θ, ϕ) (the pair denotes the latitude and longitude on the sphere) inside a sphere of radius r, then, r_d is defined as the highest radius of the macromolecule such as $\bar{n}_{\theta,\phi,r} = \bar{n}_{\theta',\phi',r}$ for $\theta \neq \theta' \; \phi \neq \phi'$. It implies that the slope of the radial averages should be measured for $r < r_d$. This phenomenon is also responsible of the sudden change in the slope of radius close to the border of the macromolecule.

2. *Existence of preferred directions*: With the only knowledge of the reconstruction, it is complicated to infer the existence of preferred directions. However, they have an impact on the reconstruction and the resolution by direction. A given voxel will take its highest resolution along a direction defined by two angles (θ, ϕ). Hence, a histogram on a polar plot can be used to represent the number of voxels that have their highest resolution along a given direction (θ, ϕ). When the polar plot shows a uniform coverage of the projection sphere, then the reconstruction is globally isotropic, otherwise, the sample may have preferred directions resulting in a global anisotropy.

6.5 MONODIR RESULTS:

The *MonoDir* was applied under different scenarios considering: (1) synthetic data to validate and measure the accuracy of the method. (2) Experimental data with the aim of evaluate the performance of the algorithm in real maps taken from EMDB [95].

6.5.1 *Tests with synthetic data*

6.5.1.1 *Test example: Introducing orientation errors*

This test is addressed to understand how the directional resolution is affected by errors in the angular assignment or the absence of information in the projection sphere. To do that, the atomic model of β-galactosidase (PDB entry 3j7h [11]) was taken and converted into density map using `xmipp_volume_from_pdb` [161] with a sampling rate of 1Å/pixel. The map was uniformly projected to obtain 500 projections and then Gaussian noise with zero mean and standard deviation of 2 a.u. was added to each projection. Thus, it is possible to simulate particle with correct and known angular assignment. Using this set of particles the following reconstructions were carried out:

1. *Error-free map*: The map is reconstructed with the correct angular assignment. This reconstruction should be perfect. *MonoDir* was used to check the quality of the map. The results can be seen in Fig. 41. First, it is noted that the ADR map resolution is very homogeneous. The angular plot is well covered and the radial resolution averages curves are completely flat (leaving the extremes out). Thus, it can be concluded that this map presents a high quality, as was expected because it is a synthetic map without errors.

2. *Missing directions*: All particle with tilt angle smaller than 40 degrees were removed from the data set. Then, a map was reconstructed with this new set of particles. Note that the experiment simulates a random conical tilt reconstruction, and the structure will be affected by a missing cone of 40 degrees, and resolution anisotropy. *MonoDir* was used to analyze this map. The result can be observed in Fig. 41. It is recommended to compare them with the ones obtained in the case of the error-free map (also in Fig. 41). Firstly, the existence of the missing cone is reported by *MonoDir* in the angular plot, in

which a bold area of exactly 40 degrees is shown in the center of the plot. It is remarkable the difference in the ADR map between this case for which the ADR shows very low resolution and the error-free map in which high resolution values are shown. Finally, because of the angular orientations of each particle is kept, the radial average of local resolution maps show present flat curves, as reported by *MonoDir*, but the effect of the missing cone shifts the curves towards low resolutions.

3. *Angular errors*: In this test, the orientation of each particle was randomly modified by introducing a random Gaussian error with zero mean and 1.2 degrees of standard deviation and then a map was reconstructed. Because each direction of the projection sphere is defined by two angles, possible combinations of angular errors were tested casting similar results. In Fig. 41 the *MonoDir* results are shown. The angular plots were similar to the obtained for the error-free map, i.e. they present a good coverage of the projection sphere. It was expected because all directions are presented with approximately the same number of particles. In contrast, the radial average plot shows a slope, which is particularly clear for the radial and tangential local directional resolution maps. It was verified, that the slope of these curve increases as the introduced error is increased. This experiment allows a method for identifying angular assignment error without no more information than the reconstructed map.

4. *Shift errors*: Similarly to the previous test, a shift Gaussian error (with zero mean and standard deviation of 1.2 pixels) in the orientation assignment of each particle was introduced keeping its angular orientation and then a map was reconstructed. In Fig. 41 the *MonoDir* results are shown. The angular plot is omitted because it is essentially the same one obtained by the error-free map. The radial average of radial and local directional tangential resolutions cast a flat profile, but the resolution gets lower than the obtained by the error-free map, as a consequence of the shift errors.

6.5.2 *Test with experimental data*

Once *MonoDir* was tested to ensure the reliability of the directional and anisotropy measurements, it was applied to analyze the quality of experimental maps deposited in EMDB [95]. In concrete, the following maps were studied: the Proteosome 20S (EMD-6287) [20], the Ribosome 80S (EMD-2275) [8], and two maps of Influenza Hemagglutinin (HA) trimer [173] taken from EMPIAR (entries 10196, 10197). These maps have been deeply analyzed, thus Proteosome 20S and the Ribosome 80S are good examples of very good and problematic angular assignment respectively [111]. Finally, the Hemagglutinin (HA) trimer maps are excellent examples about how *MonoDir* can be used to identified reconstruction in a case for which the same macromolecule presents two reconstructions with different angular assignment coverage.

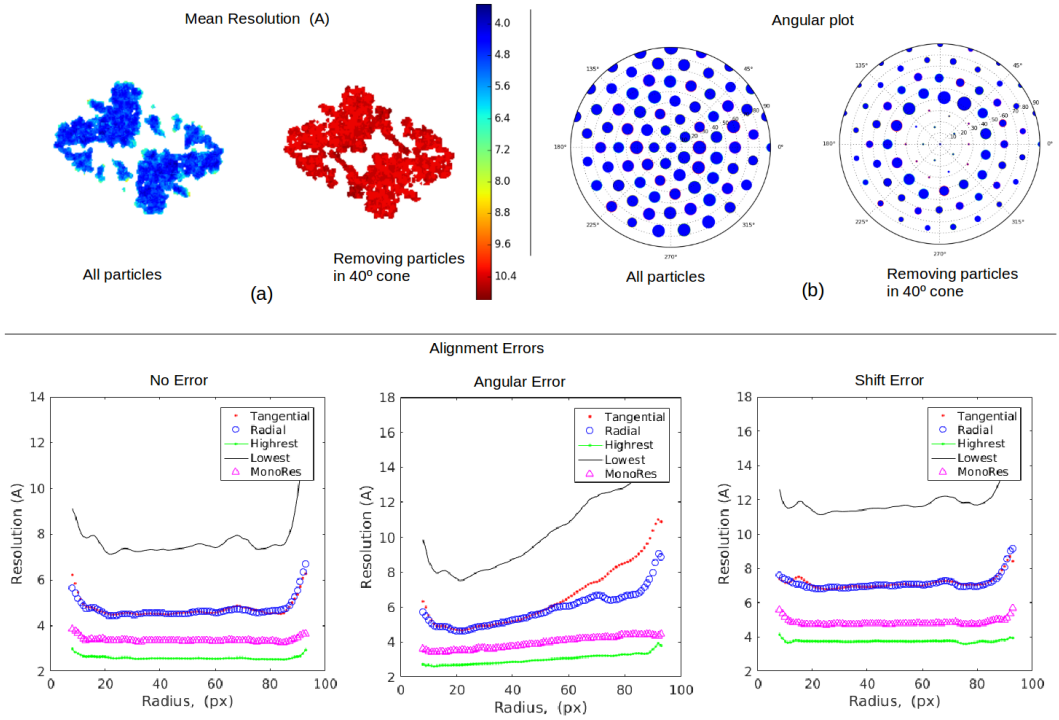

Figure 41: *MonoDir* results for the β-galactosidase obtained from synthetic particles: (a) ADR map for two reconstructions first with an uniform coverage of the projection sphere and then removing particles, (b) their respective angular plots, and (c) radial average of local resolution maps when random errors are introduced in particle orientation (in Euler angles and shift) in all cases considering a uniform coverage of the projection sphere.

.

6.5.2.1 *Experimental Map 1: Proteasome 20S*

The first experimental case makes use of the experimental map of the Proteasome 20S taken from EMDB (EMDB entry-6287)[20]. Note that it was previously analyzed with *MonoRes* in Chapter 5. Remind that its reported resolution was 2.8 Å (measured as a FSC gold standard at 0.143).

This map is a well known example of reconstruction with good angular assignment [111]. *MonoDir* was applied to check it, the results can be observed in Fig. 42. The local conclusions provided by *MonoDir*: the highest and lowest local-directional resolution maps present values very close each other, in the range [2.2,4.9] Å. The ADR values were between [2.7,3.9] Å, a range close to the interval of [2.0, 4.6] Å reported by *MonoRes* (see Fig. 30). In addition, the half interquartile range map shows a very low resolution dispersion. As a consequence, the results imply a high isotropy in terms of local directional resolution. The global conclusions of *MonoDir* are extracted from the polar plot and the radial average curves. The polar plot points out a good coverage of the projection sphere in the center of the plot with a worse coverage of top views. It implies the existence of preferred directions (side view). This conclusion is in agreement with the reality, most of particles are side views [111]. Note that it is possible to reconstruct only with side views be-

cause the Fourier space is fully covered. Finally, the radial average plot casts very flat curves, it is an indicator of a good angular assignment, in particular, for the radial and tangential curves.
The 3DFSC that provides information about global resolution anisotropy, was also calculated. The results can be seen in Fig. 45. They are in agreement to the ones obtained by *MonoDir*. In particular, it is noteworthy the very low dispersion that the resolution presents.

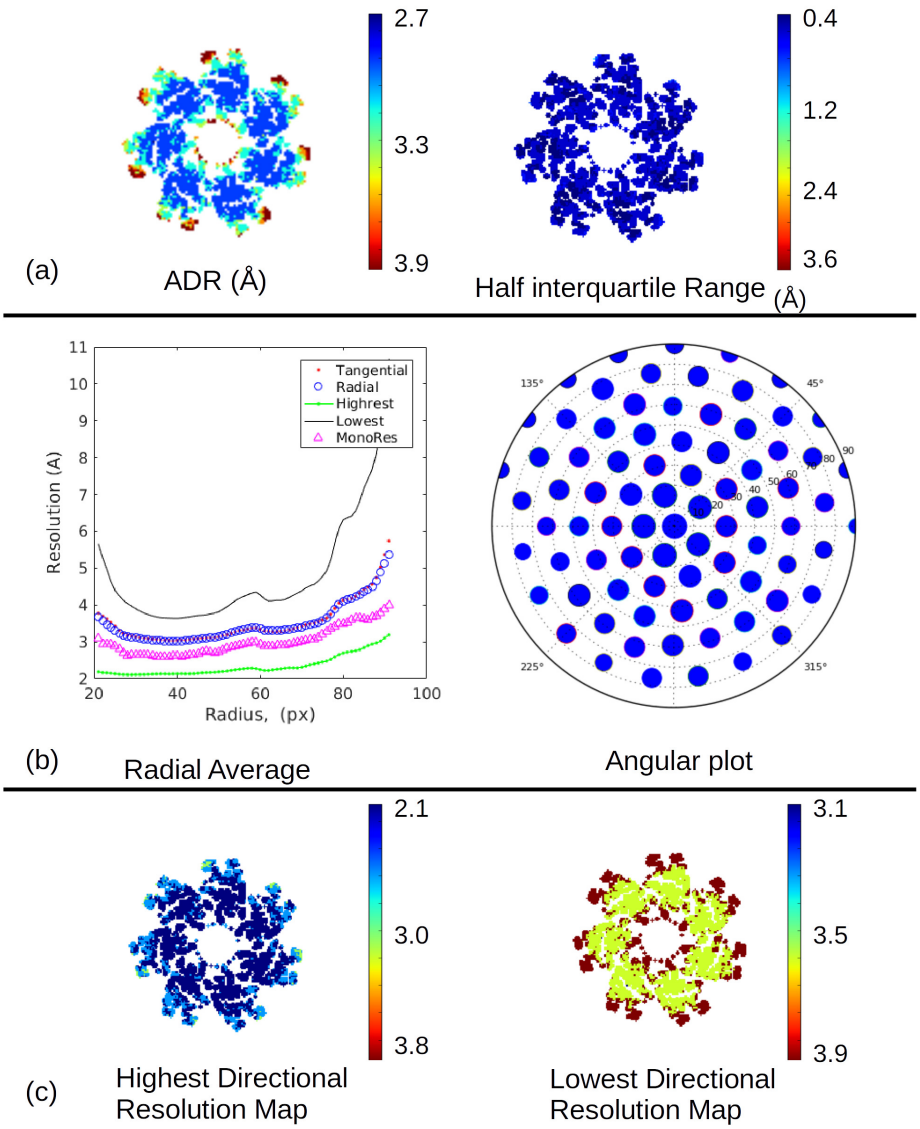

Figure 42: Results for Proteasome 20S: (a) Average Directional Resolution (ADR) and its half interquartile range, (b) radial average of local tangential (pink points) and radial (blue circles) directions plotted along radii, together with the highest, lowest and MonoRes local resolution estimations and polar plot, (c) highest and lowest local directional resolution maps.

6.5.2.2 *Experimental Map 2: Ribosome 80S*

The second experimental case considers the experimental map of the Ribosome 80S taken from EMDB (entry-2275) [8]. The map has a pixel size of 1.22 Å and box dimensions of $240 \times 240 \times 240$ voxels with reported resolution FSC gold standard resolution at 0.143 of 4.5 Å.
Despite the map allows a good understanding of the macromolecule, it was shown that it had problems in the angular assignment and the coverage of the projection sphere [111]. It is highlighted that this ribosome is assembled and presents heterogeneity, which complicates the alignment and decreases the resolution. *MonoDir* was applied to estimate the map quality, determine the existence of resolution anisotropy, and check the angular assignment. In Fig. 43 the results are shown. In this case the measurement of highest an lowest local-directional resolution map is omitted, the ADR map provides enough information with values in the range of [2.7,3.9] Å and the half interquartile range shows dispersion from 1 to 6 Å, which indicates a high variability in the directional resolution. In contrast, when global parameters are analyzed, the angular plot shows a non-uniform coverage of the projection sphere, this fact is in agreement with previous studies of this map [111]. Finally, the radial average curves present a high slope and therefore, there is a problem in the angular assignment. However, the identification of this error allows to solve it by assigning angles again up to achieve the flatter the better radial average curve.

6.5.2.3 *Experimental Map 3: Influenza Hemagglutinin (HA) trimer*

The third experimental example considers the structure of the Influenza Hemagglutinin (HA) trimer. This case is of special interest for *MonoDir*, because there are two reported maps of this structures acquired under different acquisition conditions with a tilt sample of 0 and 40 degrees and a pixel size of 1.49 Å. These data can be found in the EMPIAR entries 10196, 10197 [173]. When the sample is untilted, the sample presents preferred directions and the reconstruction with this dataset presents anisotropy and a bad coverage of the projection sphere. This effect can be alleviated by tilting the sample, it was experimentally checked that a tilt of 40 degrees was enough to that purpose [173]. The result was a better angular assignment and coverage of the projection sphere. To check this two facts, the 3DFSC was calculated and it is shown in Fig. 45. Note how the dispersion in FSC curve is reduced when the sample is tilted.
These hypotheses were confirmed by *MonoDir*, the results can be found in Fig. 44. Both maps present similar ADR maps, it is noted that the tilted map cast slightly lower ADR values. However, the ADR map must be complemented with the half interquartile range map, that clears up the debate about which map is better. The untilted map shows higher directional resolution dispersion than the tilted one. This fact support a better quality but is not the only one. The global results are in agreement with this conclusion, the angular plot, shows differences in the coverage of the projection sphere, the map obtained by tilting the sample shows a better coverage of the angular plot than the map obtained at 0^{o} tilt. It means that by tilting the sample the global anisotropy is reduced. Finally, the radial average plot

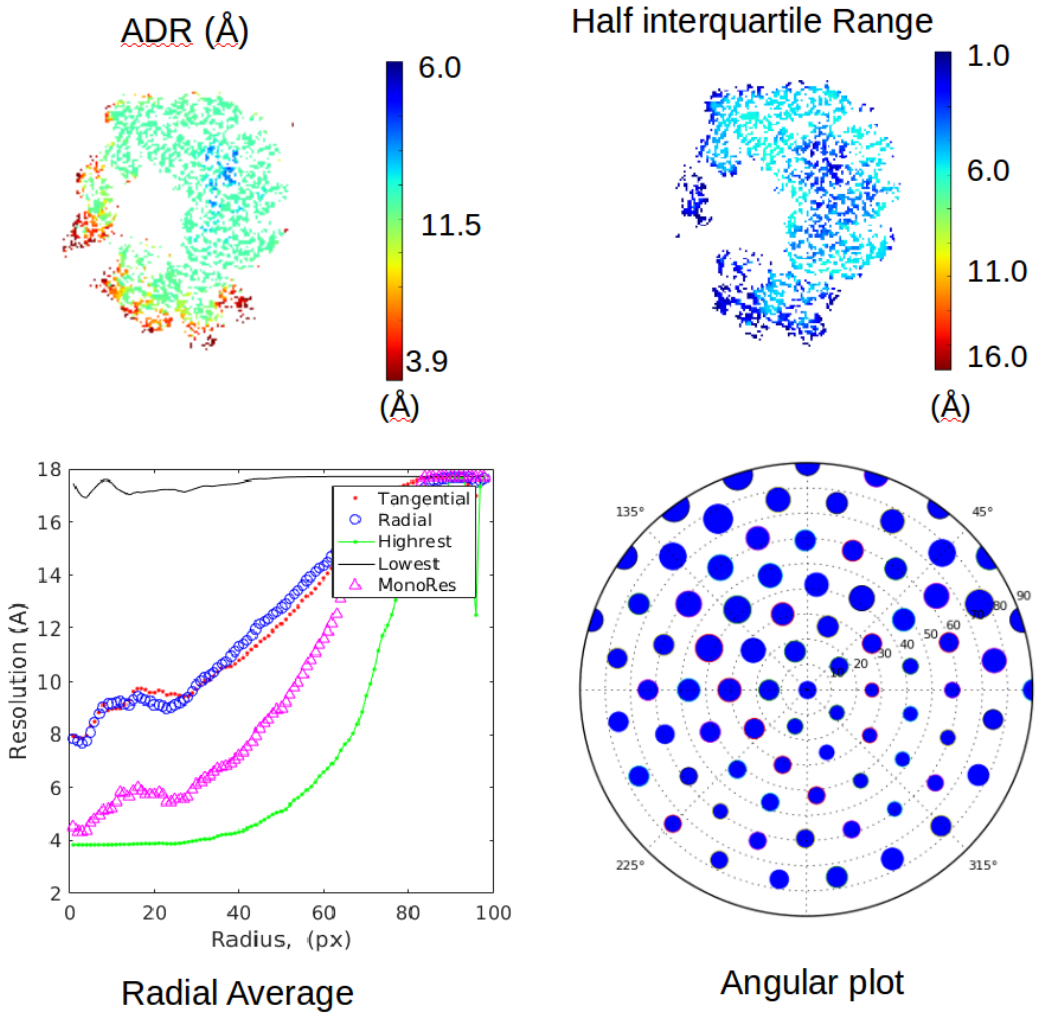

Figure 43: Results for Ribosome 80S: (a) Average Directional Resolution (ADR) and its half interquartile range, (b) radial average of local tangential (pink points) and radial (blue circles) directions plotted along radii, together with the highest, lowest and MonoRes local resolution estimations and the polar plot.

to indicates the existence of angular alignment errors, showing that the map reconstructed at 0° tilt presents higher slopes in the radial average resolution curves, in contrast to the map obtained at 40° tilt with flatter curves.

6.6 DISCUSSION OF MONODIR RESULTS

The measurement of local resolution has been expanded by introducing the directionality concept to measure local-directional resolution map. In this sense, resolution results in a tensor rather than a local value. Local resolution is the combination of many factors that were commented in Chapter 5. Unfortunately, it is not possible to isolate them. The method that was proposed in this chapter is an attempt to do it in a local manner. When local-directional resolution is measured along set of directions that uniformly covers the projection sphere, then very rich infor-

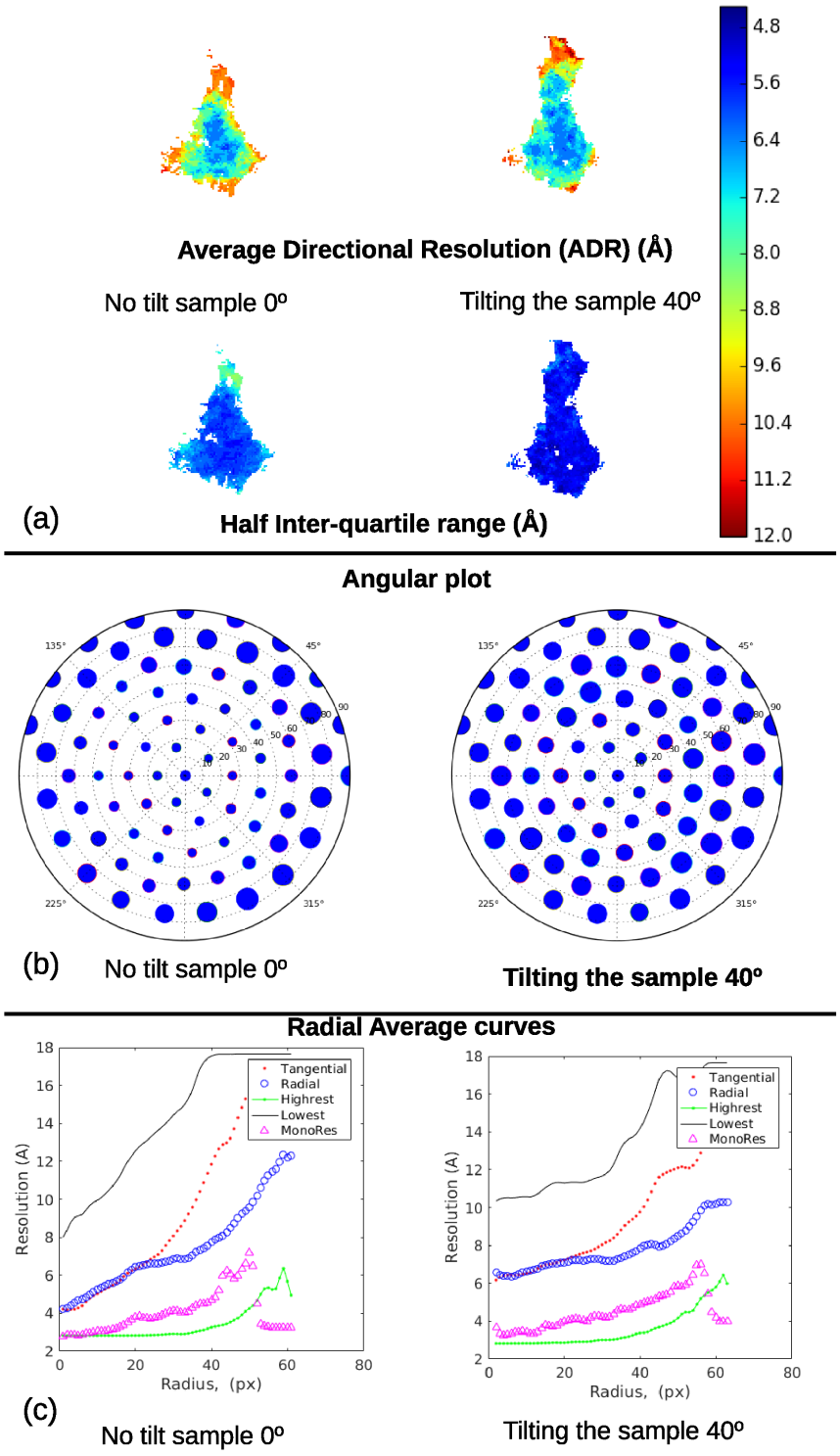

Figure 44: Results for Influenza Hemagglutinin (HA) trimer when the sample is untilted and tilted: (a) ADR and the half interquartile range, (b) angular plot, and (c) radial averages.

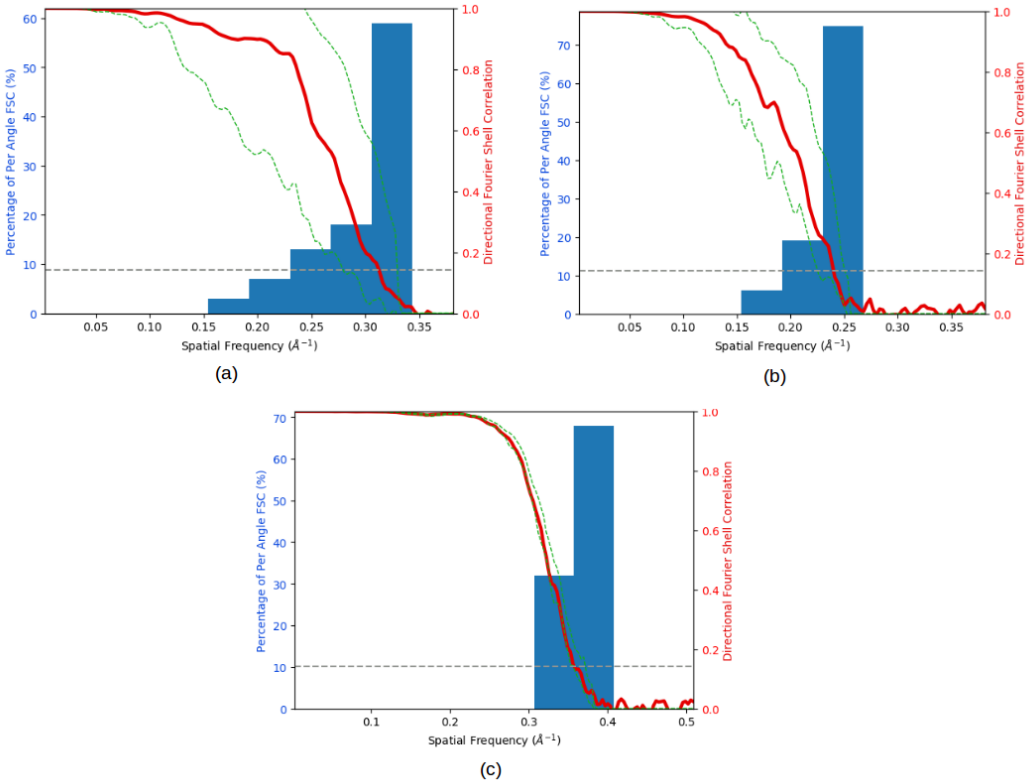

Figure 45: Results of the 3DFSC method for Influenza Hemagglutinin (HA) trimer when the sample is (a) untilted, (b) tilted. Results of the 3DFSC method for the Proteasome 20S.

mation about the reconstruction process can be extracted. To analyze that, a new algorithm named *MonoDir* was proposed. Up to our knowledge, it is the first local-directional resolution algorithm in cryoEM. Hence, with no more information than the reconstruction, *MonoDir* is able to determine the existence of preferred directions, global anisotropies, angular assignment errors, and local-directional resolution anisotropies. This measures are only a few possible applications. As a pioneer method, it is expected that it opens new horizons in the field. Thus, the measurement of local-anisotropy might be used to perform local-anisopropic sharpening or filters, map validation, or chain tracing among others.

Results with experimental data confirm the performance of *MonoDir* as validation map quality method. The examples with synthetic data illustrate how the angular alignment errors leave a fingerprint in the resolution. Hence, thanks to them it was possible to identify angular assignment errors in experimental data sets. The ribosome and the reconstruction of the Influenza hemagglutinin trimer without tilting the sample are two good experimental confirmations. The synthetic example in which some particles were removed to artificially produce a missing cone helps in the understanding of the maximum resolution angular plot. In that case, the missing cone can be clearly observed. Similarly to what happen with the Influenza hemagglutinin trimer, in which the existence of preferred direction can be clearly observed. In addition, the introduced error in the synthetic map are taken into

account by the ADR map. In particular, the loss of resolution is specially significant in the case of the missing cone, that casts a very low resolution ADR map.

6.7 CONCLUSIONS

1. The concept of local resolution has been generalized adding directionality, resolution is then a tensor instead of a scalar.
2. A fully algorithm, *MonoDir*, for computing the local-directional resolution and existence of anisotropy in SPA reconstructions has been proposed.
3. The resolution anisotropy informs about the local and global map quality with no more information than the reconstructed map.
4. Angular assignment error can be identified as a slope in the radial average of the local resolution maps, in particular in the radial and tangential.
5. Existence of preferred directions and global anisotropy are shown as a non uniform coverage of the angular plot.

7 ON THE B-FACTOR AND ITS CORRECTION

This chapter should be understood as a transition chapter. In Chapter 5 the estimation of local resolution maps in SPA was exposed. In the following chapter it will be shown the *LocalDeBlur* algorithm that by means of local resolution information, it allows to enhance locally the visualization of the reconstructed density map as a post-processing step, this enhancement is known as sharpening. However, before introducing the sharpening algorithm proposed in this thesis, the current state of the art will be exposed.
The tendency nowadays is to sharpen the reconstructed density map using a kind of transformation called B-factor correction. This transformation is applied to the structure to alleviate the low of contrast (in particular at high frequency) as a consequence of many factors, for example: heterogeneity, flexibility, beam induced movement, inelastic scattering, charging or radiation damage among others. To do that, what B-factor correction does is a boosting of high frequencies by flattening the spectrum of the reconstructed structure.
Despite the good results that B-factor correction provides, it will be shown that its use is not supported by the scattering theory. Hence, the aim of this chapter is to provide a clear, theoretical and experimental proofs about why the B-factor correction should be revisited. To do that the scattering theory is considered obtaining the called Guinier law for which its validity range and connection with the B-factor flattening will be analyzed. To be more specific, this chapter can be summarized in the following items:

1. The Guinier approximation is valid only at very low frequency and it is basically related to the overall shape of the macromolecule.

2. The amplitude spectrum of a macromolecule has, in general, a non-flat slope at all frequency ranges of interest.

3. 3D reconstruction algorithms do not always over-dampen the spectrum of the macromolecule.

In the following chapter an alternative to B-factor correction based on local resolution will be proposed. Because of the theoretical nature of this chapter, many similarities can be found in our own publication [190]. However, here they are deeply explained adding many proofs to support the theoretical results which are out of that publication.

7.1 GLOBAL SHARPENING. B-FACTOR CORRECTION

The sharpening technique known as B-factor correction is a global transformation that has its origin with the publication of P.B. Rosenthal and R. Henderson [135]. Usually, the high frequency information is hidden by the low frequency terms, the

idea is to boost the high frequencies to get better the map visualization, but avoiding the noise amplification.
The method begins by calculating the structure factor of the macromolecular complex. The structure factor will be rigorously defined in the following section, but it essentially represents the radial average of the power spectra of the macromolecule. Thus, it is assumed a Gaussian decay of the Fourier coefficients with the frequency given by

$$|F(q)|^2 = |F(0)|^2 e^{-\frac{1}{3}R_g^2 q^2}. \tag{89}$$

where, $F(q)$ represents the structure factor at the frequency q, and R_g the gyration radius (it will be derived in next section). The exponential term $e^{-\frac{1}{3}R_g^2 q^2}$ is responsible of reducing the contrast at high frequencies and therefore it complicates the interpretability of the map. What B-factor correction does, it to remove this exponential term by making a flat spectra. It is achieved by linearizing Eq. (89) as it follows

$$\log|F(q)|^2 = \log|F(0)|^2 - \frac{1}{3}R_g^2 q^2. \tag{90}$$

Then, the slope of the fitted line, $B = -\frac{1}{3}R_g^2$ is estimated to multiply the structure factor by $C(q)e^{Bq^2}$, where $C(q)$ is a weight factor explained below. It pretends to be a flat structure factor i.e. the decay term, $e^{-\frac{1}{3}R_g^2 q^2}$, is removed. Regarding to the frequency range in which the linear fitting is fulfilled, the most spread criterion considers a the fitting in the interval, $[FSC, 10-15]$ Å, see [41]. The weight, $C(q)$, is responsible of avoiding the noise amplification, by considering the SNR at the given frequency, or alternatively the FSC as

$$C(q) = \frac{2FSC}{1+FSC}. \tag{91}$$

7.2 SCATTERING THEORY

Consider an electron beam lighting a macromolecule composed by N atoms. The beam direction is defined by the the unitary vector $\mathbf{s}_0$ and incident electrons will be scattered as a consequence of their interaction with the charge density of the macromolecule. Let scattering direction be defined by other unitary vector $\mathbf{s}$, which forms an angle 2θ with respect to the incident direction, $\mathbf{s}_0$. In Fig. 46 a simplified scheme of the experiment is shown. This scenario represents the electron-sample interaction inside the electron microscope. Thus, the objective will be to calculate the intensity of the scattered beam which is the measured.

The scattering process is a consequence of the electron matter interaction. Thus, each atom of the macromolecule can be understood as an scatterer element. Hence, the macromolecule can be described by its electronic density $\rho(\mathbf{r})$. Therefore, the problem turns into calculating the intensity of the scattered beam by a set of N small scatterer elements defined by their positions $\mathbf{r}_k$, with $k = 1, ..., N$. The electron-matter interaction mainly depends on the atomic number of atoms in the

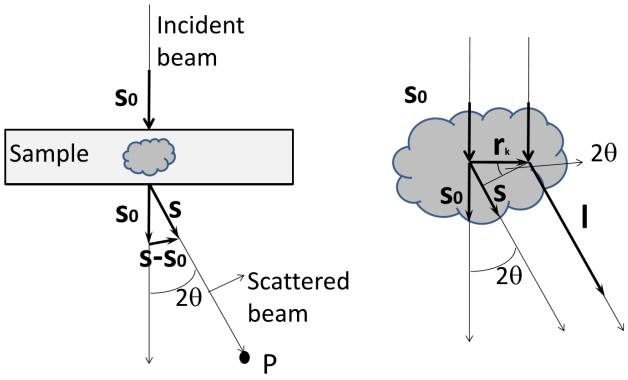

Figure 46: Scheme of the system under study. An incident electron beam is scattered as a consequence of the interaction with an scatterer element. The incident beam with direction $\mathbf{s}_0$ and the scattered one with direction, $\mathbf{s}$, forms an angle 2θ .

sample, sample thickness, and the energy of the incident electrons [112]. Note that the scattering/interaction probability is directly proportional to the atomic number and inversely proportional to the atomic number. Moreover, the thicker samples, the higher probabilities of interaction. The electron energies provided by the electron guns for measuring biological samples are fixed in TEM between 100 keV to 1 MeV. Moreover, the thin sample and the low atomic number of atoms that compose the biological samples make the probability of interaction very low. Thus, it is considered that the electron beam passes through the sample without suffering any interaction at all or it interacts once. It is assumed that the interaction is elastic, being the electron energy and momentum conserved magnitudes. Despite these assumptions, reality is slightly different, in all electron-matter interactions some inelastic processes occur. A proof of that is the measurement of X-rays in the microscope column as Brehmstralung or characteristic radiation. However, in the sake of simplicity these inelastic interactions are neglected.
The intensity of the scattered beam will be measured in a far point P, where the set of N scattered plane waves from the N scatterer elements will interfere. The interference is characterized by the the optical path length between scattered rays. Assuming the reference system in the center of the macromolecule, the optical path difference between an incident electron and any scattered electron will be given by

$$\mathbf{s}\cdot\mathbf{r}_k - \mathbf{s}_0\cdot\mathbf{r}_k = (\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{r}_k \tag{92}$$

Note that the optical path length only depends on the scattering direction but it is independent of the distance to the scatterer element (assuming that the point P is enough far, i.e. Fraunhofer regime). Electrons with energy E or alternatively (see Eq. (9)) after interact elastically with the sample will be characterized by the *scattering vector*, $\mathbf{q}$, defined as

$$\mathbf{q} = \frac{2\pi}{\lambda}(\mathbf{s}-\mathbf{s}_0). \tag{93}$$

It is noteworthy that the magnitude, q, of the scattering vector present frequency $(1/\text{Å})$ units because of its dependence on the wavelength,

$$q = \frac{4\pi}{\lambda}\sin\theta. \tag{94}$$

Each electron will be modeled by a wave function, with amplitude A and frequency $2\pi/\lambda$. In particular, the incident electron at the position $\mathbf{r}_{k,\lambda}$ will present wave function

$$\Psi_{in} = Ae^{i\frac{2\pi}{\lambda}\mathbf{s}_0\cdot\mathbf{r}_k}. \tag{95}$$

Note that the amplitude of the scattered electron will be kept after the scattering process (for being elastic the energy is kept, which is contained in the amplitude term). The far point P along the scattering direction, $\mathbf{s}$, is defined by the position vector, $\mathbf{l}$ with origin in the scatterer element $\mathbf{r}_k$, see Fig. 46. The relation between the incident, Ψ_{in} and scattered wave, $\Psi_k(\mathbf{s})$ is given by the linear relation

$$\begin{aligned}\Psi_k(\mathbf{s}) &= \Psi_{in} f_{k,\lambda} e^{-i\frac{2\pi}{\lambda}\mathbf{s}\cdot(\mathbf{r}_k+\mathbf{l})} \\ &= Ae^{i\frac{2\pi}{\lambda}\mathbf{s}_0\cdot\mathbf{r}_k} f_{k,\lambda} e^{-i\frac{2\pi}{\lambda}\mathbf{s}\cdot(\mathbf{r}_k+\mathbf{l})}\end{aligned} \tag{96}$$

where $f_{k,\lambda}$ is the electron atomic scattering factor at wavelength λ. It is remarkable that the optical path difference, $(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{r}_k$ allows to rewrite the last expression in terms of the scattering vector, $\mathbf{q}$ as

$$\Psi_k(\mathbf{q}) = Af_k(q)e^{-i\mathbf{q}\cdot\mathbf{r}_k}e^{-i\frac{2\pi}{\lambda}\mathbf{s}\mathbf{l}}. \tag{97}$$

To work with scattering vectors instead of scattering directions is convenient as will be shown in next calculus. Moreover, a change of variable in the electron atomic scattering factor was performed, writing now its dependence on q instead on the wavelength, Eq. (94) allows this duality.
The aim is to calculate the intensity of the electron scattered beam in the far point P. Thus, the intensity will be the interference of all scattered waves. Thus, the amplitude in the point P can be calculated by applying the superposition principle,

$$\Psi(\mathbf{q}) = \sum_{k=1}^{N}\Psi_k(\mathbf{q}) = Ae^{-i\frac{2\pi}{\lambda}\mathbf{s}\mathbf{l}}\sum_{k=1}^{N} f_k(q)e^{-i\mathbf{q}\cdot\mathbf{r}_k}, \tag{98}$$

where the sum is extended up to the number N of scatterers, i.e. atoms in the macromolecule. This equation is directly related with the electron density, $\rho(\mathbf{r})$, of the macromolecule. There are two ways to establish this relationship. First, it can be assumed a probabilistic model under the hypothesis that the scattering amplitudes are proportional to the local charge, $df_k(\mathbf{r}) = CdQ(\mathbf{r}) = C\rho(\mathbf{r})dV$, [52], where C is the proportionality constant. Moreover, assuming that the phase of the scattered waves, is essentially the same between neighboring atoms, i.e. $\mathbf{q}\mathbf{r}_k \approx \mathbf{q}\mathbf{r}_j$ for $j \neq k$, then, (98) the sum can be converted into an integral

$$A(\mathbf{q}) = \sum_k \Psi_k = A\sum_k^N f_k e^{i\mathbf{q}\mathbf{r}_k} \qquad \rightarrow \qquad A(\mathbf{q}) = AC\int \rho(r)e^{i\mathbf{q}\mathbf{r}}dV. \tag{99}$$

Leaving out proportionality the constant, the amplitude is essentially given by structure factor which is defined as

$$F(\mathbf{q}) = \int \rho(\mathbf{r}_k) e^{i\mathbf{q}\mathbf{r}_k} dV. \tag{100}$$

The second approach is to consider the definition of electron density from the superposition of electron scattering form factors

$$\rho(\mathbf{r}) = \sum_{k=1}^{N} b_k(\mathbf{r} - \mathbf{r}_k) = \sum_{k=1}^{N} b_k(\mathbf{r}) \star \delta(\mathbf{r} - \mathbf{r}_k), \tag{101}$$

where $b_k(\mathbf{r})$ is the function resulting from the inverse Fourier transform of the $f_k(q)$ coefficients:

$$b_k(\mathbf{r}) = \int f_k(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{q} \tag{102}$$

then by performing the Fourier transform of Eq. (102) a variable F is obtained which squared is named *structure factor*

$$F(\mathbf{q}) = \int \rho(\mathbf{r}) e^{-i\mathbf{q}\mathbf{r}} d\mathbf{r} = \sum_{k=1}^{N} f_k(\mathbf{q}) e^{-i\mathbf{q}\cdot\mathbf{r}_k} \tag{103}$$

As a consequence, Eq. (98) con be rewritten in term of F as

$$\Psi(\mathbf{q}) = A e^{-i\frac{2\pi}{\lambda}\mathbf{s}\mathbf{l}} F(\mathbf{q}) \tag{104}$$

However, it should be noteworthy that, detectors record intensity, and therefore, it is preferred to work is measurable magnitudes, i.e. intensity,

$$I(\mathbf{q}) = |\Psi(\mathbf{q})|^2 = A^2 |F(\mathbf{q})|^2 = A^2 \int\int \rho(\mathbf{r}') \rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}'} e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} d\mathbf{r}'. \tag{105}$$

7.3 STRUCTURE FACTOR

The structure factor defined in Eq. (103) is directly related to the charge density via Fourier transform. In particular, at zero frequency $\mathbf{q} = \mathbf{o}$, the structure factor represents the total charge, Q, of the macromolecule and the number of charges M

$$Q = F(\mathbf{o}) = \int \rho(\mathbf{r}) dV \qquad \Rightarrow \qquad M = \frac{F(\mathbf{o})}{e} \tag{106}$$

where, e is the electron charge. However, instead of working with the structure factor it is more common to work with its squared modulus, $|F(\mathbf{q})|^2$, which for language extension is also referred as structure factor

$$|F(\mathbf{q})|^2 = \int\int \rho(\mathbf{r}') \rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}'} e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} d\mathbf{r}'. \tag{107}$$

Just by performing the simple variable change, $\hat{\mathbf{r}} = \mathbf{r}' - \mathbf{r}$, Eq. (107) is considerably simplified

$$|F(\mathbf{q})|^2 = \int\int \rho(\mathbf{r})\rho(\hat{\mathbf{r}}+\mathbf{r})e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}}d\mathbf{r}d\hat{\mathbf{r}}. \tag{108}$$

It should be remarked that $|F(\mathbf{q})|^2$ is an even function, because $|F(-\mathbf{q})|^2 = |F^*(\mathbf{q})|^2 = |F(\mathbf{q})|^2$. The integral is determined by the autocorrelation function of the electron density, denoted by

$$\gamma(\hat{\mathbf{r}}) = \int_0^{2R} \rho(\mathbf{r})\rho(\hat{\mathbf{r}}+\mathbf{r})d\mathbf{r}. \tag{109}$$

The integral limits are bounded by the maximum distance of the macromolecular, thus, considering that the macromolecule is tight to the interior of a sphere, then, the sphere radius will be the maximum radius of the macromolecular. Therefore the autocorrelation function is limited to 2R. Introducing the autocorrelation function into Eq. (108), then, the modulus squared of the structure factor results as the Fourier Transform of the autocorrelation function,

$$|F(\mathbf{q})|^2 = \int \gamma(\hat{\mathbf{r}})e^{-i\mathbf{q}\cdot\hat{\mathbf{r}}}d\hat{\mathbf{r}}. \tag{110}$$

The expression of the structure factor as Fourier transform of the autocorrelation function of the electron density map shows how the structure factor is determined by the particle geometry, it means by the electron density. Unfortunately, the structure factor of only a few geometries can be calculated analytically. The goal will then be to provide a general expression for integrating the structure factor under specific conditions. To do that, firstly it is rewritten in spherical coordinates, being β the angle between $\hat{\mathbf{r}}$ and $\mathbf{q}$

$$|F(\mathbf{q})|^2 = \int\int\int \gamma(\hat{r}\sin\beta\cos\phi, \hat{r}\sin\beta\sin\phi, \hat{r}\cos\beta)e^{-iq\hat{r}\cos\beta}\hat{r}^2\sin\beta\, d\hat{r}d\beta d\phi. \tag{111}$$

Here, an assumption of the autocorrelation function is performed, we assume that it presents radial symmetry or alternatively calculating the radial average along a given frequency. Then, the structure factor can be calculated as

$$\begin{aligned} |F(q)|^2 &= \int\int\int \gamma(\hat{r})e^{-iq\hat{r}\cos\beta}\hat{r}^2\sin\beta\, d\hat{r}d\beta d\phi \\ &= 4\pi\int \hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r} \end{aligned} \tag{112}$$

As a consequence, the radial average of the structure factor at zero frequency will be

$$|F(0)|^2 = 4\pi\int \hat{r}^2\gamma(\hat{r})d\hat{r} \tag{113}$$

which simplifies the structure factor expression

$$|F(q)|^2 = |F(0)|^2\frac{\int \hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r}}{\int \hat{r}^2\gamma(\hat{r})d\hat{r}}. \tag{114}$$

7.4 GUINIER APPROXIMATION. GUINIER LAW

The structure factor given by Eq. (114) can hardly ever be integrated in terms of elemental functions, and usually requires numerical integration. The electron density determines the shape of the macromolecule. Therefore, the structure factor of only a few geometries can be calculated analytically. To avoid that, the small angle scattering approximation can be used. Thus, the Guinier idea was to expand the sinc function in Taylor series up to third order [59]

$$\operatorname{sinc}(q\hat{r}) \approx 1 - \frac{1}{3!}(q\hat{r})^2 + O((q\hat{r})^4). \tag{115}$$

It allows to integrate Eq. (114) to get a simple the simpler linear expression

$$|F(q)|^2 \approx |F(0)|^2 \left(1 - \tfrac{1}{3!}\frac{\int \hat{r}^4 \gamma(\hat{r})\,d\hat{r}}{\int \hat{r}^2 \gamma(\hat{r})\,d\hat{r}} q^2\right), \tag{116}$$

we may define R_g as the *radius of gyration*

$$R_g^2 = \frac{1}{2}\frac{\int \hat{r}^4 \gamma(\hat{r})\,d\hat{r}}{\int \hat{r}^2 \gamma(\hat{r})\,d\hat{r}}. \tag{117}$$

Note that the gyration radius can be understood as the ratio between moments of fourth and second order. Introducing this in the structure factor it results

$$|F(q)|^2 \approx |F(0)|^2 \left(1 - \frac{R_g^2}{3} q^2\right). \tag{118}$$

The last polynomial coincides with the Taylor expansion of third order of an exponential function, and therefore it will be substituted by the exponential. Note how, the use of Taylor polynomial can be applied in both directions, i.e. to approximate a function by a polynomial, or to approximate a polynomial by a function. The essence of Taylor expansion is that the function and polynomial are locally equal around the point chosen to expand in Taylor series. Thus, the second Guinier step consisting in substituting Eq. (118) by the exponential is justified casting

$$|F(q)|^2 \approx |F(0)|^2 e^{-\frac{R_g^2}{3} q^2}. \tag{119}$$

The exponent defines the so-called B-factor, $B = \frac{1}{3}R_g^2$. The negative exponent implies a decay of the structure factor $F(q)$. As it was commented in Section 7.1, the sharpening process consists in determining B, to get a flat structure factor curve.

7.5 RANGE OF GUINIER APPROXIMATION

There are two facts about Guinier approximation that ought to be discussed. The first one is the Taylor expansion of the sinc function in Eq. (115). The second one is the Guinier approximation of the sinc function by an exponential in the Guinier Law, see Eq. (118) and Eq. (119). These two approximations constrain the frequency

range in which the Guinier law is verified.
The Taylor expansion in Eq. (114) is traditionally justified considering very low scattering angles. However, a rigorous proof should take into account the product $q\hat{r} \to 0$. To check the validity of the range of the Taylor expansion, the relative error, ϵ, between the sinc function and its Taylor series up to third order can be calculated

$$\epsilon = |\frac{\mathrm{sinc} q\hat{r} - (1 - \frac{1}{6}(q\hat{r})^2)}{\mathrm{sinc} q\hat{r}}| \tag{120}$$

If it is assumed a 10% of relative error, then,

$$q\hat{r} \leqslant \frac{5}{3} \tag{121}$$

Despite that this inequality provides a criterion for the product, what it really of interest is the frequency range. Thus, considering that $\hat{r}$ is limited by the particle radius, R, then, $qR \leqslant 5/3$, or equivalently, for resolutions $1/q \geqslant 3R/5$. A simple numerical example is given, for a macromolecule with radius $R = 100\text{Å}$ the sinc approximation by a polynomial will be valid for resolutions lower than 60Å(assuming a relative error of 10%). A graphical solution by plotting relative error between the sinc function and its Taylor expansion to first order was also performed, see Fig. 47.

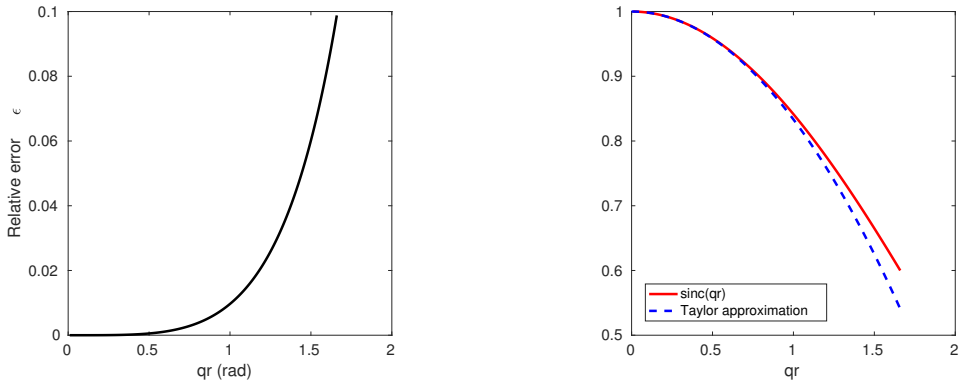

Figure 47: (a) Relative error between the sinc function and its Taylor expansion up to third order. (b) sinc function (continuous - red) and Taylor approximation up to third order (dashed - blue). (c). Relative error ξ, derived from the Guinier approximation is carried out, i.e. when the polynomial is approximated by an exponential. (d) exponential function (continuous - red) and Taylor approximation up to third order (dashed - blue).

The second item to be discussed is the Guinier approximation, i.e. the conversion step from the Taylor expansion (Eq. (115)) into the exponential (Eq. (119)). The validity of this Guinier step means to evaluate the difference between the exact expression and the performed approximation

$$\epsilon = \left| \frac{1 - \frac{1}{3}R_g^2 q^2 - e^{-\frac{1}{3}R_g^2 q^2}}{1 - \frac{1}{3}R_g^2 q^2} \right|. \tag{122}$$

The goal will be then to assume an error ϵ, and then to solve the equation for analyzing the range of validity of Guinier approximation. The W-Lambert [27] function allows to calculate that solution, in particular, the solution of

$$e^{-cz} = a_0(z-r) \quad \Rightarrow \quad z = r + \frac{1}{c}W\left(\frac{c}{a_0}e^{-cr}\right). \tag{123}$$

Eq. (122) is rewritten to establish a direct comparison to Eq. (123), thus,

$$e^{-\frac{1}{3}R_g^2q^2} = -(1+\epsilon)\left(\frac{1}{3}R_g^2q^2 - 1\right) \tag{124}$$

which cast, as solution

$$\frac{1}{3}R_g^2q^2 = 1 + W\left(-\frac{1}{(1+\epsilon)e}\right) \tag{125}$$

This relationship allows knowing the value of the product $R_g^2q^2$ for a given error ϵ. An alternative is to plot Eq. (47), see Fig. **??**. In particular, it shows how the committed error is less than 10% when

$$\frac{1}{3}R_g^2q^2 \leqslant 0.3755 \quad \Rightarrow \quad R_gq \leqslant 1.06 \tag{126}$$

As a consequence, a criterion can be approximately established as $R_gq \leqslant 1$. Note that this solution is independent of the scatterer shape. Finally, as well as it was performed with the approximation of the sinc function by the polynomial, if a particle with gyration radius, $R_g = 100$Å is considered, then, the Guinier approximation is only valid for frequencies lower than 94 Å. In Fig. 122 the relative error given by Eq. (122) is plotted in terms of the product $1/3R_g^2q^2$, in agreement with analytic results. It is noteworthy that the gyration radius is usually around the macromolecular radius. As a consequence, the Guinier approximation is only valid at very low frequencies. If higher error are assumed, for instance, $\epsilon = 0.2$, the frequency range is increased being the Guinier approximation only valid for frequencies lower than 77 Å.

7.6 POROD APPROXIMATION. HIGH FREQUENCIES

Guinier approximation is limited to a range of low frequencies, constrained to the inequality $qR_g \leqslant 1$. In this section the structure factor will be analyzed for higher frequencies, in particular at high frequencies giving as result the Porod invariant. The general expression for the structure factor is given by (114), if Eq. (113) is substituted

$$|F(q)|^2 = |F(0)|^2\frac{\int \hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r}}{\int \hat{r}^2\gamma(\hat{r})d\hat{r}} = 4\pi\int_0^{2R}\hat{r}^2\gamma(\hat{r})\mathrm{sinc}(q\hat{r})d\hat{r} =$$

$$\frac{4\pi}{q}\int_0^{2R}\hat{r}\gamma(\hat{\mathbf{r}})\sin(q\hat{r})d\hat{r},$$

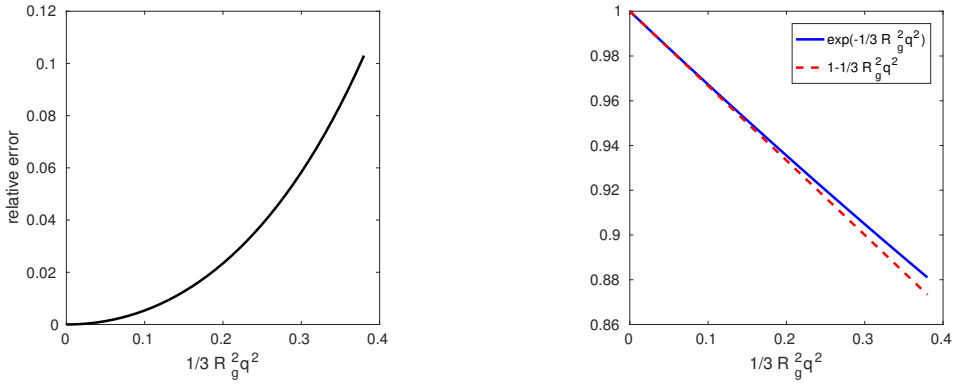

Figure 48: (a) Relative error between the exp function and its Taylor expansion.. (b) exponential function (continuous - red) and Taylor approximation (dashed - blue). (c). Relative error ξ derived from the Guinier approximation is carried out, i.e. when the polynomial is approximated by an exponential. (d) exponential function (continuous - red) and Taylor approximation up to third order (dashed - blue).

Hence, the structure factor can be integrated twice by parts, first integration casts

$$= \frac{4\pi}{q}\left[-\frac{1}{q}\hat{r}\gamma(\hat{\mathbf{r}})\cos q\hat{r}\right]_0^{2R} + \frac{4\pi}{q}\frac{1}{q}\int_0^{2R}(\hat{r}\gamma(\hat{r}))'\cos q\hat{r} =$$

and the second integration gives

$$= \frac{4\pi}{q}\left[-\frac{1}{q}\hat{r}\gamma(\hat{\mathbf{r}})\cos q\hat{r} + \frac{1}{q^2}(\hat{r}\gamma(\hat{r}))'\sin q\hat{r}\right]_0^{2R} - \frac{4\pi}{q}\frac{1}{q^2}\int_0^{2R}(\hat{r}\gamma(\hat{r}))''\sin q\hat{r}$$

Reminding that, $\gamma(2R) = 0$, then

$$= \frac{4\pi}{q^3}\left[2R\gamma'(2R)\sin 2qR - \int_0^{2R}(\hat{r}\gamma(\hat{r}))''\sin q\hat{r}\right],$$

which integrating once more time

$$= \frac{4\pi}{q^3}\left[2R\gamma'(2R)\sin 2qR + \left[\frac{1}{q}(\hat{r}\gamma(\hat{\mathbf{r}}))''\cos q\hat{r}\right]_0^{2R} - \frac{1}{q}\int_0^{2R}(\hat{r}\gamma(\hat{r}))'''\cos q\hat{r}\right] =$$

$$= \frac{4\pi}{q^3}\left[2R\gamma'(2R)\sin 2qR + \left[\frac{1}{q}(2\gamma'(\hat{r}) + \hat{r}\gamma''(\hat{r})\cos q\hat{r}\right]_0^{2R} - \frac{1}{q}\int_0^{2R}(\hat{r}\gamma(\hat{r}))'''\cos q\hat{r}\right] =$$

$$= \frac{4\pi}{q^3}\left[2R\gamma'(2R)\sin 2qR + \frac{1}{q}(2\gamma'(2R) + 2R\gamma''(2R))\cos 2qR + \right.$$
$$\left. -\frac{1}{q}2\gamma'(0) - \frac{1}{q}\int_0^{2R}(\hat{r}\gamma(\hat{r}))'''\cos q\hat{r}\right]$$

Assuming that γ is a smooth and differentiable function, there is only one term non oscillating that therefore at high scattering q-values the structure factor can be approximated by the *Porod law*

$$|F(q)|^2 = -\frac{8\pi}{q^4}\gamma'(0). \tag{127}$$

The Porod law implies than even at very high frequencies the structure factor will present a decay.

CONCLUSION FROM ANALYTIC EXPRESSIONS

In the last two sections it was calculated the shape of the structure factor at two different spectral ranges. It was found that the Guinier law is only valid at very low frequencies, meanwhile at high frequencies the behaviour of the structure factor is given by the Porod law. In both cases, the expression of the structure factor decays with the frequency. The consequence, is that the flattening spectra that B-factor does, is not supported by the scattering theory, (at least in the spectral range given by the Guinier and Porod law), i.e. the structure factor must decrease with the frequency.

7.7 STRUCTURE FACTOR OF EXPERIMENTAL DATA

To analyze the validity of the exposed theoretical results, a set of tests with experimental data were carried out. These experiments involve two kinds of data: first, by making use macromolecular atomic models (converted into density maps), and second by considering the reconstructed 3D density map. In both cases the profile of the structure factor is analyzed.

7.7.1 *Experimental validation of the Guinier law*

The aim of this section is to validate the interval of frequencies in which the Guinier law is valid. To do that, the atomic model of the β−galactosidase [1] (PDB-3j7h) is considered. The model was converted into density map by means of `xmipp_volume_from_PDB` [161] with a sampling rate of 1 Å/pixel. Then, the structure factor was calculated and represented in the Guinier plot, see Fig. 49. For sake of simplicity, it was normalized, it means $|F(q)|^2/|F(0)|^2$.
The plot shows exactly what the theory predicts. The structure factors shows a linear behaviour up to about 100 Å. In Fig. 49, this breakpoint was highlighted with a vertical line. From this breakpoint, the Guinier law cannot be applied anymore, and another approach should be used to analyze the structure factor. However, it is observed that the curve does not present a flat profile. Hence, the B-factor correction is not justified beyond visualization purposes. In other words, the B-factor correction enhance the visualization, but the obtained map does not represent a real protein.

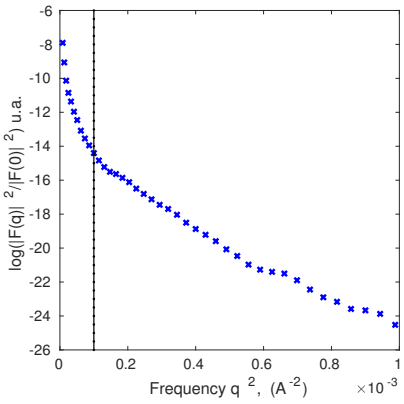

Figure 49: Guinier plot for the atomic model PDB 3j7h converted into density map, on the low frequency range up to 30 Å. The Guinier law is verified for resolutions lower than 100Å.

7.7.2 *Experiments with secondary structure elements*

This second experiment is addressed to show the dependence of the structure factor on the secondary structure elements. To do that, the following atomic models were considered: the structure of β-galactosidase [1] (PDB-3j7h), TRPV1 [99] (PDB-3j9j), a triple mutant of the NHAA dimer [19] (PDB-4atv) and the Yeast 20S proteasome in complex with Ac-PAE-ep [33] (PDB-4y6v) were used. They were converted into into density maps with a sampling rate of 1Å/pixel using `xmipp_volume_from_PDB` [161]. Then, the structure factor was calculated, normalized, and plotted $\log(|F(q)|^2/|F(0)|^2) - q$. In Fig. 50 these plots were represented.
To understand how the elements of secondary structure affect to the structure factor curve, a set of experiments were performed by modifying the atomic model

1. *Experiment 1. Relationship to the atom description*: In this experiment it is obtained a less accurate density map by considering as atom descriptors gaussian functions with variance proportional to the atomic numbers. This transformation should not have too much impact in the structure, as can be seen in Fig. 50, green-continuous-dotted.

2. *Experiment 2. Relationship to the atom nature*: A more dramatic scenario is considered by substituting all carbon atoms in the atomic model to iron atoms. This transformation casts a heavy macromolecule with the same shape. The results on the structure factor can be seen in Fig. 50, blue dashed line. It is noted that the structure factor profile is essentially the same but shifted up, because iron scatters more than carbon.

3. *Experiment 3. Relationship to the overall macromolecule shape*: In this test, the internal structure of the macromolecule is completely change, keeping its overall shape. The density map was binarized and then the structure factor was calculated. Note that, this transformation keeps overall shape of macromolecule but completely destroys the atomic information. The results can be observed in Fig. 50, red triangles line. The result shows that the structure

factor curve is kept and shifted slight slight differences in particular at high resolution.

4. *Experiment 4. Relationship to the relative atomic positions*: In this last test, the secondary structure of the macromolecule was destroyed. The atomic positions were randomly displaced from their original positions. It was performed by modifying the pdb. Thus, it was considered a uniform distribution with maximum shift of 6Å which implies movements from -6 to 6 Å in all directions. The radius of an α-helix is 6Å, and the sideway distance between α carbons in a β sheet is around 5Å, the proposed displacements completely destroys the secondary structure. The results on the structure factor are shown in Fig. 50, cyan dots. Again, the structure factor profile keeps the shape of the original one, with an slight shift.

In all these experiments the structure factor of the atomic models converted into density maps, presents a non flat profile. As a consequence, the structure factor must decay with the frequency as it was pointed out in the theory.

7.7.3 *B-factor correction to compensate the overdampened spectra*

In Fig. 51 we show the structure factor profile for β-galactosidase obtained by processing with Relion the data from EMPIAR entry 10013, and the structure factor profile of the corresponding PDB entry 3j7h. As can be seen from the figure, the structure factor of the PDB falls faster (implying a larger B-factor) than the one of the Relion reconstruction. Although, this is not necessarily the case for all reconstructions, this means that the purpose of the B-factor correction normally applied to 3D reconstructions cannot be to compensate for an extra filtering allegedly introduced by the reconstruction algorithm.

7.7.4 *B-factor correction on atomic models*

All theoretical results and experiments are in agreement with the theory, however, the biological understanding of the results can also be affected. To show that, the atomic model of a structure of the β-galactosidase (PDB-3j7h) [1] was converted into density map using `xmipp_volume_from_pdb` [161]. The B-factor correction was then applied. Note that this density map represents the perfect shape of the macromolecule, and therefore it should be invariant under sharpening corrections because it cannot be enhanced. In Fig. 52 a comparison in the density map of the same region of the β-galactosidase are shown. It is noted a slight but significant density variation between both maps. This simple experiment shows that the global B-factor correction might be not the best sharpening strategy, and in agreement with the scattering theory, it should be revisited.

7.8 CONCLUSIONS

1. Guinier law is only valid at very low frequencies, the theory predicts an exponential decay of the structure factor.

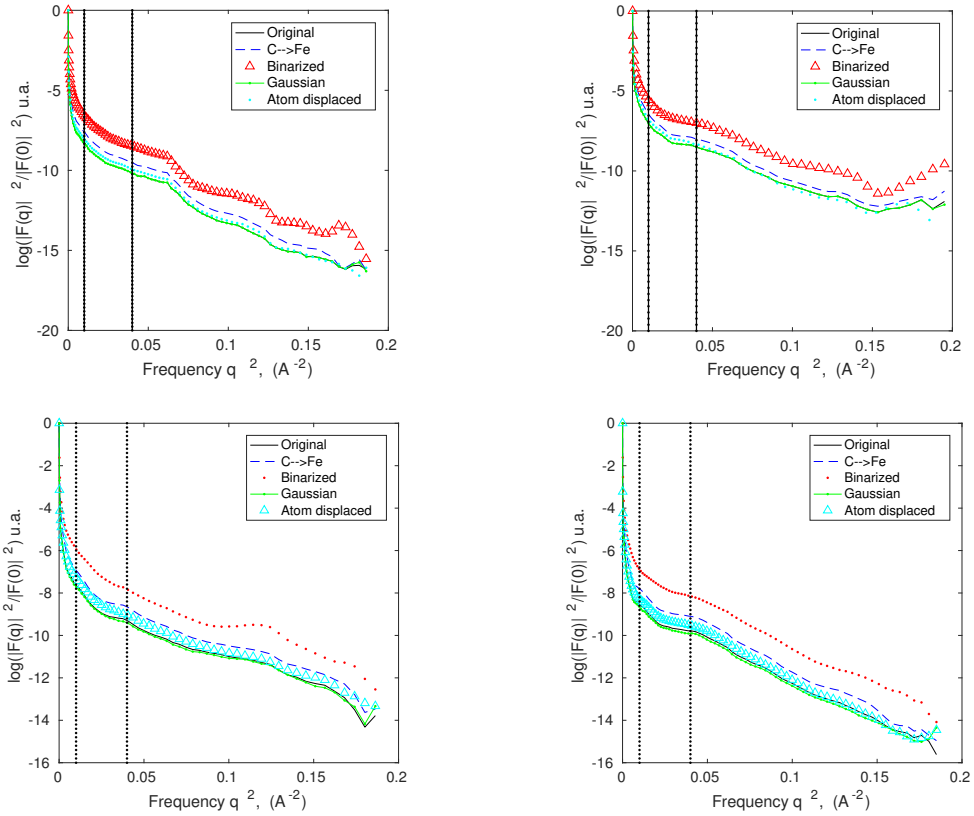

Figure 50: Normalized structure factors up to high resolution for (up-left) PDB 3j7h, (up-right) 3j9j, (down-left) PDB 4atv and (down-right) PDB4 y6v. a) (Continuous-black) original converted density map. b) (blue-dashed) Substituted C atoms by Fe atoms. c) (red-triangles) Binarized density map. d) (green-continuous-dotted) Map converted from PDB substituting atoms by Gaussian functions. e) (cyan-dotted) Random displacement of atoms. The vertical lines determine resolutions of 10 and 5 Å.

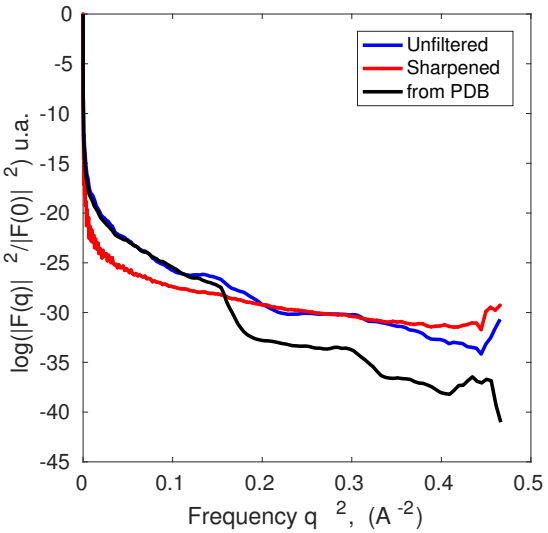

Figure 51: Normalized structure factor from a) (Continuous-black) Relion map. b) (red-dashed) atomic model (pdb entry 3j7h) converted into electron density map. The vertical lines determines the frequencies of 10 and 5 Å.

2. Up to our knowledge, the first derivation of its validity range has been derived.

3. The Porod low is only valid at high resolution range, and also predicts a decay of the structure factor.

4. The structure factor of a macromolecule has a non-flat slope with independence of the spectral range.

5. The B-factor correction helps in the visualization of the macromolecular complex but it does not represent the macromolecular complex. It is not justified by the scattering theory.

6. 3D Reconstruction algorithms do not always over-dampen the spectrum of the macromolecule.

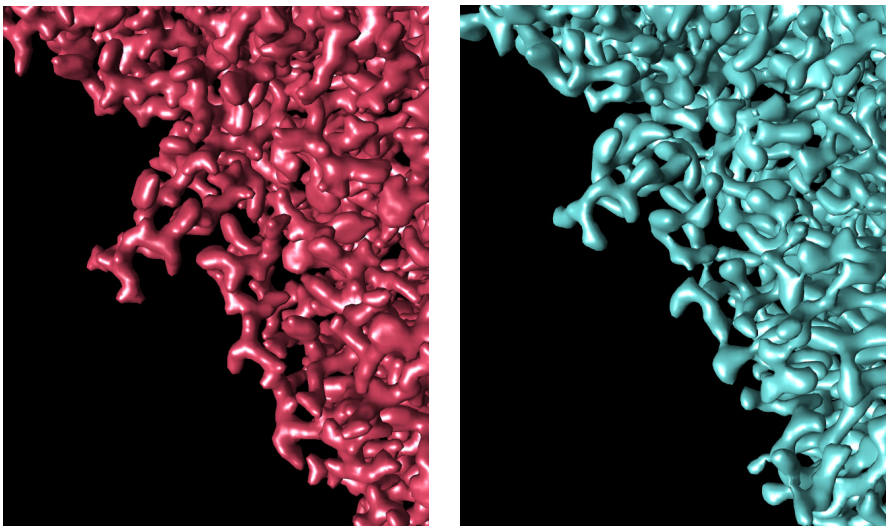

Figure 52: (left) Map converted into electron density map. (right) B-factor corrected map.

8

LOCAL SHARPENING BASED ON LOCAL RESOLUTION MEASURES

Once an electron density map is obtained, the final purpose is to create an atomic model. The reliability of the atomic model is higher as the resolution of the map increases. In this sense resolution plays the role of quality parameter of the map. Thus, it is wished to get high resolution maps, unfortunately it involves some drawbacks, which, leaving out measurement errors (as they are radiation damage, microscope aberrations, or defocus among others), can be mainly summarized in: sample heterogeneity, computational accuracy or alignment. They are responsible of the low contrast at high frequencies. This fact affects to the reconstructed map introducing a blurring effect that low weights the high resolution content. Traditionally this lose of contrast at high resolution has been compensated by the so-called *sharpening* techniques, where B-factor correction is the most spread as was shown in the previous chapter. However, despite the good results can be obtained with the B-factor correction, its use is not supported by the scattering theory. In addition, to the best of our knowledge, sharpening methods neglect the local resolution of the map. For that reason, in this chapter the local resolution information is used as an aggregated value to enrich the degree of detail of the reconstructed structure by means of a local sharpening procedure. Hence, a new local sharpening method, named *LocalDeBlur*, is proposed as an application of the local resolution measurement. It is a map restoration method based on a Wiener filter based on local resolutions. Hence, the structure factor flattening is avoided. The results show a significant improvement in the interpretability of the sharpened map, which has special impact in the biological understanding of the macromolecule and the tracing of the atomic model. In particular, in maps with a broad local resolution variation, for instance membrane proteins.

8.1 CURRENT LOCAL SHARPENING METHODS

In the previous chapter, the B-factor correction was introduced with the aim of increasing interpretability of density maps. It consists in a global transformation that boosts the high frequencies, depending on the algorithm certain constrains are applied [135, 178]. However, as local resolution shows, density maps might present different degree of detail, which suggests local sharpening tools. The current state of the art is dominated by *LocScale* [81] that is briefly introduced in this section before proposing our alternative named *localDeBlur*.

8.1.1 *Local sharpening - LocScale*

The fundamentals of *LocScale* [81] is the search of similarity between the structure factor (amplitude values of Fourier components) of the reconstructed density map

and its corresponding atomic model (converted into density map) in a local sense. It is achieved by means of correcting the structure of the protein to improve the reconstructed density map, but according to the atomic model. Thus, the flattening spectra is avoided. To do that, a prior information is required, in this case an atomic model that is converted into density map. A moving window is used to determine the radial average of the structure factor inside it. Then, the ratio between the radial averaged of the structure factor of model, F_{model}, and map, F_{obs}, is calculated.

$$k(\omega) = \sqrt{\frac{\sum_{\omega \pm \Delta\omega} |F_{model}|^2}{\sum_{\omega \pm \Delta\omega} |F_{obs}|^2}}. \tag{128}$$

Now, it is possible to use this scale factor to enhance the contrast of the measured map by locally scaling its structure to be similar to the converted map, it means

$$F_{corr}(\omega) = k(\omega) F_{obs}(\omega), \tag{129}$$

where $F_{corr}(\omega)$ is the corrected structure factor. By doing this process iteratively, the map is refined in terms of contrast but keeping the same resolution.

8.2 SHARPENING BASED ON LOCAL RESOLUTION INFORMATION - LOCALDEBLUR

The sharpening process tries to restore the high resolution information contained by modifying the map contrast at different frequencies. In this section, a fully automatic method, named *LocalDeBlur*, that performs that adjustment of map in a local manner is proposed. The objective *LocalDeBlur* is then to enhance the map visualization which increases the biological understanding of the macromolecule. To do that, the local resolution information can be exploited. Each voxel of the reconstructed map will be the superposition with different weights (lineal combination) of many waves of different frequencies. Note that local resolution represents the highest frequency statistically measurable above the noise level. Coarsely talking *LocalDeBlur* will try to increase the weight of the limit frequency wave with respect to the rest waves. This mechanism is mathematically modelled as it follows: Given a reconstructed map $V(\mathbf{r})$, and its local resolution map, $H(\mathbf{r})$ is it possible to find a new map $V_{sh}(\mathbf{r})$, named sharpened map, such as

$$V(\mathbf{r}) = \Omega(H(\mathbf{r})) * V_{sh}(\mathbf{r}), \tag{130}$$

where the symbol * represents the convolution operation, and $\Omega(H(\mathbf{r}))$ is a local filter such as it filters the voxel in position $\mathbf{r}$ of $V_{sh}(\mathbf{r})$ at the frequency given by $H(\mathbf{r})$. In other words, the operator $\Omega(H(\mathbf{r}))$ is a local filter defined by the local resolution values.
To really understand the effect of the local filter, the original reconstructed map can be decomposed as sum of band pass filtered maps,

$$V(\mathbf{r}) = \sum_{\omega} V_{BP,\omega}(\mathbf{r}). \tag{131}$$

The effect of the local filter on the map $V(\mathbf{r})$ is to assign resolution and spatially dependent weights in the volume decomposition as it follows

$$\Omega(H(\mathbf{r})) * V_{sh}(\mathbf{r}) = \frac{\sum_\omega w_\omega(\mathbf{r}) V_{BP,\omega}(\mathbf{r})}{\sum_\omega w_\omega(\mathbf{r})}, \tag{132}$$

where $w_\omega(\mathbf{r})$ is a spatially dependent weight at resolution ω, and the denominator just represents a normalization factor. In particular, the weight factor is a Gaussian function that measured the distance between the filtering frequency, ω and the local resolution, ω_{loc} of the voxel in position $\mathbf{r}$,

$$w_\omega(\mathbf{r}) = e^{-K(\omega_{loc}-\omega)^2}, \tag{133}$$

being $K = 0.025$ a constant obtained empirically.
Once all elements of Eq. (130) have been defined, it rests to find the sharpened map $V_{sh}(\mathbf{r})$. It can be obtained in an iterative way via steepest descent approach [100]

$$V_{sh}^{(i+1)} = V_{sh}^{(i)} + \lambda \left[\Omega^T \left(V - \Omega V_{sh}^{(i)} \right) - \frac{1}{SNR} V_{sh}^{(i)} \right], \tag{134}$$

where the spatial dependence has been omitted for sake of simplicity. It is noteworthy that, when the SNR is enough high, the term $1/SNR\ V_{sh}^{(i)}$ can be neglected, and therefore, a simpler iterative expression is obtained,

$$V_{sh}^{(i+1)} = V_{sh}^{(i)} + \lambda \Omega^T \left(V - \Omega V_{sh}^{(i)} \right). \tag{135}$$

Because the sharpening process is usually, carried out once a reconstruction is obtained, the SNR term can be neglected and therefore, this last equation can be used as iterative solution of the ill posed problem, Eq. (130). The convergence of the algorithm depends on the λ parameter, the greater λ, the faster convergence. Unfortunately, there exists a compromise between convergence speed and sharpening. In the performed test, it was found that

$$\lambda = \frac{\|V(\mathbf{r})\|}{10\|\Omega(V(\mathbf{r}))\|}. \tag{136}$$

The criterion used to stop the iterative process considers a good convergence when the norm of the difference, $\|V^{i+1} - V^i\| \leqslant \|V^i\|/100$, between the current iteration and the previous one is smaller or equal to 1%.

8.3 RESULTS

LocalDeBlur was tested with experimental data sets taken from EMDB [95] and atomic models from PDB [13]. In particular, the the capsaicin receptor TRPV1 (EMDB-5778, PDB-3j5p) [99] and the *Plasmodium falciparum* 80S ribosome [198]. The local resolution map of the first one was estimated in Chapter 5. It was pointed out that it presents a broad local resolution range, which makes it a good candidate for a local sharpening approach.

To test the validity of the algorithm, their atomic model is known, and therefore, the ground-truth will be better matching between map and model. To measure the performance of *LocalDeBlur* and to establish a comparison with other sharpening algorithms, the following methods were also applied: the postprocessing - RELION [206] and *Autosharpen* - Phenix [178] as global methods, and *LocScale* [81] as local sharpening method.

8.3.1 *Experimental Map 1: capsaicin receptor TRPV1*

The first experimental example considers the capsaicin receptor TRPV1 (EMDB entry 5778) [99]. This volume has a dimensions of $256 \times 256 \times 256$ voxels, a pixel size of 1.1 Å/pixel, and the reported FSC resolution at 0.143 was 3.8 Å. The corresponding atomic model was taken from PDB (entry 3j5p). The local resolution map of this macromolecule was calculated in Chapter 5, and it can be observed in Fig. 32. Note that, the EMDB entry 5778 provides the original reconstruction and a B-factor corrected map.
LocalDeBlur was applied to sharpen the raw map, its results can be observed in Fig. 53. The global sharpening methods of postprocessing - RELION and Autosharpen - Phenix, were applied, as well as the local one, *LocScale*. In Fig. 53 the sharpened map with all methods and its comparison with the original reconstruction is shown. To expose the sharpening effect on the map of each method, an α-helix and its corresponding atomic model was overlapped, they can also be observed in Fig. 53. All methods exhibit a considerably improvement respect to the original reconstruction, and allow a better understanding of the structure. This improvement is specially observed in the represented α-helix, note how the post-processing is able to recover the real shape, but most of the elements of the side chain are out of the sharpened map. The *Autosharpen* of Phenix, could restore the α-helix and most of the side chain elements. the local methods of *LocalDeBlur* and *LocScale*, cast similar results. However, the *LocalDeBlur* map is slightly tighter to the atomic model than *LocScale*.
Finally, a combination of *LocalDeBlur* and *LocScale* was performed. The objective of this test was to: first with *LocalDeBlur* restore the high resolution components based on local resolution, and then with *LocScale*, to adjust the contrast to be as similar as possible to the atomic model. This combination seems to be the most properly sharpening method.
It remains to analyze the Guinier plot of these maps. They are represented in Fig. 54, where the dashed line represents the structure factor of the atomic model converted into density map using `xmipp_volume_from_pdb` [161]. Note how the structure factor of the converted map presents a decay with the frequency, as it was pointed out in Chapter 7. The global sharpening methods of postprocessing - RELION and *Autosharpen* of Phenix attempts to flatten this sprectrum, and the result is a divergence between the structure factor of the converted map and the sharpened map. Local approaches, *LocalDeBlur* and *LocScale*, keep the strucuture factor, in particular, the decay, in agreement with the scattering theory. As well as it occurs with the combination *LocalDeBlur* and *LocScale*. However, it must be

highlighted that *LocalDeBlur* achieves that without prior knowledge of the atomic model.

8.3.2 *Experimental Map 2: Plasmodium falciparum 80S ribosome*

The second experimental case of use considers map of the *Plasmodium falciparum* 80S ribosome taken from EMDB (EMDB entry-2660) [198]. This volume has a dimensions of $360 \times 360 \times 360$ voxels, a pixel size of 1.34 Å/pixel, and the reported FSC resolution at 0.143 was 3.2 Å. The corresponding atomic model was taken from PDB (entry 3j79). As it happened with the previous example, this entry of EMDB provides the half maps, and the full map. The local resolution map of this macromolecule was calculated with *MonoRes* using a single volume, the result can be observed in Fig. 55. The original map was then sharpened with *LocalDeBlur* using as input the obtained local resolution map. Other sharpening methods were also applied, as postprocessing - RELION and Autosharpen - Phenix and *LocScale*, all results are summarized in Fig. 55. To cast more light on the sharpening effect, a region of interest was magnified and overlapped with the atomic model. As it occurred with the TRPV1, the ribosome presents a broad resolution range, however, the differences between sharpened maps is clearer in this case of use. It indicates that sharpening methods seems to be map dependent. In particular, the global method of *Autosharpen* of Phenix was able to restore even better than the local method of *LocScale* the secondary structure. However, *LocalDeBlur* got again an excellent result in comparison to their alternatives, its results show an almost perfect matching with the atomic model, it is highlighted how much tight is the density map with the side chains. In particular, in this case the combination *LocalDeBlur* with *LocScale* considerably enriched the elucidation of the protein structure.
Finally, the structure factor of these sharpened maps was calculated, and represented in the Guinier plot. The results can be observed in Fig. 56. To do that, the atomic model was converted into density map by means of `xmipp_volume_from_pdb` [161]. Then the structure factor was calculated an represented in the Guinier plot (dashed line). As it happened in the previous example, the structure factor of the sharpened maps with local approaches present an excellent correspondence between the structure factor of the converted map and the sharpened one, in contrast with the local approaches. Again, it is remarkable the similarity of *LocalDeBlur* results with the converted map, without any knowledge about the model.

8.4 DISCUSSION

A new and fully automatic algorithm for local sharpening, called *LocalDeBlur*, has been developed. This method makes use of local resolution information to deblur the density map, in this sense and leaving the local filters out, it represents the first application of local resolution. Thus, the B-factor correction is avoided solving the problem of structure factor flattening, in agreement with the scattering theory. In contrast with global approaches, *LocalDeBlur* is specially suitable for map that presents a broad local resolution range. It is a logical result. Indeed, if resolution is understood as a measurement of SNR, then global sharpening approaches try to

sharp the map weighting by the SNR. When the SNR is more less constant in the whole map, the global sharpening works properly. In contrast, if it is spatially variant, it is necessary to weight locally. It is exactly what *LocalDeBlur* does, it modifies the local weights of the different frequencies highlighting the closer to the local resolution, and neglecting higher frequencies.
The results with experimental maps show a considerable improvement with respect to global sharpening approaches, and slightly better results than the current local approach, *LocScale*. However, the best result is obtained when a combination of both methods is considered in the order, first *LocalDeBlur* and then *LocScale*. This combination restores the high resolution elements, by means of a deblurring based on local resolution, to then adjust the contrast with *LocScale* trying to get the highest similarity with the atomic model.
The proper elucidation of the structure by means of sharpening process results critical for modelling process and tracing atomic models. In particular, it currently represents a hot topic in the field of structural biology, with applications to drug development.

8.5 CONCLUSIONS

1. A new and fully automatic free-parameter local sharpening method, named *LocalDeBlur* based on local resolution information has been proposed.

2. Up to our knowledge it represents the first application of local resolution (leaving out local filters)

3. *LocalDeBlur* avoids the B-factor quasi-flattening in agreement with the scattering theory.

4. It is compatible with *LocScale*, in fact the best result are achieved with a combination of both methods.

5. The proposed algorithm significantly improves the interpretability of density maps helping in the interpretability of the map and in elucidation of atomic models.

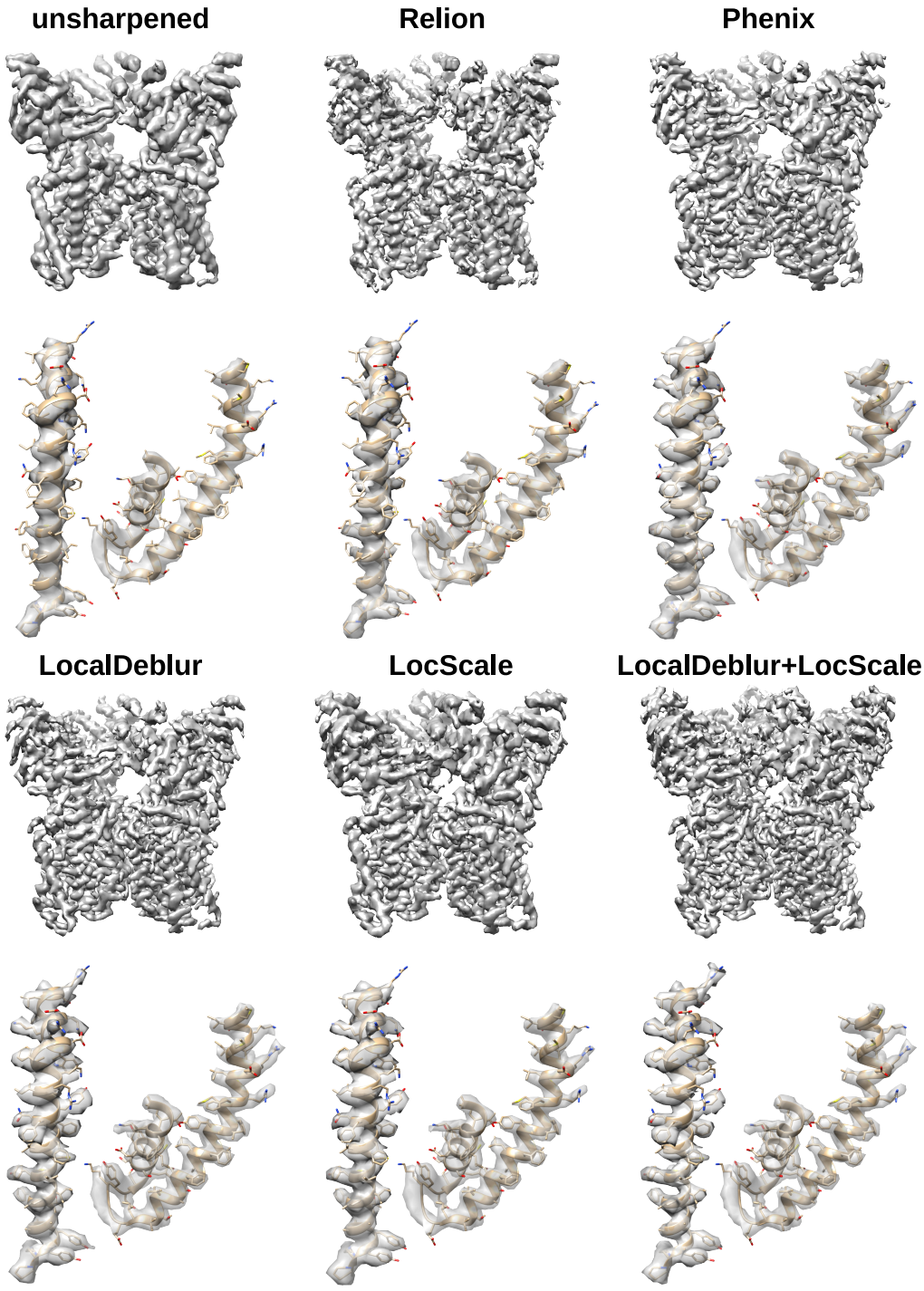

Figure 53: Capsaicin receptor TRPV1 and an example of α-helix overlapped with the atomic model. The original map and the sharpened maps obtained with post-processing - RELION, Autosharpen - Phenix, *LocalDeBlur*, *LocScale* and a combination of *LocalDeBlur* and *LocScale* are shown.

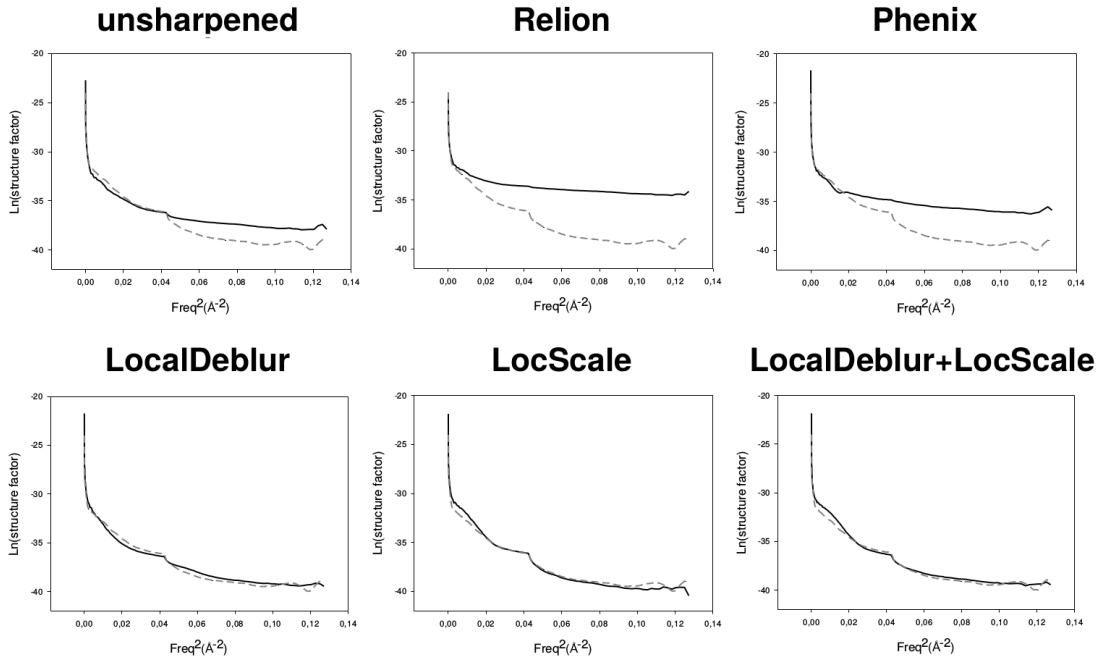

Figure 54: The Guinier plots for the Capsaicin receptor TRPV1 of the original map and the sharpened maps obtained with postprocessing - RELION, Autosharpen - Phenix, *LocalDeBlur*, *LocScale* and a combination of *LocalDeBlur* and *LocScale* are shown (continuous line). In dashed line it is represented the Guinier for the atomic model converted into density map.

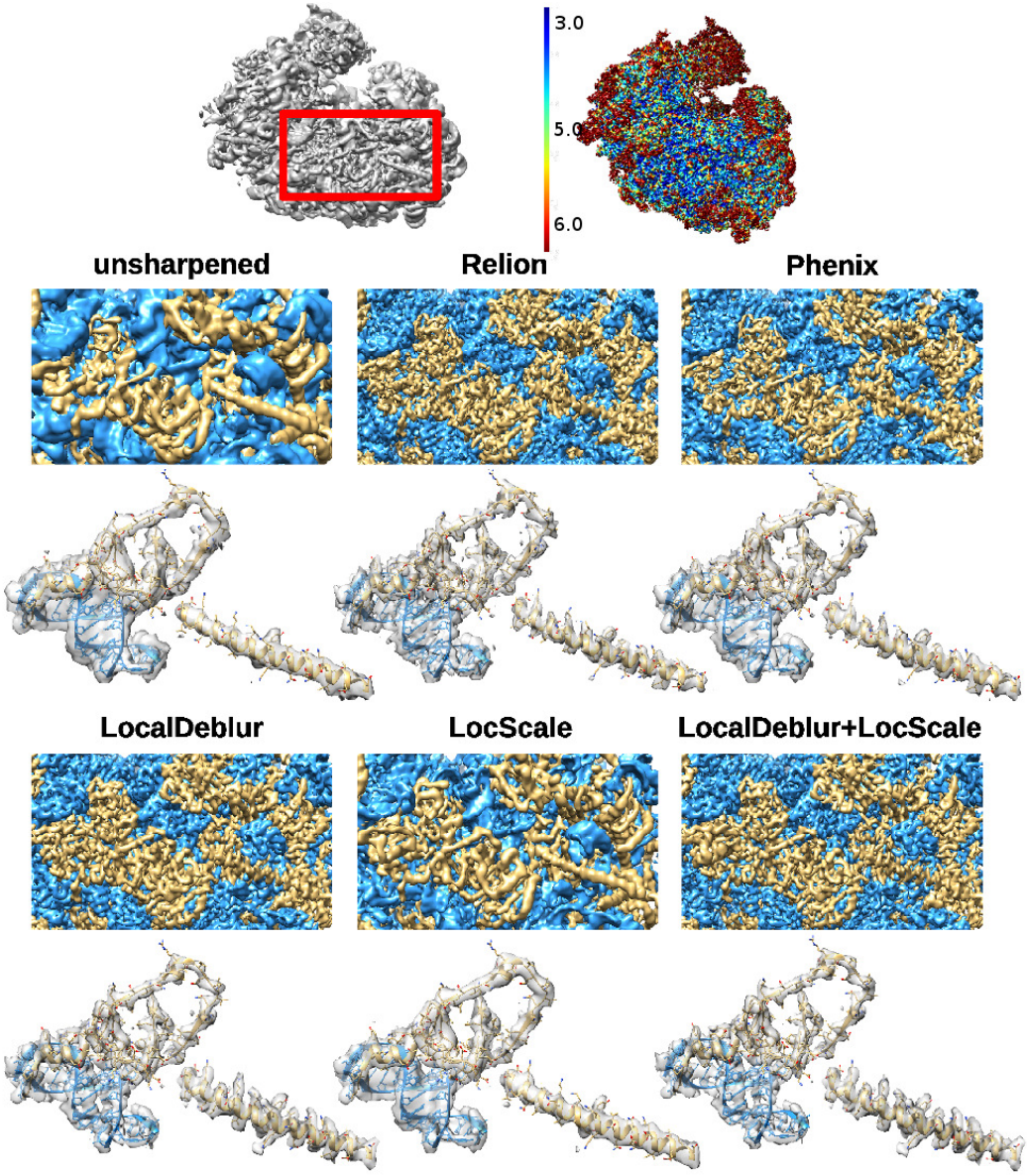

Figure 55: *Plasmodium falciparum* 80S ribosome and region of interest overlapped with the atomic model. The original map, its local resolution map (resolution in Å), and a region of the sharpened maps obtained with postprocessing - RELION, Autosharpen - Phenix, *LocalDeBlur*, *LocScale* and a combination of *LocalDeBlur* and *LocScale* are shown.

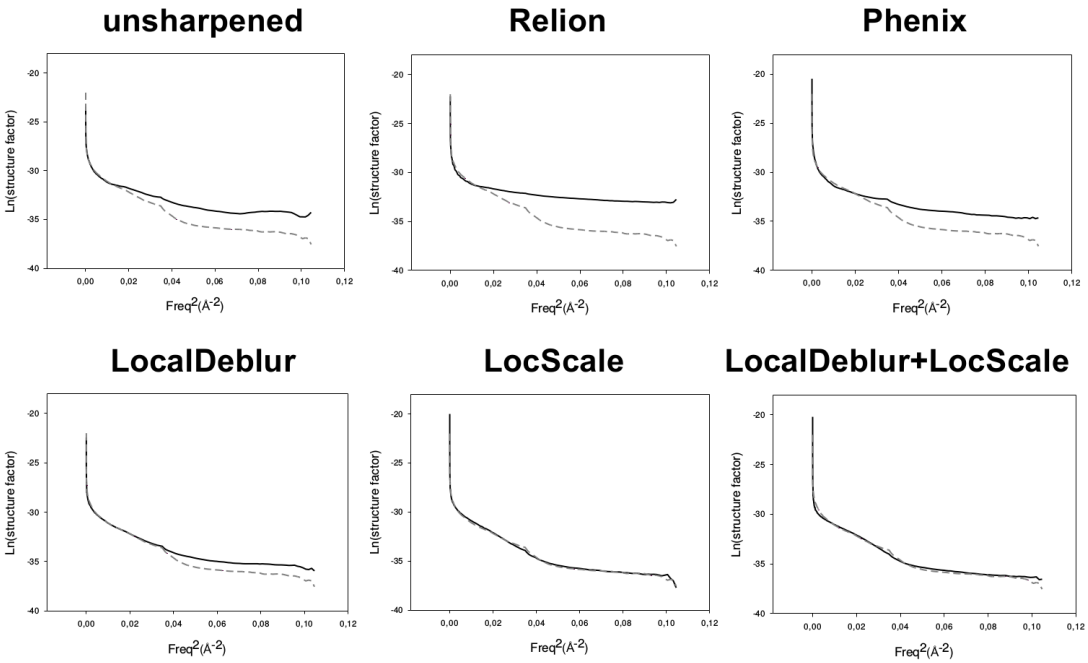

Figure 56: The Guinier plots for *Plasmodium falciparum* 80S ribosome of the original map and the sharpened maps obtained with postprocessing - RELION, Autosharpen - Phenix, *LocalDeBlur*, *LocScale* and a combination of *LocalDeBlur* and *LocScale* are shown (continuous line). In dashed line it is represented the Guinier for the atomic model converted into density map.

9 CONCLUSIONS

This section summarizes the main conclusions of the present thesis. The main objectives of the this work were exposed in the first chapter, and the conclusions will be presented following the same order and referencing them.

9.1 OBJECTIVE 1: MEASUREMENT OF LOCAL RESOLUTION IN SPA

- It was proposed a new local resolution method, *MonoRes*, for computing the local resolution in SPA. This method is fully automatic, and only requires a mask and the density map.
- The local resolution values provided by *MonoRes* consist on a measurement of the SNR. Resolution values are assigned when the local amplitude of the signal cannot be statistically distinguished (hypothesis test) from the amplitude of noise.
- *MonoRes* algorithm is invariant under B-factor correction. Sharpened map and original reconstruction present the same resolution. Sharpening only affects the visualization of the map.
- The proposed algorithm casts high accuracy on resolution estimation. This was confirmed with synthetic data.
- *MonoRes* allows to estimate the local resolution of large maps in short computational times.
- The proposed method is also able to locally filter maps by the local resolution values.

9.2 OBJECTIVE 2: MEASUREMENT OF LOCAL RESOLUTION IN ELECTRON TOMOGRAPHY

1. Up to our knowledge *MonoTomo* represents the first local resolution method in electron tomography.
2. The input of *MonoTomo* consists on two independent tomograms, usually obtained by odd-even splitting of the tilt series.
3. Splitting frames approach is introduced as a new method for splitting datasets.
4. The noise in electron tomograms is spatially dependent. *MonoTomo* uses the core of *MonoRes*adapted to handle this dependency.
5. *MonoTomo* might open new possibilities in the field, like subtomogram averaging, alignment, segmentation, sharpening or local filtering.

9.3 OBJECTIVE 3: MEASUREMENT OF LOCAL-DIRECTIONAL RESOLUTION AND LOCAL RESOLUTION ANISOTROPY IN SPA

1. The concept of local resolution has been generalized adding directionality: resolution is represented with a tensor instead of a scalar.
2. A fully algorithm, *MonoDir*, for computing the local-directional resolution and analyze the existence of anisotropy in SPA reconstructions has been proposed.
3. The resolution anisotropy provides information about the local and global map quality needing only the reconstructed map for that.
4. Angular assignment error can be identified as a slope in the radial average of the local resolution maps, in particular in the radial and tangential components.
5. The existence of preferred directions and global anisotropy are shown as a non uniform coverage of the angular plot.

9.4 OBJECTIVE 4: ANALYSIS OF B-FACTOR CORRECTION

1. Guinier law is only valid at very low frequencies, the theory predicts an exponential decay of the structure factor.
2. Up to our knowledge, the first analytic calculation of Guinier's validity range has been presented.
3. The Porod law is only valid at high resolutions, and also predicts a decay of the structure factor.
4. The structure factor of a macromolecule has a non-flat spectrum. This fact is independent of the spectral range.
5. The b-factor correction only contributes to the visualization of the macromolecular complex.
6. Map with flattening/quasi-flattening sprectra does not represent macromolecular complexes. Proteins must present a decay in the structure factor.
7. 3D Reconstruction algorithms do not always over-dampen the spectrum of the macromolecule.

9.5 OBJECTIVE 5: DEVELOPMENT OF A LOCAL SHARPENING METHOD

1. A new local sharpening method, *LocalDeBlur* based on local resolution has been developed.
2. Up to our knowledge, it represents the first application of local resolution (leaving out local filters).

3. *LocalDeBlur* avoids the b-factor quasi-flattening in agreement with the scattering theory.
4. It is compatible with *LocScale*, in fact the best results are achieved with combining both methods.
5. The proposed algorithm significantly improves the interpretability of density maps in terms of the of the map understanding and in the elucidation of atomic models.

9.6 OBJECTIVE 6: INTRODUCTION TO THE PHYSICS OF THE ELECTRON MICROSCOPE AND SPA WORKFLOW

This was a side objective, out of the scope of this thesis and was originated because of the multidisciplinarity of the field of cryoem: on the microscope side, it is required knowledge of Optics, Electromagnetism, Material Physics and Quantum Mechanics. On the other hand, the development of SPA methods needs knowledge of image processing, and mathematical background, but the microscope physics is also desirable. Part of the introduction of this thesis was conceived as a concise and enough detailed approach to Microscope Physics, the SPA assumptions and the current methods, in the search of the providing of a more accesible starting point to this document for the people who belong to one of the others disciplines involved in this topic. Usually all this information is spread in very specific books and reviews.

BIBLIOGRAPHY

[1] S. Banerjee D. Matthies X. Wu J.L.S. Milne S. Subramaniam A. Bartesaghi A. Merk. "2.2 A resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor." In: *Science* 348 (2015), pp. 1147–1151.

[2] "A negative staining method for high resolution electron microscopy of viruses." In: *Biochimica et Biophysica Acta* 34 (1959), pp. 103 –110.

[3] V. Abrishami, A. Zaldivar-Peraza, J. M. de la Rosa-Trevin, J. Vargas, J. Oton, R. Marabini, Y. Shkolnisky, J. M. Carazo, and C. O. S. Sorzano. "A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs." In: *Bioinformatics* 29.19 (2013), pp. 2460–2468.

[4] V. Abrishami, J. Vargas, X. Li, Y. Cheng, R. Marabini, C. O. S. Sorzano, and J. M. Carazo. "Alignment of direct detection device micrographs using a robust Optical Flow approach." In: *Journal of Structural Biology* 189 (2015), pp. 163–176.

[5] M. Adrian, J. Dubochet, J. Lepault, and A.W. McDowall. "Cryo-electron microscopy of viruses." In: *Nature* 308 (1984), pp. 32–36.

[6] P. Afanasyev et al. "A posteriori correction of camera characteristics from large image data sets." In: *Scientific Reports* 5.10317 (2015).

[7] L.A. Amos and J.T. Finch. "Aaron Klug and the revolution in biomolecular structure determination." In: *Trends in Cell Biology* 14 (3 2004), pp. 148 –152.

[8] X. Bai, I. S. Fernandez, G. McMullan, and S. Scheres. "Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles." In: *eLife* 2 (2013).

[9] X.C. Bai, G. McMullan, and Scheres S.H. "How cryoEM is revolutionizing structural biology." In: *Trends Biochem Sci.* 40 (2015), pp. 49–57.

[10] L.E. Ballentine. *Quantum Mechanics, a modern development*. 1999.

[11] A. Bartesaghi, D. Matthies, S. Banerjee, A. Merk, and S. Subramaniam. "Structure of beta-galactosidase at 3.2-A resolution obtained by cryo-electron microscopy." In: *Proceedings of the National Academy of Sciences* 111.32 (2014), pp. 11709–11714.

[12] J. Baumgardner and P.O. Frederickson. "Icosahedral Discretization of the Two-Sphere." In: *SIAM J. Numer. Anal.* 22 (1985), pp. 1107–1115.

[13] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. "The Protein Data Bank." In: *Nucleic Acids Research* 28.1 (2000), pp. 235–242. URL: https://www.rcsb.org/.

[14] N. Bershad and A. Rockmore. "On estimating signal-to-noise ratio using the sample correlation coefficient." In: *IEEE Transactions on Information Theory* 20 (1 1974), pp. 112–113.

[15] T. Bhamre, T. Zhang, and A. Singer. "Denoising and covariance estimation of single particle cryo-EM images." In: *Journal of Structural Biology* 195.1 (2016), pp. 72–81.

[16] J. Bogaerts, B. Dierickx, G. Meynants, and D. Uwaerts. "Total Dose and Displacement Damage Effects in a Radiation-Hardened CMOS APS." In: *IEEE Transactions on electron devices* 50.1 (2003), pp. 84–90.

[17] M. Born and E. Wolf. *Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 7th-edition, 1999.

[18] A.F. Brilot, J.Z. Chen, A. Cheng, J. Pan, S.C. Harrison, C.S. Potter, B. Carragher, R. Henderson, and N. Grigorieff. "Beam-induced motion of vitrified specimen on holey carbon film." In: *Journal of Structural Biology* 177.3 (2012), pp. 630–637.

[19] D.L. Dotson P. Uzdavinys S. Iwata M.S. Sansom C. von Ballmoos O. Beckstein D. Drew C. Lee S. Yashiro and A.D. Cameron. "Systematic Analyses of Substrate Preferences of 20S Proteasomes Using Peptidic Epoxyketone Inhibitors." In: *J. Gen. Physiol* 1444 (2014), pp. 529–544.

[20] M.G. Campbell, D. Veesler, A. Cheng, and B. Carragher. "2.8 Angstrom resolution reconstruction of the Thermoplasma acidophilum 20S proteasome using cryo-electron microscopy." In: *eLife* 4 (2015), e06380.

[21] G. Cardone, K. Grunewald, and A.C. Steven. "A resolution criterion for electron tomography based on cross-validation." In: *Journal of Structural Biology* 151 (2005), pp. 117–129.

[22] G. Cardone, J. B. Heymann, and A. C. Steven. "One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions." In: *Journal of Structural Biology* 184 (2013), pp. 226–236.

[23] Bridget Carragher, Nick Kisseberth, David Kriegman, Ronald A. Milligan, Clinton S. Potter, James Pulokas, and Amy Reilein. "Leginon: An Automated System for Acquisition of Images from Vitreous Ice Specimens." In: *Journal of Structural Biology* 132.1 (2000), pp. 33 –45.

[24] Y. W. Chang, L. A. Rettberg, A. Treuner-Lange, J. Iwasa, L. Sogaard-Andersen, and G. J. Jensen. "Architecture of the type IVa pilus machine." In: *Science* 351 (2016), p. 1165.

[25] Y.W. Chang, A. Kjaer, D.R. Ortega, G. Kovacikova, J.A. Sutherland, L.A. Rettberg, R.K. Taylor, and G.J. Jensen. "Architecture of the Vibrio cholerae toxin-coregulated pilus machine revealed by electron cryotomography." In: *Nature Microbiology* 2 (2017), p. 16269.

[26] A Chicano, E. Crosas, J. Oton, R. Melero, B.D. Engel, and J.R Daban. "Frozen-hydrated chromatin from metaphase chromosomes has an interdigitated multilayer structure." In: *EMBO Journal* (2019), e99769.

[27] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth. "On the Lambert W Function." In: *Adv. Comput.Math.* 5 (1996), pp. 329–359.

[28] M.G.S. Costa, P.R. Batista, P.M. Bisch, and D. Perahia. "Exploring Free Energy Landscapes of Large Conformational Changes: Molecular Dynamics with Excited Normal Modes." In: *Journal of Chemical Theory and Computation* 11.6 (2015), pp. 2755–2767.

[29] A. Doerr. "Single.particle cryo-electron microscopy." In: *Nature Methods* 13 (2016), p. 23.

[30] E. Driest and H.O. Muller. "Electron microscopic photographs (electron micrograms) of Chitin specimens." In: *Z. Wiss. Mikrosk* 52 (1935), pp. 53–57.

[31] J. Dubochet, J. Lepault, R. Freeman, J. A. Berriman, and J.-C. Homo. "Electron microscopy of frozen water and aqueous solutions." In: *Journal of Microscopy* 128.3 (1982), pp. 219–237.

[32] J. Dubochet, M. Adrian, J.J Chang, J.C Homo, J. Lepault, A.W. McDowall, and P. Schultz. "Cryo-electron microscopy of vitrified specimens." In: *Quarterly Review of Biophysics* 21.2 (1988), pp. 129–228.

[33] W. Heinemeyer G. Paniagua Soriano H.S. Overkleeft E.M. Huber G. de Bruin and M. Groll. "Systematic Analyses of Substrate Preferences of 20S Proteasomes Using Peptidic Epoxyketone Inhibitors." In: *J.Am.Chem.Soc* 137 (2015), pp. 7835–7842.

[34] M. Eisenstein. "The field that came in from the cold." In: *Nature Methods* 13 (2016), pp. 19–22.

[35] H. Elmlund, D. Elmlund, and S. Bengio. "PRIME: probabilistic initial 3D model generation for single-particle cryo-electron microscopy." In: *Structure* 21.8 (2013), pp. 1299–1306.

[36] G.Y. Fan and M.H. Ellisman. "Digital imaging in transmission electron microscopy." In: *Journal of Microscopy* 200 (2000), pp. 1–13.

[37] A.R. Faruqi and R. Henderson. "Electronic detector for electron microscopy." In: *Current Opinion in Structural Biology* 17.5 (2007), pp. 549–555.

[38] A.R. Faruqi, R. Henderson, and A. J. Holmes. "Radiation damage studies on STAR250 CMOS sensor at 300 keV for electron microscopy." In: *Nuclear Instruments and Methods in Physics Reseach A* 565 (2006), pp. 139–143.

[39] M. Felsberg and G. Sommer. "The Monogenic Signal." In: *IEEE Transactions on Signal Processing* 49 (2001), pp. 3136–3144.

[40] J.J. Fernandez. "Computational methods for electron tomography." In: *Micron* 43 (2012), pp. 1010–1030.

[41] J.J. Fernandez, D. Luque, J.R. Caston, and J.L. Carrascosa. "Sharpening high resolution information in single particle electron cryomicroscopy." In: *Journal of Structural Biology* 164 (2008), pp. 170–175.

[42] J. Frank. "Generalized single-particle cryo-EM - a historical perspective." In: *Microscopy* 65 (1 2016), pp. 3–8.

[43] J. Frank, Shimkin B., and H. Dowse. "SPIDER—A modular software system for electron image processing." In: *Ultramicroscopy* 6.4 (1981), pp. 343–357.

[44] J. Frank and M van Heel. "Correspondence-analysis of aligned images of biological particles." In: *Journal of Molecular Biology* 161 (1982), pp. 134–137.

[45] J. Frank, A. Verschoor, and M. Boublik. "Computer averaging of electron micrographs of 40S ribosomal subunits." In: *Science* 214 (1981), pp. 1353–1355.

[46] N. Voss A. Cheng D. Fellmann J. Pulokas C. Yoshioka C. Irving A. Mulder P. Lau D. Lyumkis C. Potter G. Lander S. Stagg and B. Carragher. "Appion: an integrated, database-driven pipeline to facilitate EM image processing." In: *Journal of Structural Biology* 166.1 (2009), pp. 95–102.

[47] D. Gabor. "Theory of Communication." In: *Proc IEE (London)* 93 (1946), pp. 429–457.

[48] J.J. Gil. "Polarimetric characterization of light and media." In: *European Physical Journal Applied Physics* 40 (2007), pp. 1–47.

[49] R.M. Glaeser. "Limitations to significant information in biological electron microscopy as a result of radiation damage." In: *Ultramicroscopy* 36.3-4 (1971), pp. 466–482.

[50] R.M. Glaeser. "Specimen Behavior in the Electron Beam." In: *Methods Enzymol.* 579 (2016), pp. 19–50.

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

[52] O. Glatter and O. Kratky. *Small Angle X-ray Scattering*. Academic Press, 1982.

[53] J. Gomez-Blanco et al. "Using Scipion for stream image processing at Cryo-EM facilities." In: *Journal of Structural Biology* 204.3 (2018), pp. 457 –463.

[54] J.W. Goodman. *Statistical Optics*. 2015.

[55] T. Grant and N. Grigorieff. "Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6." In: *eLife* 4 (2015), e06980.

[56] I. Greenberg and Y. Shkolnisky. "Common lines modeling for reference free Ab-initio reconstruction in cryo-EM." In: *Journal of Structural Biology* 200.2 (2017), pp. 106–117.

[57] N. Grigorieff. "Resolution measurement in structures derived from single particles." In: *Acta Crystallographica D* 56 (2000), pp. 1270–1277.

[58] N. Grigorieff. "FREALIGN: high-resolution refinement of single particle structures." In: *Journal of Structural Biology* 157 (2007), pp. 117–125.

[59] A. Guinier and G. Fournet. *Small-Angle Scattering of X-Rays*. John Wiley and Sons, 1955.

[60] M. Gur, J.D. Madura, and I. Bahar. "Global transitions of proteins explored by a multiscale hybrid methodology: application to adenylate kinase." In: *Biophysical Journal* 105 (7 2013), pp. 1643–1652.

[61] J.F. Hamilton and J.C. Marchant. “Image recording in electron microscopy.” In: *Journal of the optical society of America* 57.2 (1967), pp. 232–239.

[62] D. Haselbach, I. Komarov, D.E. Agafonov, K. Hartmuth, B. Graf, O. Dybkov, H. Urlaub, B. Kastner, R. Luhrmann, and H. Stark. “Structure and Conformational Dynamics of the Human Spliceosomal Bact Complex.” In: *Cell* 172 (3 2018), pp. 454–464.

[63] R.W. Hawkes and E. Kasper. *Principles of Electron Optics*. 1996.

[64] van Heel M. and Hollenberg J. *The stretching of distorted images of two-dimensional crystals*. 1980, pp. 256–260.

[65] M. van Heel. “Imagic and its results.” In: *Ultramicroscopy* 4 (1979), pp. 117–117.

[66] M. van Heel. “Similarity measures between images.” In: *Ultramicroscopy* 21.1 (1987), pp. 95–100.

[67] M. van Heel and J. Frank. “Use of multivariate statistics in analyzing the images of biological macromolecules.” In: *Ultramicroscopy* 6 (1981), pp. 187–194.

[68] M. van Heel and W. Keegstra. “Imagic - a fast, flexible and friendly image-analysis software system.” In: *Ultramicroscopy* 7 (1981), pp. 113–130.

[69] M. van Heel, R. Portugal, and M. Schatz. *Multivariate Statistical Analysis in Single Particle (Cryo) Electron Microscopy*. 2009.

[70] A. Heimowitz, J. Anden, and A. Singer. “APPLE picker: Automatic particle picking, a low-effort cryo-EM framework.” In: *Journal of Structural Biology* 204.2 (2018), pp. 215 –227.

[71] R. Henderson. “The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules.” In: *Quarterly Reviews of Biophysics* 28.2 (1995), pp. 171–193.

[72] R. Henderson. “Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise.” In: *PNAS* 110.45 (2013), pp. 18037–18041.

[73] R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckmann, and H.K. Downing. “Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy.” In: *Journal of Molecular Biology* 213 (1990), pp. 899–929.

[74] R. Henderson, S. Chen, J.Z. Chen, N. Grigorieff, L.A. Passmore, L. Ciccarelli, J.L. Rubinstein, R.A. Crowther, P.L. Stewart, and P.B. Rosenthal. “Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy.” In: *Journal of Molecular Biology* 413.5 (2011), pp. 1028–1046.

[75] B. Heymann. “Validation of 3DEM Reconstructions: The phantom in the noise.” In: *AIMS Biophysics* 2 (2015), pp. 21–35.

[76] J.B. Heymann. “Bsoft: Image and molecular processing in electron microscopy.” In: *Journal of Structural Biology* 133 (2001), pp. 156–169.

[77] J.B. Heymann and D.M. Belnap. "Bsoft: Image and molecular processing in electron microscopy." In: *Journal of Structural Biology* 157.1 (2007), pp. 3–18.

[78] T. V. Hoang, X. Cavin, P. Schultz, and D. W. Ritchie. "gEMpicker: A highly parallel GPU-accelerated particle picking tool for cryo-electron microscopy." In: *BMC Structural Biology* 13.1 (2013), p. 25.

[79] G.C. Holst. *CCD arrays cameras and displays*. 1967.

[80] *J. Kai Zhang – Gautomatch*. http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/. Accessed: 2018-05-16.

[81] A.J. Jakobi, M. Wilmanns, and C. Sachse. "Model-based local density sharpening of cryo-EM maps." In: *elife* 6 (2017), e27131.

[82] C. Jiko, K.M. Davies, K. Shinzawa-Itoh, K. Tani, S. Maeda, D.J. Mills, T. Tsukihara, Y. Fujiyoshi, W. Kuhlbrandt, and C Gerle. "Bovine F_1F_o ATP synthase monomers bend the lipid bilayer in 2D membrane crystals." In: *eLife* 4 (2015), e06119.

[83] J. Johnson. "Analysis of Image Forming Systems." In: *Image Intensifier Symposium* (1958), pp. 244–273.

[84] D. Kimanius, B.O. Forsberg, E. Lindahl, and S.H.W. Scheres. "Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2." In: *elife* 5 (2016), e18722.

[85] Dari Kimanius, Björn O Forsberg, Sjors HW Scheres, and Erik Lindahl. "Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2." In: *eLife* 5 (2016), e18722.

[86] E. J. Kirkland. *Advanced Computing in Electron Microscopy*. Springer, 2010.

[87] C. Kittel. *Introduction To Solid State Physics*. 2005.

[88] A. Klug. "From Macromolecules to Biological Assemblies." In: *Nobel Lecture* (1983).

[89] J.R. Kremer, D.N. Mastronarde, and J.R. McIntosh. "Computer Visualization of Three-Dimensional Image Data Using IMOD." In: *Journal of Structural Biology* 166 (1 1996), pp. 71–76.

[90] A. Kucukelbir, F. J. Sigworh, and H. D. Tagare. "Quantifying the local resolution of cry-EM density maps." In: *Nature Methods* 11 (2014), pp. 63–65.

[91] M. Kudryashev. *Cellular Imaging*. Springer, 2018. Chap. Chapter 10: Resolution in Electron Tomography.

[92] W. Kuhlbrandt. "The resolution revolution." In: *Science* 343.6178 (2014), pp. 1443–1444.

[93] Z. Kurkcuoglu, I. Bahar, and P. Doruker. "ClustENM: ENM-Based Sampling of Essential Conformational Space at Full Atomic Resolution." In: *Journal of Chemical Theory and Computation* 12 (9 2016), pp. 4549–4562.

[94] K. G. Larkin, D. J. Bone, and M. A. Oldfield. "Natural demodulation of two-dimensional fringe patterns. I. General background of the spiral phase quadrature transform." In: *Journal of the Optical Society of America A* 18 (2001), pp. 1862–1870.

[95] C.L. Lawson et al. "EMDataBank unified data resource for 3DEM." In: *Nucleic Acids Research* 44.D1 (2016), pp. D396–D403. URL: http://www-ebi.emdatabank.org/.

[96] A. Leschziner. "Chapter Nine - The Orthogonal Tilt Reconstruction Method." In: *Cryo-EM, Part B: 3-D Reconstruction*. Ed. by G.J. Jensen. Vol. 482. Methods in Enzymology. Academic Press, 2010, pp. 237 –262.

[97] X. Li, P. Mooney, S. Zheng, C.R. Booth, M.B. Braunfeld, S. Gubbens, D.A. Agard, and Y. Cheng. "Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-em." In: *Nature Methods* 10.6 (2013), pp. 584–590.

[98] S.J. Ludtke, P.R. Baldwin, and W. Chiu. "EMAN: semiautomated software for high-resolution single-particle reconstructions." In: *Journal of Structural Biology* 128 (1999), pp. 82 –97.

[99] D. Julius Y. Cheng M. Liao E. Cao. "Structure of the TRPV1 ion channel determined by electron cryo-microscopy." In: *Nature* 504 (2013), pp. 107–112.

[100] D.J.C. Mackay. *Information Theory, Inference, and Learning Algorithms*. 2004.

[101] V.N. Mahajan. *Aberration theory made simple*. SPIE, 2011.

[102] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. 1995.

[103] B.R. Masters. *History of the Electron Microscope in Cell Biology*. 2009.

[104] D.N. Mastronarde. "SerialEM: A program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position." In: *Microscopy and Microanalysis* 9.2 (2003), p. 1182.

[105] D.N. Mastronarde. "Automated electron microscope tomography using robust prediction of specimen movements." In: *Journal of Structural Biology* 152 (2005), pp. 36–51.

[106] J. C Maxwell. "On the General Laws of Optical Instruments." In: *Quart. J. of Pure and Appl. Math* (1858), pp. 233–246.

[107] G. McMullan, A.R. Faruqi, and R. Henderson. "Direct Electron Detectors." In: *Methods in Enzymology* 579 (2016), pp. 1–17.

[108] G. McMullan, S. Chen, R. Henderson, and A.R. Faruqi. "Detective quantum efficiency of electron area detectors in electron microscopy." In: *Ultramicroscopy* 109 (2009), pp. 1126–1143.

[109] Alan Merk et al. "Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery." In: *Cell* 165.7 (2016), pp. 1698 –1707.

[110] J.A. Mindell and N. Grigorieff. "Accurate determination of local defocus and specimen tilt in electron microscopy." In: *Journal of Structural Biology* 142 (2003), pp. 334–347.

[111] Katerina Naydenova and Christopher J Russo. "Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy." In: *Nature Communications* 8 (1 2017), p. 629.

[112] R.G. Newton. *Scattering Theory of Waves and Particles*. 2013.

[113] William V. Nicholson and Robert M. Glaeser. "Review: Automatic Particle Detection in Electron Microscopy." In: *Journal of Structural Biology* 133.2 (2001), pp. 90 –101.

[114] Eva Nogales. "The development of cryo-EM into a mainstream structural biology technique." In: *Nature Methods* 13.1 (2016), pp. 24–27.

[115] L. W. Nordheim. "The Effect of the Image Force on the Emission and Reflexion of Electrons by Metals." In: *Proc. R. Soc. Lond. A* 121 (1928), pp. 626–639.

[116] R. K. Pathria and P. D. Beale. *Statistical Mechanics*. 2011.

[117] M. Paur, B. Stoklasa, Z. Hradil, L.L. Sanchez-Soto, and J. Rehacek. "Achieving the ultimate optical resolution." In: *Optica* 3.10 (2016), pp. 1144 –1147.

[118] M. Paur, B. Stoklasa, J. Grover, A. Krzic, L.L. Sanchez-Soto, Z. Hradil, and J. Rehacek. "Tempering Rayleigh's curse with PSF shaping." In: *Optica* 5.10 (2018), pp. 1177 –1180.

[119] P. A. Penczek. "Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms." In: *Journal of Structural Biology* 138 (2002), pp. 36–46.

[120] P. Penczek. "Resolution measures in molecular electron microscopy." In: *Methods Enzymol* 482 (2010), pp. 73–100.

[121] P. Penczek and J. Frank. *Electron Tomography: Methods for three-dimensional visualization of structures in the cell*. 2nd. Springer, 2006. Chap. Resolution in electron tomography.

[122] P. Penczek, M. Radermacher, and J. Frank. "Three-dimensional reconstruction of single particles embedded in ice." In: *Ultramicroscopy* 40.1 (1992), pp. 33 –53.

[123] P.A. Penczek, R.A. Grassucci, and J. Frank. "The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles." In: *Ultramicroscopy* 53 (1994), pp. 251–270.

[124] A. Potamianos and P. Maragos. "A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation." In: *Signal processing* 37 (1994), pp. 95–120.

[125] G. Pragier, I. Greenberg, X. Cheng, and Y. Shkolnisky. "A Graph Partitioning Approach to Simultaneous Angular Reconstitution." In: *IEEE Transactions on Computational Imaging* 2 (3 2016), pp. 323–334.

[126] A. Punjani, J.L. Rubinstein, D.J. Fleet, and M. Brubaker. "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." In: *Nature Methods* 14 (2017), pp. 290 –296.

[127] M. Radermacher and W. Hoppe. "Properties of 3-D reconstruction from projections by conical tilting compared to single axis tilting." In: *Proc. Seventh European Congress on Electron Microscopy* I (1980), pp. 132–133.

[128] M. Radermacher, R. Wagenknecht, A. Verschoor, and J. Frank. "3-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50s-Ribosomal subunit of Escherichia-Coli." In: *Journal of Microscopy-Oxford* 146 (1987), pp. 113–136.

[129] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 779–788.

[130] L. Reimer. *Transmission Electron Microscope. Physics of Image Formation and Microanalysis*. Springer Verlag, 1989.

[131] Z A Ripstein and J L Rubinstein. "Processing of Cryo-EM Movie Data." In: *Methods in Enzymology* 579 (2016), pp. 103–124.

[132] P.T.E. Roberts, J:N: Chapman, and A.M. MacLeod. "A CCD-Based image recording system for the CTEM." In: *Ultramicroscopy* 8 (1982), pp. 385–396.

[133] A. Rohou and N. Grigorieff. "CTFFIND4: Fast and accurate defocus estimation from electron micrographs." In: *Journal of Structural Biology* 192.2 (2015), pp. 216–221.

[134] J.M. de la Rosa-Trevin, J. Oton, R. Marabini, A. Zaldivar, J. Vargas, J.M. Carazo, and C.O.S. Sorzano. "Xmipp 3.0: An improved software suite for image processing in electron microscopy." In: *Journal of Structural Biology* 184.1 (2013), pp. 321–328.

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

[136] John L. Rubinstein and Marcus A. Brubaker. "Alignment of cryo-EM movies of individual particles by optimization of image translations." In: *Journal of Structural Biology* 192.2 (2015), pp. 188–195.

[137] E. Ruska. "The Development of the Electron Microscope and of Electron Microscopy (Nobel Lecture)." In: *Angewandte Chemie* 26, 7 (1987), pp. 595–605.

[138] E. Ruska and M. Knoll. "Das Elektronenmikrokop." In: *Zeitschrift fur Physik* 78, 5 (1932), pp. 318–339.

[139] C. J. Russo and L. A. Passmore. "Robust evaluation of 3D electron cryomicroscopy data using tilt-pairs." In: *Journal of Structural Biology* 187.2 (2014), pp. 112–118.

[140] C.J. Russo and R. Henderson. "Charge accumulation in electron cryomicroscopy." In: *Ultramicroscopy* 187 (2018), pp. 43 –49.

[141] C.J. Russo and L.A. Passmore. "Progress towards an optimal specimen support for electron cryomicroscopy." In: *Current Opinion in Structural Biology* 37 (2016), pp. 81 –89.

[142] C.J. Russo and L.A. Passmore. "Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens." In: *Journal of Structural Biology* 193.1 (2016), pp. 33 –44.

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

[144] W.O. Saxton. *Computer Techniques for Image Processing of Electron Microscopy*. Academic Press, 1978.

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

[146] S.H. Scheres. "Beam-induced motion correction for sub-megadalton cryo-EM particles." In: *eLife* 3 (2014), e03665.

[147] S.H.W. Scheres. "Chapter Eleven - Classification of Structural Heterogeneity by Maximum-Likelihood Methods." In: *Cryo-EM, Part B: 3-D Reconstruction*. Ed. by G.J. Jensen. Vol. 482. Methods in Enzymology. Academic Press, 2010, pp. 295 –320.

[148] S.H.W. Scheres and S. Chen. "Prevention of overfitting in cryo-EM structure determination." In: *Nature Methods* 9 (2012), pp. 853–854.

[149] Sjors H W. Scheres. "Semi-automated selection of cryo-EM particles in RELION-1.3." In: *Journal of Structural Biology* 189.2 (2014).

[150] Sjors H.W. Scheres, Mikel Valle, Rafael Nuñez, Carlos O.S. Sorzano, Roberto Marabini, Gabor T. Herman, and Jose-Maria Carazo. "Maximum-likelihood Multi-reference Refinement for Electron Microscopy Images." In: *Journal of Molecular Biology* 348.1 (2005), pp. 139 –149.

[151] W. Schottky. "Über kalte und warme Elektronenentladungen." In: *Zeitschrift für Physik A* 63 (1923), pp. 63–106.

[152] F.K.M. Schur, M. Obr, W.J.H. Hagen, W. Wan, A.J. Jakobi, J.M. Kirkpatrick, C. Sachse, H.G. Krausslich, and J.A.G. Briggs. "An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation." In: *Science* 353.6298 (2016), pp. 506–508.

[153] M. Shatsky, R.J. Hall, S.E. Brenner, and R.M. Glaeser. "A method for the alignment of heterogeneous macromolecules from electron microscopy." In: *Journal of Structural Biology* 166.1 (2009), pp. 67 –78.

[154] Y. Shkolnisky and A. Singer. "Viewing direction estimation in cryo-EM using synchronization." In: *SIAM Journal on Imaging Sciences* 5.3 (2012), pp. 1088–1110.

[155] F.J. Sigworth. "A maximum-likelihood approach to single-particle image refinement." In: *Journal of Structural Biology* 122 (2003), pp. 329–339.

[156] F.J. Sigworth, P.C. Doerschuk, J.M. Carazo, and S.H.W. Scheres. "Chapter Ten - An Introduction to Maximum-Likelihood Methods in Cryo-EM." In: *Cryo-EM, Part B: 3-D Reconstruction*. Ed. by G.J. Jensen. Vol. 482. Methods in Enzymology. Academic Press, 2010, pp. 263 –294.

[157] C. O. S. Sorzano, R. Marabini, N. Boisset, E. Rietzel, R. Schröder, G. T. Herman, and J. M. Carazo. "The effect of overabundant projection directions on 3D reconstruction algorithms." In: *Journal of Structural Biology* 133 (2001), pp. 108–118.

[158] C. O S Sorzano, J. R. Bilbao-Castro, Y. Shkolnisky, M. Alcorlo, R. Melero, G. Caffarena-Fernández, M. Li, G. Xu, R. Marabini, and J. M. Carazo. "A clustering approach to multireference alignment of single-particle projections in electron microscopy." In: *Journal of Structural Biology* 171.2 (2010), pp. 197–206.

[159] C O S. Sorzano, J. Vargas, J. M. de la Rosa-Trevín, J. Otón, A. L. Álvarez-Cabrera, V. Abrishami, E. Sesmero, R. Marabini, and J. M. Carazo. "A statistical approach to the initial volume problem in Single Particle Analysis by Electron Microscopy." In: *Journal of Structural Biology* 189.3 (2015), pp. 213–219.

[160] C O S. Sorzano et al. "Cryo-EM and the elucidation of new macromolecular structures: Random Conical Tilt revisited." In: *Scientific Reports* 5.14290 (2015), pp. 1–6.

[161] C. O. S. Sorzano, J. Vargas, J. Otón, V. Abrishami, J. M. de la Rosa Trevín, S. del Riego, A. Fernández-Alderete, C. Martínez-Rey, R. Marabini, and J. M. Carazo. "Fast and accurate conversion of atomic models into electron density maps." In: *AIMS Biophysics* 2 (2015), pp. 8–20.

[162] C.O.S. Sorzano, R. Marabini, J. Velazquez-Muriel, J.R. Bilbao-Castro, S.H.W. Scheres, and Pascual-Montano A. Carazo J.M. and. "XMIPP: a new generation of an open-source image processing package for electron microscopy." In: *Journal of Structural Biology* 148.2 (2004), pp. 194–204.

[163] C.O.S. Sorzano, S. Jonic, R. Nunez-Ramirez, N. Boisset, and J.M. Carazo. "Fast, robust, and accurate determination of transmissionelectron microscopy contrast transfer function." In: *Journal of Structural Biology* 160 (2007), pp. 249–262.

[164] C.O.S. Sorzano, A.L. Alvarez-Cabrera, M. Kazemi, J.M. Carazo, and S. Jonić. "StructMap: Elastic Distance Analysis of Electron Microscopy Maps for Studying Conformational Changes." In: *Biophysical Journal* 110.8 (2016), pp. 1753–1765.

[165] C.O.S. Sorzano et al. "A Survey of the Use of Iterative Reconstruction Algorithms in Electron Microscopy." In: *BioMed Research International* 6482567 (2017).

[166] C.O.S. Sorzano, J. Vargas, J. Oton, V. Abrishami, J.M. de la Rosa-Trevin, J. Gomez-Blanco, J.L. Vilas, R Marabini, and J.M. Carazo. "A review of resolution measures and related aspects in 3D Electron Microscopy." In: *Progress in Biophysics and Molecular Biology* 124 (2017), pp. 1–30.

[167] C.O.S. Sorzano et al. "A new algorithm for high-resolution reconstruction of single particles by electron microscopy." In: *Journal of Structural Biology* 204.2 (2018), pp. 329–337.

[168] C.O.S. Sorzano et al. "Blind estimation of DED camera gain in Electron Microscopy." In: *Journal of Structural Biology* 203.2 (2018), pp. 90–93.

[169] C.O.S. Sorzano et al. "Swarm optimization as a consensus technique for Electron Microscopy Initial Volume." In: *Applied Analysis and Optimization* 2 (2018), pp. 299–313.

[170] J.C.H. Spence and J.M. Zuo. "Large dynamic range, parallel detection system for electron diffraction and Imaging." In: *Review of Scientific Instruments* 59.9 (1988), pp. 2101–2105.

[171] M.T. Swulius and G.J. Jensen. "The Helical MreB Cytoskeleton in *Escherichia coli* MC1000/pLE7 Is an Artifact of the N-Terminal Yellow Fluorescent Protein Tag." In: *Journal of Bacteriology* 194 (2012), pp. 6382–6386.

[172] Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, and D. Lyumkis. "Addressing preferred specimen orientation in single-particle cryo-EM through tilting." In: *Nature Methods* 14 (2017), pp. 793–796.

[173] Y.Z. Tan, P.R Baldwin, J.H. Davis, J.R Williamson, C.S. Potter, B. Carragher, and D. Lyumkis. "Addressing preferred specimen orientation in single-particle cryo-EM through tilting." In: *Nature Methods* 14.8 (2017), pp. 793–796.

[174] G. Tang, L. Peng, P.R. Baldwin, D.S. Mann, W. Jiang, I. Rees, and S.J. Ludtke. "EMAN2: An extensible image processing suite for electron microscopy." In: *Journal of Structural Biology* 157.1 (2007), pp. 38 –46.

[175] K.A. Taylor and R.M. Glaeser. "Electron-diffraction of frozen, hydrated protein crystals." In: *Science* 186 (1974), pp. 1036 –1037.

[176] K.A. Taylor and R.M. Glaeser. "Electron-microscopy of frozen hydrated biological specimens." In: *Journal of Ultrastructure Research* 55 (1976), pp. 448 –456.

[177] Kenneth A. Taylor and Robert M. Glaeser. "Electron microscopy of frozen hydrated biological specimens." In: *Journal of Ultrastructure Research* 55.3 (1976), pp. 448 –456.

[178] T.C. Terwilliger, O.V. Sobolev, P.V Afonine, and Adams P.D. "Automated map sharpening by maximization of detail and connectivityg." In: *IUCr* 74.6 (2019), pp. 545–559.

[179] F. Thon. "Zur defokussierunsabhaengigkeit des phasenkontrastes bei der elektronenmikroskopischen abbildung." In: *Z. Naturforsch* 21 (1966), pp. 476–478.

[180] M. Unser, D. Sage, and D. Van De Ville. "Multiresolution Monogenic Signal Analysis Using the Riesz-Laplace Wavelet Transform." In: *IEEE Transactions on Image Processing* 18 (2009), pp. 2402–2418.

[181] M. Unser, B. Trus, and A. Steven. "A new resolution criterion based on spectral signal-to-noise ratio." In: *Ultramicroscopy* 23 (1982), pp. 39–52.

[182] P.N.T. Unwin and R. Henderson. "Molecular-structure determination by electron-microscopy of unstained crystalline specimen." In: *Journal of Molecular Biology* 94 (1975), pp. 425 –440.

[183] M. Van Heel. "Angular reconstitution: A posteriori assignment of projection directions for 3D reconstruction." In: *Ultramicroscopy* 21 (1987), pp. 111–123.

[184] J. Vargas, J. Otón, R. Marabini, S. Jonic, J.M. de la Rosa-Trevín, J.M. Carazo, and C.O.S. Sorzano. "FASTDEF: Fast defocus and astigmatism estimation for high-throughput transmission electron microscopy." In: *Journal of Structural Biology* 181 (2013), pp. 136–148.

[185] J. Vargas, V. Abrishami, R. Marabini, J.M. de la Rosa-Trevín, A. Zaldivar, J.M. Carazo, and C.O.S. Sorzano. "Particle quality assessment and sorting for automatic and semiautomatic particle-picking techniques." In: *Journal of Structural Biology* 183 (2013), pp. 342–353.

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

[187] J. Vargas, J. Otón, R. Marabini, J. M. Carazo, and C. O. S. Sorzano. "Particle alignment reliability in single particle electron cryomicroscopy: a general approach." In: *Scientific Reports* 6 (2016), p. 21626.

[188] J. Vargas, R. Melero, J. Gómez-Blanco, J.M. Carazo, and C.O.S. Sorzano. "Quantitative analysis of 3D alignment quality: its impact on soft-validation, particle pruning and homogeneity analysis." In: *Scientific Reports* 7 (1 2017), p. 6307.

[189] J. L. Vilas et al. "MonoRes: automatic and accurate estimation of Local Resolution for electron microscopy Maps." In: *Structure* 26 (2018), pp. 337–344.

[190] J.L. Vilas et al. "Re-examining the spectra of macromolecules. Limits of the Guinier approximation and of the current practice of spectral quasi B-factor flattening." In: *submitted* xx.xx (2019), p. xx.

[191] K.R. Vinothkumar and R. Henderson. "Single particle electron cryomicroscopy: trends, issues and future perspective." In: *Quarterly reviews of biophysics* 49 (2016), pp. 1–25.

[192] N.R. Voss, C.K. Yoshioka, M. Radermacher, C.S. Potter, and B. Carragher. "DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy." In: *Journal of Structural Biology* 166.2 (2009), pp. 205–213.

[193] M. Vulovic, L.M. Voortman, L.J. van Vliet, and B. Rieger. "When to use the projection assumption and the weak-phase object approximation in phase contrast cryo-EM." In: *Ultramicroscopy* 136 (2014), pp. 61–66.

[194] T. Wagner, F. Merino, M. Stabrin, T. Moriya, C. Gatsogiannis, and S. Raunser. "SPHIRE-crYOLO: A fast and well-centering automated particle picker for cryo-EM." In: *biorxiv* (2018).

[195] Feng Wang, Huichao Gong, Gaochao Liu, Meijing Li, Chuangye Yan, Tian Xia, Xueming Li, and Jianyang Zeng. "DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM." In: *Journal of Structural Biology* 195.3 (2016), pp. 325–336.

[196] L. Wang, A. Singer, and Z. Wen. "Orientation determination of cryo-EM images using least unsquared deviations." In: *SIAM Journal on Imaging Sciences* 6.4 (2013), pp. 2450–2483.

[197] Sebastian Wasilewski and Peter B. Rosenthal. "Web server for tilt-pair validation of single particle maps from electron cryomicroscopy." In: *Journal of Structural Biology* 186.1 (2014), pp. 122–131.

[198] W. Wong, X. Bai, A. Brown, I.S. Fernandez, E. Hanssen, M. Condron, Y.H. Tan, J. Baum, and S. Scheres. "Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine." In: *eLife* 3 (2014), e03080.

[199] K. Zhang. "Gctf: real-time CTF determination and correction." In: *Journal of Structural Biology* 193 (2016), pp. 1–12.

[200] X. Zhang, Q. Fang, W. H. Hui, and Z. Hong Zou. "3.3 Å Cryo-EM Structure of a Nonenveloped Virus Reveals a Priming Mechanism for Cell Entry." In: *Cell* 141 (2010), pp. 472–482.

[201] S.R. Zhao and H. Halling. "A new Fourier method for fan beam reconstruction." In: *1995 IEEE Nuclear Science Symposium and Medical Imaging Conference Record*. IEEE. 1995, pp. 1287–1291.

[202] S.Q. Zheng, E. Palovcak, J.P. Armache, K.A. Verba, Y. Cheng, and D.A. Agard. "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy." In: *Nature Methods* 14 (2017), pp. 331–332.

[203] Shawn Q Zheng, Eugene Palovcak, Jean-Paul Armache, Kliment A Verba, Yifan Cheng, and David A Agard. "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy." In: *Nature Methods* 14 (2017), pp. 331–332.

[204] Yanan Zhu, Qi Ouyang, and Youdong Mao. "A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy." In: *BMC Bioinformatics* 18.1 (2017).

[205] Z. Zivanov, N. Nakane, D. Forsberg B.O. Kimanius, W.J.H Hagen, E. Lindahl, and S.H.W. Scheres. "New tools for automated high-resolution cryo-EM structure determination in RELION-3." In: *elife* 7 (2018), e42166.

[206] S.H.W. cheres. "RELION: Implementation of a Bayesian approach to cryo-EM structure determination." In: *Journal of Structural Biology* 180.3 (2012), pp. 519–530.

[207] J. M. de la Rosa-Trevín et al. "Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy." In: *Journal of Structural Biology* 195.1 (2016), pp. 93–99.

[208] M. van Heel. "Finding trimeric HIV-1 envelope glycoproteins in random noise." In: *PNAS* 110.45 (2013), E4175–E4177.

[209] M. van Heel and M. Schatz. "Fourier shell correlation threshold criteria." In: *Journal of Structural Biology* 151.3 (2005), pp. 250 –262.