CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	2D image classification and alignment in single-particle Cryo-EM
Authors	ZiJian, Bai
Publication date	2024
Original Citation	Bai, Z. 2024. 2D image classification and alignment in single-particle Cryo-EM. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2024, Zijian Bai. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-11-02 15:41:25
Item downloaded from	https://hdl.handle.net/10468/16961

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

2D Image Classification and Alignment in Single-Particle Cryo-EM

ZiJian Bai



NATIONAL UNIVERSITY OF IRELAND, CORK
DEPARTMENT OF STATISTICS

Thesis submitted for the degree of
Doctor of Philosophy

Head of Department: Dr. Michael Cronin

Supervisors: **Dr. Jian Huang**
Prof. Finbarr O'Sullivan

Date: December 16, 2024

Contents

List of Figures 5
List of Tables 8
List of Algorithms 9
Acknowledgements 11
Abstract 13

1 Introduction **17**
1.1 Motivation 17
1.2 Main Objectives 18
1.3 Outline and Contribution 19

2 Background and Literature review **21**
2.1 Structural biology 21
2.2 Cryogenic Electron Microscopy (Cryo-EM) 22
2.2.1 Imaging Process of Cryo-EM 27
2.3 Single-particle Cryo-EM 30
2.3.1 Development Timeline of Single-Particle Cryo-EM 30
2.3.2 Imaging Principle of Single-Particle Cryo-EM 31
2.4 2D Classification 33
2.5 2D Alignment 36
2.6 Applications of Fourier transform 37
2.7 Applications of Non-uniform discrete Fourier transform 39
2.8 Applications of Spectral Clustering 40

3 NUDFT based 2D classification **43**
3.1 Introduction 43

3.2 Methodology . . . 44
3.2.1 Non-uniform discrete Fourier transform . . . 44
3.2.2 Rotation-invariant information . . . 45
3.2.3 Principle component analysis . . . 49
3.2.4 K-means clustering . . . 50
3.2.5 Experimental studies . . . 51
3.3 Experimental result . . . 52
3.3.1 Synthetic data generation . . . 52
3.3.2 PCA and K-means classification result examination . . . 54
3.3.3 Final result comparison . . . 57
3.4 Summary and discussion . . . 61

4 Spectral clustering based on NUDFT features 63
4.1 Introduction . . . 63
4.2 Methodology . . . 64
4.2.1 Amplitude and power spectrum . . . 64
4.2.2 Spectral clustering . . . 65
4.2.3 Algorithms . . . 70
4.3 Experimental result . . . 74
4.3.1 Synthetic data generation . . . 74
4.3.2 Projection images generation . . . 75
4.3.3 Details in applying spectral clustering method . . . 77
4.3.4 Spectral clustering and K-means overall result examination 79
4.4 Conclusion and discussion . . . 80

5 NUDFT based 2D in-plane rotation angle estimation 85
5.1 Introduction . . . 85
5.2 Methodology . . . 87
5.2.1 Fourier alignment . . . 87
5.2.2 Algorithms . . . 87
5.3 Experimental result . . . 90
5.3.1 Synthetic data generation . . . 90
5.3.2 Projection images generation . . . 91

5.3.3 2D alignment performance comparison result 92
5.4 Summary and discussion 96

6 Spectral clustering featured NUDFT based 2D in-plane rotation angle estimation 99
6.1 Introduction 99
6.2 Methodology 100
6.2.1 Introducing spectral clustering 100
6.2.2 Introducing Z-score outlier detection 101
6.2.3 Algorithms 102
6.3 Experimental result 106
6.3.1 Projection images generation 106
6.3.2 Spectral clustering application result 109
6.3.3 Final result after Z-score method 111
6.3.4 Performance comparison 114
6.4 Conclusion and discussion 115

7 Discussion, conclusion and future work 117

List of Figures

2.1 An example of a complex 3D structure of protein 23
2.2 Timeline of key contributions to the field of 3D electron microscopy highlighted in this Commentary.[65] 24
2.3 Cumulative number of particle entries released in the last 17 years (statistical data comes from EMDataBank). 25
2.4 Comparison among different technologies in structural biology. . . 26
2.5 Typical Cryo-EM workflow. 28
2.6 The three main steps in the single particle analysis workflow. . . . 29
2.7 Timeline of key events in the development of single-particle Cryo-EM methods.[37] 30
2.8 A sample of single-particle Cryo-EM imaging process. 31
2.9 Image translation in single-particle Cryo-EM.[17] 34
2.10 A sample of the 2D Fourier transform's application on single-particle Cryo-EM projection images. 37
2.11 An example of image rotation alignment.[17] 38
2.12 A sample of spectral clustering using the idea of graph theory. . . 41

3.1 Different signal-to-noise ratios. 52
3.2 EMDB 1080 52
3.3 Constructed noise is similar with real noise. 53
3.4 Different classes of projection images are specified by different projection angles. 54
3.5 The cumulative explained variance by each PCs, the figure shows one case under SNR=0.1 55

3.6 Within-Cluster sum of squares under different numbers of clusters, the figure shows one case under SNR=0.1 57
3.7 Power Spectrum generated from NUDFT 58
3.8 Classification accuracy of different methods 60

4.1 Proposed algorithm pipeline 70
4.2 Phantom constructed 75
4.3 Different signal-to-noise ratios 76
4.4 Different classes of projection images are specified by different projection angles. 77
4.5 Power spectrum and amplitude generated from NUDFT. 78
4.6 Eigenvertors of the images after spectral embedding. 80
4.7 Distance matrix of the data. 81
4.8 Left: error rate of two algorithms under each SNR. Right: accuracy of two algorithms under each SNR. 81

5.1 Phantom constructed 91
5.2 Different signal-to-noise ratios. 93
5.3 Power spectrum generated by the NUDFT. 95
5.4 Left: relative error rate of two algorithms. Right: correntropy calculated from two algorithms 95

6.1 Proposed algorithm pipeline. 107
6.2 Phantom constructed 107
6.3 Different signal-to-noise ratios 108
6.4 Samples of projection images. 109
6.5 Result of spectral clustering and Z-score method. 110
6.6 Part of the result of the similarity matrix. 111
6.7 Frequency information generated from the NUDFT. 112
6.8 Z-score application result. 113
6.9 Correntropy obtained from two algorithms. 116

List of Tables

3.1 Mean classification accuracy of each method generated from equation 3.20, in the bracket is the 95% confidence interval from 100 simulations. 59

4.1 Mean classification accuracy of each method generated from equation 4.31, in the bracket is the 95% confidence interval from 100 simulations. 82

5.1 Mean relative error rate of each method generated from equation 5.8, in the bracket is the 95% confidence interval from 100 simulations. 96

6.1 Mean accuracy of the classification before and after Z-score application generated from 6.9, in the bracket is the 95% confidence interval from 100 simulations. 114

6.2 Final correntropy of each algorithm 115

List of Algorithms

1 Non-uniform discrete Fourier transform based rotation-invariant feature extraction algorithm 49

2 Non-uniform discrete Fourier transform based spectral clustering featured 2D image classification algorithm 73

3 Non-uniform discrete Fourier transform based K-means clustering featured 2D image classification algorithm 74

4 Non-uniform discrete Fourier transform based power spectrum featured in-plane rotation angle estimation algorithm 88

5 Conventional 2D Fourier transform based power spectrum featured in-plane rotation angle estimation algorithm 89

6 Non-uniform discrete Fourier transform based spectral clustering featured 2D in-plane rotation angle estimation algorithm 105

7 Non-uniform discrete Fourier transform based 2D in-plane rotation angle estimation algorithm 106

Acknowledgements

ZiJian Bai's PhD study was funded by the China Scholarship Council (File NO.202108300031).

Abstract

Background and Objective Single-particle Cryo-EM is a powerful tool for structural biologists to determine structures of macro-molecules at atomic resolution. In single-particle Cryo-EM, many individual biological molecular particles with identical structures are rapidly frozen at unknown, random orientations, and then imaged by transmission electron microscopy to generate a set of two-dimensional (2D) projection images. These generated images are treated as projections of one particle from various directions and used to reconstruct the 3D structure of the particle.

Reconstruction in single-particle Cryo-EM is extremely challenging due to the lack of knowledge on the projection directions and low signal to noise ratios (SNR). To tackle this difficulty, most methods start by estimating an ab-initio model from class-averaged particle images. Then, this initial model is refined iteratively until a high resolution map is obtained, a task named '3D refinement'. There are two challenges in reconstructing a 3D particle structure from projection images in single-particle Cryo-EM. One is that as the particles are oriented randomly within the ice, the projection angle of the image is unknown. Another is the very low SNR, it is extremely hard to estimate projection angles. The main reason for the low SNR is the low electron doses allowed to be used for imaging. To improve the SNR in single-particle Cryo-EM, projection images are classified and averaged to generate the class-average images. This thesis focus on identifying and developing methods that can accurately and effectively compute class-averages.

Methodology Two important steps in class averaging are: (1) classifying images with similar viewing directions but with different in-plane rotations; (2) rotation alignment of images in each class.

To perform rotation-invariant classification and rotation alignment, we first propose a non-uniform discrete Fourier transform (NUDFT) to calculate the Fourier transform of an image in polar coordinates. Based on the proposed NUDFT, we develop a rotation-invariant feature extraction algorithm. After using principal component analysis to reduce the dimension of the extracted rotation-invariant features and defining a distance between images using their corresponding features, K-means is employed to classify images. We also investigate combining spectral clustering with our NUDFT based rotation-invariant features for the image classification and compare its performance with the K-means.

We build an algorithm for estimation of rotation between a pair of images on the base of proposed NUDFT. We also develop a novel spectral clustering method to improve the accuracy of image alignment in the same class. To demonstrate the efficiency of the algorithms, extensive simulations and experiments are performed to compare with rotation-invariant classification and rotation alignment based conventional FT and some existing algorithms in single-particle Cryo-EM.

Results and conclusions The results for the first set of simulation studies show that combining NUDFT with K-means has superior performance in rotation-invariant classification compared with K-means using features extracted by the classical FT and CL2D, displaying enhanced noise resistance, particularly in very low signal-to-noise ratio environments. The second set of simulation studies show that combining NUDFT with spectral clustering can further improve classification performance. In third set of simulation studies, the results show that NUDFT has superior performance in image rotation estimation compared with classical FT, it shows enhanced noise resistance in different SNR levels. The fourth set of simulation studies show that applying spectral clustering and Z-score method on the frequency information generated from NUDFT can further enhanced alignment performance. The overall studies show that NUDFT is effective and efficient in 2D class-averaging, and proposed algorithms can realize the potential of NUDFT and improve its performance.

Chapter 1

Introduction

1.1 Motivation

Structural biology is a field of science that uses a variety of techniques to determine the 3D structures of macro-molecules such as proteins and protein complexes. Knowledge of the structures of is of major importance for improving our understanding of fundamental life processes, since cells, considered the basic units of life, are made of many complex molecules called macro-molecules, such as proteins, and nucleic acids (RNA and DNA). Structural biology has made important contributions to the study of pathological mechanisms, drug development, and the promotion of biotechnology [103]. Nowadays, single-particle Cryo-EM is often the first choice of structural biologists to determine structures of macro-molecules at atomic resolution[71].

In single-particle Cryo-EM many individual biological molecular particles with identical structures are rapidly frozen at unknown, random orientations to below -150°C [2, 98]. Subsequently, a beam of electrons passes through frozen-hydrated particles, generating a set of two-dimensional (2D) projections. These generated images are treated as projections of one particle from various directions and used to reconstruct the 3D shape of the particle.

Reconstruction of the 3D object from its multiple projection images is the reverse process of obtaining 2D projection images from 3D objects. The problem is well researched in medical imaging such as computed tomography (CT)

and positron emission tomography (PET). Various algorithms have been developed for the reconstruction of a 3D object from its 2D projection image when projection angles are known[45]. However, in single-particle Cryo-EM, as the particles are oriented randomly within the ice, the projection angles of images are unknown, and one faces the problem of reconstruction of a 3D object from random unknown projections. One solution is that estimate projection angles from data first, and then reconstruct the 3D particle from projection images with the estimated projection angles using reconstruction methods well developed in medical imaging such as CT and PET.

Because of the low SNR, it is extremely hard to estimate projection angles. The main reason for the low SNR is the low electron doses allowed to irradiate the macro-molecules of interest, otherwise severe radiation damage accumulates, deforming the macro-molecules and affecting the projections. To improve the SNR, tens of thousands of images are collected in single-particle Cryo-EM, classified with similar projection directions and averaged to generate the class-average images. Since the randomly oriented and positioned particles are imaged, images from the same projection direction can differ by rotation around the projection direction (in-plan rotation) and translation. Thus, the classification should be rotation and translation invariant, the images in the same class should be translationally and rotationally aligned to compute the class-averaging image. In this thesis, we explore existing rotation and translation invariant classification and alignment techniques with the aim of developing strategies to deal with the problem of alignment and clustering of single-particle EM images.

1.2 Main Objectives

Construction of the 3D shape of the particle from large sets of single-particle Cryo-EM projections is challenging due to the low SNR. Classification, alignment and averaging of the 2D projection images, the process called “class averaging”, is commonly used to improve SNR. This work aims to identify and develop statistical methods that can accurately and effectively compute class-averages. Its main objectives are to compare and test rotation and translation invariant classification and alignment techniques using intensive simulation studies, thereby

developing a novel approach for computing class-averages. Our overarching goal is to improve the reconstruction of the 3D particle from 2D Cryo-EM images.

1.3 Outline and Contribution

The main contributions of this thesis is list as follow:

I. Innovatively apply non-uniform discrete Fourier transform (NUDFT) to generate Fourier transform of the image (in Cartesian coordinate system) in polar frequency domain, this can convert the rotational transformation between images into translational transformation without interpolation.

II. Develop an algorithm to generate rotation-invariant features based on NUDFT.

III. Develop classification algorithm based on NUDFT, combining spectral clustering method while defined a novel distance metric of images using both amplitude and power spectrum.

IV. Develop algorithm based on NUDFT to compute in-plane rotation angle by determining translational transformation between images.

V. Develop in-plane rotation angle estimation algorithm based on NUDFT, combining Z-score method and spectral clustering while defined a novel distance metric using both amplitude and power spectrum, this algorithm can delete images with minimal changes to improve the resolution, by using an unsupervised clustering method to improved the performance of quantitative analysis problem.

The thesis structure is outlined as follow:

Chapter 2 introduce single-particle cryo-electron microscopy (Cryo-EM) and its recent development history, especially in 2D image problem. Different methods including traditional method using 2D Fourier transform in 2D projection and 2D alignment was reviewed, also, the limitations was discussed. The chapter ends by identifying the application of non-uniform discrete Fourier transform and spectral clustering in single-particle Cryo-EM and their potential in improving the performance.

Chapter 3 focus on NUDFT in 2D classification, we develop algorithm to generate rotation-invariant features based on NUDFT, followed by applying PCA and K-means clustering. A real-world data was used to examine the proposed

algorithm, and constructed comparison with traditional FT and CL2D used in the Xmipp software.

Chapter 4 further develop the NUDFT in 2D classification. Based on the rotation-invariant feature extraction algorithm in Chapter 3, we combined spectral clustering to realize the potential of NUDFT. Also, we defined a novel distance metric using both amplitude and power spectrum generated from NUDFT. Comparison was constructed based on synthetic data to explore the performance of proposed algorithm under different signal-to-noise ratios.

Chapter 5 considered NUDFT in 2D alignment. We developed algorithm to use the power spectrum generated from the NUDFT to determine the in-plane rotation angle between images. Comparison between traditional 2D FFT using interpolation was constructed to examine the improved performance of the NUDFT.

Chapter 6 further fulfil the application of NUDFT in 2D alignment. We combined spectral clustering and Z-score method with NUDFT to use an unsupervised clustering method to enhance the performance of image rotation estimation. With our defined novel distance metric using both amplitude and power spectrum generated from NUDFT, plus the application of Z-score method, the algorithm can much improved the resolution of the algorithm. And we conducted comparison based on synthetic data to examine the performance under different SNR.

Chapter 2

Background and Literature review

2.1 Structural biology

In this section, we give a brief introduction to structural biology. For a formal introduction see [4].

Structural biology is a field of science that uses a variety of techniques to determine the 3D structures of biomolecules such as proteins and protein complexes. It has made important contributions to the study of pathological mechanisms, drug development, and the promotion of biotechnology[103]. Discovering, observing, and understanding the structure of bio-molecules can give us an in-depth understanding of the fundamentals of life. The initial establishment of the concept of structural biology may not be traced back to a precise point in time, because it gradually evolved into an independent field of discipline with the development of science and technology. However, looking back at its development history, we can also find that people's in-depth exploration of biological molecular structures and the development and evolution of science and technology complement each other. In fact, people's exploration of the molecular structure of organisms began as early as the beginning of the 20th century. In 1940, X-ray crystallography[97] enabled people to deduce the structure of some small molecules, but it was still very difficult to analyze the structure of biological macromolecules such as pro-

teins. In the 1950s-60s Protein crystallography[58] allowed the use of X-rays to resolve the structures of biological macromolecules. At the same time, Watson and Crick proposed the double helix structure of DNA, marking the beginning of the revolution in molecular biology. The development of nuclear magnetic resonance (NMR) technology[61] (1960-1970) first began to introduce liquid solutions as carriers in sample preparation. Although there were still some limitations to analyze complex structures, it provided new possibilities for future development. The development of computer simulation and modelling technology (1980-1990) and the breakthrough of Cryo-EM technology (2000-)[65], ushered in a complete revolution in structural biology, and now we can analyze the high-resolution structure of biological macromolecules, to reveal the structures and mechanisms of biomolecules such as viruses and membrane proteins, an example of the structure of the protein is shown in figure 2.1. There are various imaging technologies used in today's biological imaging field. People believe the powerful resolution capabilities and unique imaging principles of Cryo-EM are poised to lead the next breakthrough and innovation in structural biology[12].

2.2 Cryogenic Electron Microscopy (Cryo-EM)

In this section, we briefly introduce Cryo-EM, including the history of its development.

The technique of imaging the three-dimensional structure of biological macromolecules under low-temperature conditions using transmission electron microscopy is known as Cryogenic Electron Microscopy, abbreviated as Cryo-EM. In recent years, Cryo-EM technology has made remarkable advances and is increasingly applied in the structural studies of biological macromolecules. It plays an irreplaceable role in studying protein complexes that are difficult to handle with other imaging techniques, and continues to rapidly evolve and iterate[60]. The rapid innovation of Cryo-EM in recent years is mainly attributed to breakthroughs in three aspects of technology. Firstly, sample preparation has been improved by enveloping biological macromolecule samples in thinner ice layers using thin film carbon or even graphene[68], leading to an improved signal-to-noise ratio. Secondly, innovations in electron detection technology, such as the invention of

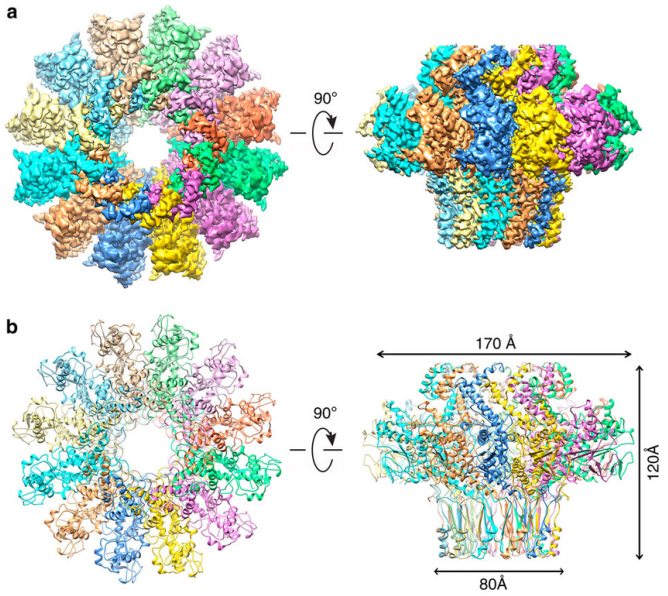


Figure 2.1: An example of a complex 3D structure of protein
(a) 3D density map of T4 portal protein assembly at 3.6Å resolution with each subunit colour-coded. Shown are the top view (left) and side view (right).
(b) Ribbon diagram of the gp20 atomic model with each subunit colour-coded. Shown are the top view (left) and side view (right). [84]

electron detectors, have significantly enhanced the resolution of cryo-electron microscopy by directly detecting the number of electrons instead of the complex process of photon generation and conversion through CCD cameras[8]. Thirdly, advancements in computational power and software algorithms have facilitated the reconstruction of Cryo-electron microscope models, which often require the analysis of a large number of projection images. In 2017, this technology was awarded the Nobel Prize in Chemistry for "developing Cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution." Its development is the result of the combined efforts of various technologies, as evidenced by its developmental history, a brief overview of the history is shown in figure 2.2. In 1924, the 'matter wave' hypothesis proposed by De Broglie[25] laid the foundation for electron microscopy. With the development of imaging technology, electron microscopes represented by the Scanning Electron Microscope (SEM)[62] and Transmission Electron Microscope (TEM)[95] have reached the level of imaging microstructure and microcomposition. However, they are unable

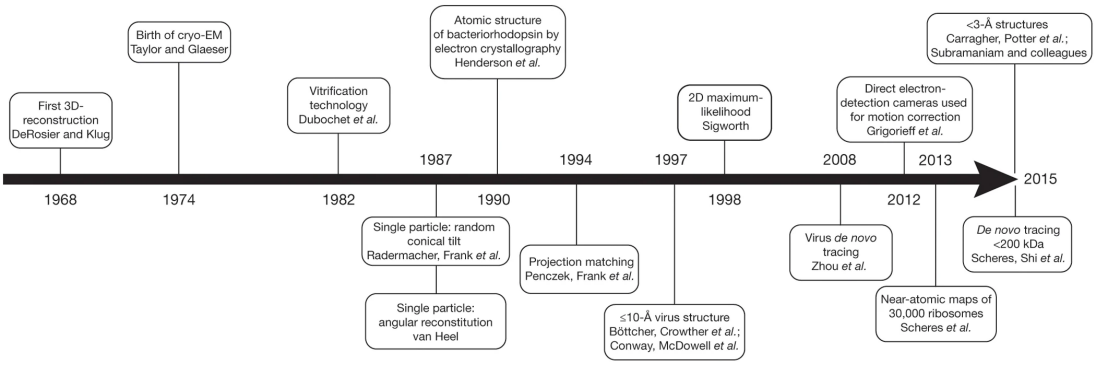


Figure 2.2: Timeline of key contributions to the field of 3D electron microscopy highlighted in this Commentary.[65]

to handle biological samples due to three major limiting factors. Firstly, the high-energy electron beam directly damages active biological samples, causing their structure to break and lose activity under radiation. Secondly, imaging needs to be conducted in a high vacuum environment due to technological limitations at the time. However, active biological samples usually contain moisture, which evaporates in this environment, leading to sample structure breakage and affecting electron beam penetration[87]. Thirdly, biological samples typically contain light elements such as C, H, N, and O, which are not conducive to imaging and seriously affect the final imaging performance. Thus, in the attempt to use electron beams for imaging biological macromolecules, people began decades of exploration, development, and innovation.

In the 1950s, negative staining techniques[66] were employed by researchers to fix biological structures and observe them under electron microscopy. This experimental method, while straightforward to operate and providing high-contrast images, yielded electron micrographs with resolutions only reaching around a few tens of angstroms.

In 1974, Kenneth A. Taylor and Robert M. Glaeser achieved a significant breakthrough by successfully obtaining high-resolution image information from frozen-hydrated catalase at -120°C[85]. During this process, they discovered that freezing the samples helped reduce radiation damage, thereby marking the beginning of Cryo-EM application in structural biology.

In 1980, Jacques Dubochet made a groundbreaking advancement in thin film

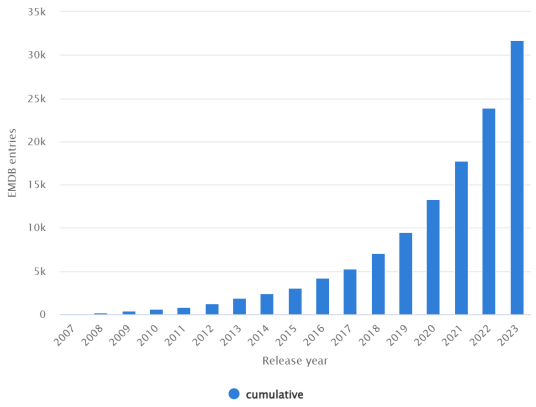


Figure 2.3: Cumulative number of particle entries released in the last 17 years (statistical data comes from EMDataBank).
Cryo-electron microscopy technology and single-particle reconstruction technology are attracting more and more attention.

of water vitrification[30]. The liquid ethane bath freezing method he introduced was fast and efficient, and it continues to be utilized in Cryo-EM sample preparation to this day.

In terms of reconstruction methods, in 1968, Aaron Klug[27] proposed the basic mathematical principle for reconstructing three-dimensional objects from two-dimensional images, also known as the Central Section Theorem (or Projection-Slice Theorem)[14], which laid the foundation for three-dimensional reconstruction technology. Klug was awarded the Nobel Prize in Chemistry in 1982 for this contribution.

In 1987, Max Radermacher, Joachim Frank and others proposed tomography imaging and single particle analysis methods[70]. Single-particle analysis technology can be used to image large protein complexes that resist crystallization. It's the basis of macromolecular structure analysis methods widely used in the field of cryo-electron microscopy today.

In 2008, Professor Zhou Zhenghong of the University of California, successfully obtained the three-dimensional structure of the CPV virus with a resolution of 3.9Å through more than 12,000 single-particle images[104]. This is the first time that the structure has been obtained with near-atomic resolution through single-

particle Cryo-EM reconstruction technology.

Since 2013, Cryo-EM has made revolutionary progress due to the development of direct electron detector device (DED). Professor Cheng Yifan's research group at the University of California, successfully used a new generation DED camera to analyze the structure of the transient receptor potential channel protein (TRPV1) at 3.4Å resolution [52]. This study breaks the resolution barrier of non-crystallizing membrane protein side chains and demonstrates the great potential of single-particle Cryo-EM for membrane protein analysis. A demonstration of different approaches is shown in figure 2.4.

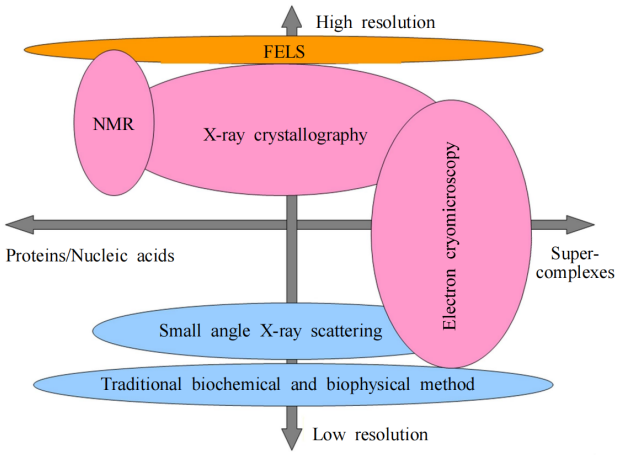


Figure 2.4: Comparison among different technologies in structural biology. The vertical axis refers to resolution (higher from bottom to top) and the horizontal axis refers to the investigated biological sample scale (larger from left to right). Small angle X-ray scattering is suitable to study biological molecules in multiple scales, but the resolution is relatively low; Nuclear Magnetic Resonance (NMR) can be used to obtain structures with high resolution, but limited to relatively small molecules, eg. below 30 KD. X-ray Crystallography has a wider investigation scale and can be used to gain high resolution, mostly atomic resolution, but crystallization is its bottleneck. Cryo-electron microscopy is powerful to solve super-complexes in medium resolution, usually on a nanometer scale, but it can reach near-atomic resolution for symmetric structures nowadays. Free electron laser scattering (FELS) is a new technology under development, which is promising to solve the structure of biological samples in multiple scales and high resolution.

Currently, Cryo-EM technology has become more mature, with a large number of biological molecular structures being elucidated (shown in figure 2.3). People's enthusiasm for research in this area is also increasing, and it is believed that this technology will continue to see significant advancements in the future. Although some may joke that "Cryo-electron microscopy is a chemistry award given to physicists for their outstanding contributions to the biological field," this statement precisely illustrates that Cryo-EM is an interdisciplinary fusion technology that combines biology, physics, chemistry, mathematics, and computer science. It is the marvellous collision of interdisciplinary inspiration, emerging from the integration of various independent disciplines, leading to the final breakthroughs in technology. Therefore, it is a pioneering technology accompanying the development of human scientific and technological progress[24].

2.2.1 Imaging Process of Cryo-EM

A typical Cryo-EM workflow is shown in figure 2.5.

Step 1: Sample Preparation. A high-purity, high-concentration solution of biological sample is pipetted onto a specially designed sample grid. The grid consists of an ultrathin amorphous carbon film punctuated with small holes and supported by a metal frame. Under the influence of surface tension, a thin water film forms across the holes. After removing the excess solution, the grid containing the thin film of the biological solution is rapidly plunged into liquid ethane cryogen to rapidly freeze it, thereby immobilizing the biomolecules dispersed within the vitrified ice[29].

Step 2: Electron Microscopy Image Acquisition. Samples with the optimal particle density and thickness of vitrified ice are more likely to produce the best images and are selected. Parameters such as defocus value, magnification, and electron dose are set optimally, and a large number of images of these sample regions are recorded[60]. These images, depicting projections formed by discrete molecules, are manually or semi-automatically selected.

Step 3: 2D projection image class-averaging. Due to the potential for electron radiation damage to sensitive samples, cryo-electron microscopy employs very low electron doses, resulting in significant background noise in 2D projection

images. To improve image resolution, researchers first propose an initial 3D model and then perform particle picking on the captured 2D images of individual particles[19]. And perform class-averaging to the classified and aligned projection images to enhance SNR. The detailed process of this step will be elaborated in subsequent sections.

Step 4: Three-dimensional Structure Reconstruction. With the enhanced SNR class-averaged projection images, the 3D structure of the particle can be reconstructed after 3D classification and iterative process.

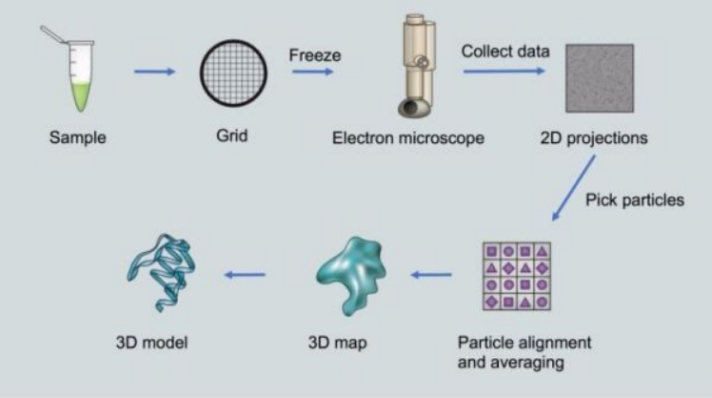


Figure 2.5: Typical Cryo-EM workflow.

Biological samples cannot be subjected to high-dose electron irradiation to achieve high imaging quality. This is because the radiation damage to biological samples is correlated with the cumulative radiation dose; as the radiation dose increases, the radiation damage to the details in the images of the sample becomes more severe[9]. Therefore, in order to obtain more details, low-dose radiation imaging of biological samples is necessary. However, reducing the electron dose to minimize damage results in noisy images with low SNR, while increasing the electron dose sufficiently to achieve a good signal-to-noise ratio leads to specimen damage reaching unacceptable levels, thus, resulting in images with high errors.

To address the aforementioned issues, cryo-electron microscopy (Cryo-EM) technology typically employs the following two methods. The first method involves imaging frozen specimens stored at liquid nitrogen or liquid helium temperatures (as previously detailed in the previous section). Imaging at liquid nitrogen temperatures can reduce radiation damage by up to sixfold compared to

room temperature[32]. The second method to improve the signal-to-noise ratio is to average images of a large number of identical biological specimen units[75]. This technique was first applied to the imaging of helical assemblies and two-dimensional protein crystals at both room and low temperatures. These two concepts, namely the concept of low-temperature imaging and the concept of averaging multiple low-dose images, constitute the foundation of modern high-resolution cryo-electron microscopy of biological.

In Cryo-EM technology, the commonly used low-dose radiation imaging methods are primarily classified into two categories based on the strategies employed during imaging: cryo-electron tomography (Cryo-ET) and single particle analysis (SPA). In Cryo-EM the sample is continuously rotated relative to the lens, and imaging is conducted at different 3D rotation angles[77]. Therefore, each imaging frame represents a two-dimensional projection of the sample in different known projection directions. The reconstruction problem is to reconstruct a 3D object from its 2D projections with known projection directions. This process is demonstrate in figure 2.6. This is well-researched in medical imaging such as computed tomography (CT) and positron emission tomography (PET). In the following section, we introduce single-particle Cryo-EM, the focus of this thesis.

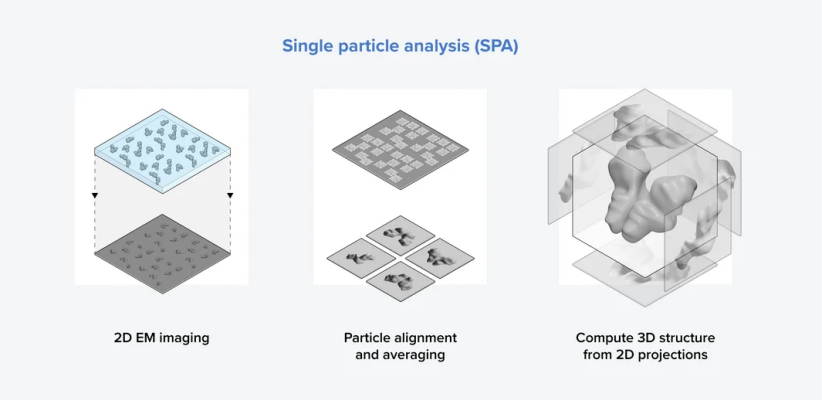


Figure 2.6: The three main steps in the single particle analysis workflow. First, an image of a sample containing identical purified proteins in various orientations is taken. Subsequently, particle alignment and averaging are conducted to generate 2D images of the protein from different orientations. Finally, the 3D structure of the protein is computed from the 2D images.

2.3 Single-particle Cryo-EM

Single-particle Cryo-EM (also known as Cryo-EM single particle analysis, SPA) is an increasingly popular Cryo-EM technique for determining the structures of large biological macromolecules (particles). This method emphasizes the analysis and reconstruction of individual particles, making it particularly suitable for the analysis of complex biological macromolecule structures such as proteins or viruses[11]. The following sections will provide an introduction of it, including its development timeline.

2.3.1 Development Timeline of Single-Particle Cryo-EM

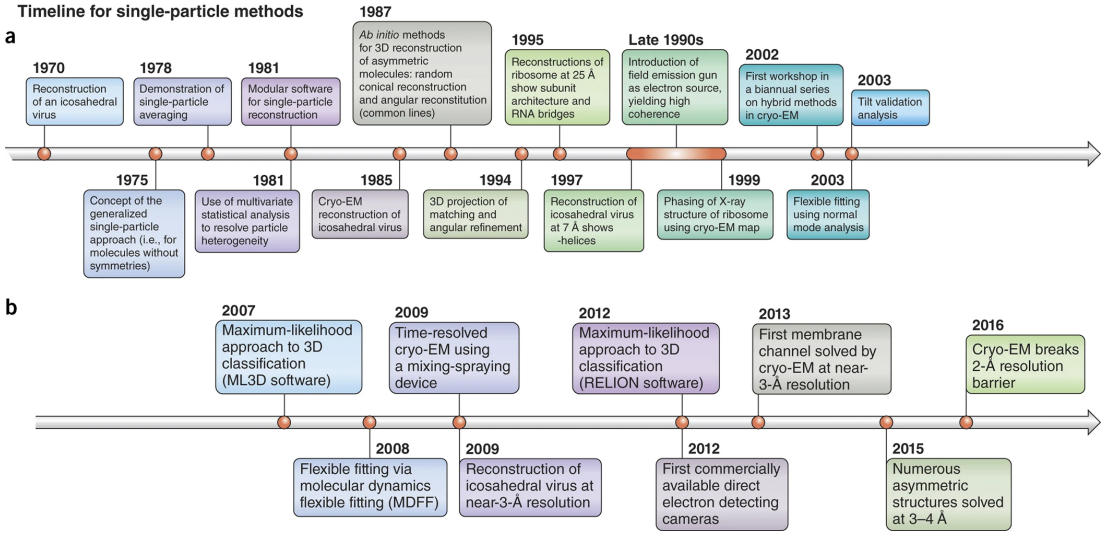


Figure 2.7: Timeline of key events in the development of single-particle Cryo-EM methods.[37]

In the 1950s, electron microscopy began to be used to observe biological structures fixed using negative staining techniques[33], marking the realization of Cryo-EM imaging technology, as detailed in the previous sections. A shortcut of the development timeline is shown in figure 2.7.

A significant milestone for single-particle Cryo-EM occurred in 1987 when Max Radermacher and Joachim Frank proposed single-particle analysis[70]. This marked the initial application of single-particle Cryo-EM, demonstrating its effec-

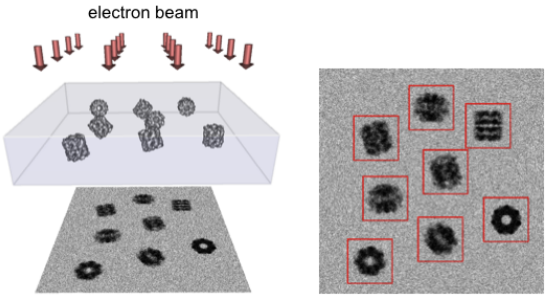


Figure 2.8: A sample of single-particle Cryo-EM imaging process.

tiveness as an analytical tool in the field of biomolecules and laying the theoretical foundation for future developments.

In the 2000s, advancements in computer technology, image processing techniques, and more advanced algorithms significantly improved the efficiency of data collection, image processing speed, and reconstruction resolution of single-particle Cryo-EM, laying the groundwork for further development of the technique.

In 2008, Professor Zhenhong Zhou of the University of California, Los Angeles, achieved near-atomic resolution structures through single-particle cryo-electron microscopy reconstruction technology[104]. This achievement marked a new level of resolution in single-particle analysis with Cryo-EM, ushering in a new stage of exploration into biomolecules.

Overall, single-particle Cryo-EM technology has undergone decades of development, from initial exploration to high-resolution structural analysis today, achieving significant progress and achievements along the way.

2.3.2 Imaging Principle of Single-Particle Cryo-EM

As shown in figure 2.8, the imaging process of single-particle Cryo-EM involves preparing many individual biological molecular particles with identical structures into frozen-hydrated samples. After applying the electron beam to the samples, many 2D projections of identical particles are collected for reconstructing the 3D structure of the particle. Its main difference from Cryo-ET lies in the way 2D projection images are obtained[20].

In Cryo-ET, the rotation of the sample during imaging is controlled, thus the projection angles of the 2D projection images are known. In contrast, in single-particle Cryo-EM, the particles being imaged frozen into random orientations in the thin film, the particle orientation relative to the projection direction is unknown. The imaging process generates a large number of projection images of the particles simultaneously with unknown orientations Theses EM images can be treated as if they are projections of a single particle in different orientations, from them we wish to estimate the 3-D structure.

Reconstructing a 3D structure from 2D projection images is a basic and important task in medical imaging, such as CT scans and Positron Emission Tomography (PET).

In CT, X-Ray projection images are acquired using a series of x-rays, which are a form of radiation on the electromagnetic spectrum. The CT scanner emits x-rays towards the patient from a variety of angles – and the detectors in the scanner measure the difference between the x-rays that are absorbed by the body, and x-rays that are transmitted through the body. The 3D CT images of X-ray attenuation of the patient are reconstructed from the X-Ray projection images.

In PET, projection images acquired by injecting radiotracer isotope into the patient. Positrons are released when a radioactive tracer isotope decays. Positrons annihilate with electrons within the body, creating a pair of photons that travel in opposite directions. Electronically coupled opposing detectors simultaneously identify the pair of photons by using coincidence detection circuits. The annihilation reaction is thus known to occur along the line joining the two detectors. PET projection data consist of a number of these coincidence lines. The PET image of the tracer distribution within a patient is reconstructed from these projection images.

Mathematically, the reconstruction problem can be formulated as recovery of a function from its line integrations (Radon transformation). A well-known analytic image reconstruction method is filtered backprojection (FBP). FBP first filters the projection data and then backprojects the filtered data to obtain an image [22]. More advanced methods based on improved model of the noise in the data, such as the expectation maximisation (EM) algorithm for maximum likelihood (ML) reconstruction and and its accelerated version, ordered subsets

EM (OSEM), have been developed and successfully used for reconstruction of CT and PET images [79, 51, 46]. However, these methods assume that projection directions are known, thus cannot directly applied to the image reconstruction in single particle Cryo-EM. In single-particle Cryo-EM, the particles being imaged frozen into random unknown orientations in the thin film, the particle orientation relative to the projection direction is unknown. The problem at hand is to reconstruct a 3D structure from its projections without knowing the projection angle values. To use above mentioned methods, one needs to estimate projection angles of the projection images.

The main difficulty in estimation of projection angles of the projection images is that very low signal-to-noise ratio (SNR) in the images due to the low electron dose being used to avoid radiation damage. A key step in the data processing is to enhance the image SNR by averaging hundreds of thousands of projection images with the same projection angle. Randomly oriented particles can have different rotation angles relative to the projection direction. Thus, there are images from the same or a similar projection direction but with a different in-plane rotation. The images from the same or similar projection directions are clustered together, aligned and averaged to increase the SNR, as images corresponding to similar projection directions tend to be very much alike (due to the smoothness of the molecular structure). This process is called 2-D classification and alignment.

2.4 2D Classification

In single particle Cryo-EM, image classification is to group projection images into classes based on projection angles. The objective of classification is to construct class averages that have higher SNR than the raw images. This process is essential for improving image SNR[72].

The purpose of the classification of 2D projection images is to sort and categorize the images, with the ultimate goal of dividing a large number of images into different categories based on their projection angles. This process is essential for improving image SNR[82]. The primary challenges encountered in this endeavour stem from the inherent in-plane rotations [81] and low SNR[10] of the projection images. These factors significantly confound the classification outcomes by in-

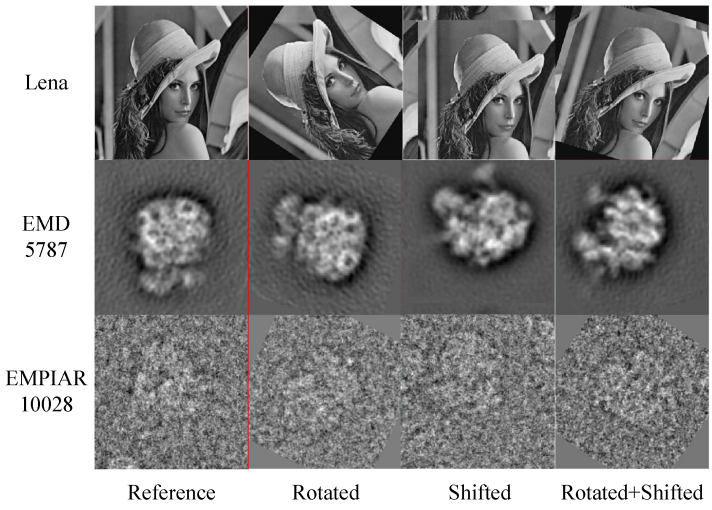


Figure 2.9: Image translation in single-particle Cryo-EM.[17]

troducing substantial ambiguity regarding the distinct projection angles of the images (shown in figure 2.9).

In single-particle Cryo-EM, it is crucial that image classification is rotation-invariant, since two images from a similar viewing direction but with a different in-plane rotation should be classified to the same class. Rotation-invariant image classification can be fulfilled by extracting rotation-invariant features from images, and then defining a distance between images based on their corresponding extracted features using conventional metrics such as a Euclidean distance.

The classification process based on the rotation-invariant features primarily involves two main approaches: the reference-free classification (RFC) method and the multi-reference alignment (MRA) method. The MRA method typically requires the random selection or establishment of one or more templates based on certain prior knowledge. While this approach may be influenced by template selection, it offers fast computation and is easier to comprehend, making it the most commonly used method. In contrast, the RFC method involves the statistical analysis of a large number of images to directly generate distinct classification results. Commonly used MSA techniques include correspondence analysis[42] and principal component analysis[1, 96]. After applying MSA techniques, similarity or Euclidean distance calculations can be performed in a relatively smaller dimensional space, used by various clustering algorithms to classify the data[90, 89, 13].

Furthermore, new methods like Vector Diffusion Maps (VDM)[21] have also been shown to be applicable to 2D projection images in Cryo-EM. Additionally, the auto-correlation function and double auto-correlation function of images [74] can also serve as highly effective rotation-invariant features. These calculated features can be directly used to classify the original data, enhancing computational efficiency and reducing noise interference to some extent.

Over the past several decades, the extraction of rotation invariant features has been extensively studied. Various methods have been developed for the extraction of rotation invariant features, including methods based on Fourier transform [17, 5, 11], methods based on combination of Radon and Fourier transforms [99], methods based on wavelet transform[63, 69] and methods based on combination of Radon and Wavelet transforms[47].

Wavelets (i.e. small waves) are mathematical functions that represent scaled and translated (shifted) copies of a finite-length waveform called the mother wavelet. A wavelet transform (WT) is based on wavelets. It is used to analyze a signal (image) into different frequency components at different resolution scales (i.e. multiresolution) [3]. Wavelet transforms share functional similarities with Fourier transforms, but unlike the Fourier transform, wavelet transforms may describe localized signals more efficiently. Their ability to extract rotation-invariant features makes them a promising alternative to Fourier transform in classifying projection images in single particle Cryo-EM. To knowledge of the author, wavelet transforms have not been used in classification of single particle Cryo-EM images. Their application in single particle Cryo-EM would worth investigating for future research. In this thesis we focus on application of Fourier transform in single particle Cryo-EM. Fourier transform and non-uniform Fourier transform will be briefly introduced in the following sections.

After extracting rotation invariant features, there are 2 types of clustering that can be performed to group projection images based on their extracted features. Hard clustering (e.g. k-means) assigns each image to a single cluster, enforcing strict categorical boundaries. In contrast, soft clustering (also known as fussy clustering) allows partial membership, assigning probabilities that indicate the likelihood of that image belonging to each of clusters. Based on an iterative algorithm, Gaussian mixture model assigns probabilities that indicate

the likelihood of that data point belonging to each of the Gaussian components (clusters) [59].

For single particle Cryo-EM, projection angles are unknown and sometimes may be slightly different from each other. Such small difference is difficult to detect in the presence of high noise. Soft clustering can be used to describe clustering uncertainty associated with high noise [83]. In this thesis we consider hard clustering due to its simplicity and computational efficiency.

It's worth mentioning that these methods are equally applicable to the 2D alignment. This is because the mathematical basis required to counteract the effects of in-plane rotation in images is also suitable for estimating the in plane rotation angle itself. We describe these two problems as two separate steps and discuss them individually, but their combination builds up the full process of 2D image class-averaging.

2.5 2D Alignment

After computing rotation-invariant features and classifying together images corresponding to the same (or quite similar) projection directions, there still exists an in-plane rotation between the images within the same class. Therefore, to calculate class-averages, images in the same class need to be aligned.

The solution to the problem shares a similar mathematical basis with computing rotation-invariant features but focuses on the estimation of the rotation angle between two images. Since rotation in Cartesian coordinates becomes translation along the angular coordinate axis of coordinates, many commonly used estimation methods utilize re-sampling the original images to polar coordinates (polar transform) and performing Fourier transform in polar coordinates. For instance, the FRM2D method proposed by Cong[23], employed in EMAN[55], this method avoids the need for zero-padding to achieve stable Fourier convolution in a one-dimensional linear space. An example is shown in figure 2.11.

Similar to the computation of rotation-invariant features, there also exist many methods that estimate rotation angles not depending on Fourier transformation. For example, methods in [48] estimate rotation angles between two images using cross correlation between two images. Methods employed in XMIPP[76]

and SPIDER[43] software directly align images by maximizing cross-correlation between images.

Most current methods involve finding a template within a class and aligning other images to it, making the final result highly susceptible to the selection of templates. Employing statistical methods can significantly mitigate this problem. With more accurate 2D alignment and well-classified images, the performance of 2D class averaging can be greatly enhanced.

2.6 Applications of Fourier transform

The significance of the Fourier transform in mathematics is indisputable, having profoundly influenced the development of mathematical theory and provided crucial mathematical tools and methods across various fields.

The Fourier transform is an important tool in image processing, it enables us to decompose the image into corresponding sets of sine and cosine waves[15]. The output of the transformation is a representation of image in the frequency domain, while the input image is the spatial domain equivalent. In the Fourier domain, each point represents a particular frequency contained in the spatial domain image. we present here the definition of Fourier transform of an image.

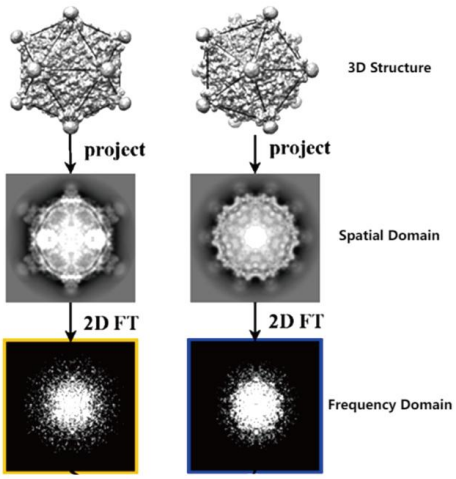


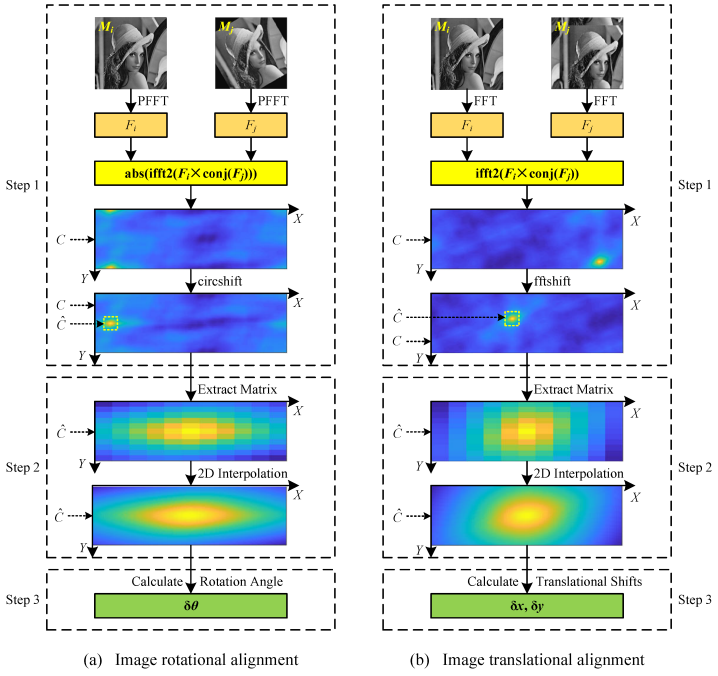
Figure 2.10: A sample of the 2D Fourier transform's application on single-particle Cryo-EM projection images.

As we are only processing digital images, we restrict the following discussion to the Discrete Fourier Transform (DFT).

While the Fourier transforms started to be applied to image processing areas in the early 1920s, the early Cryo-EM analysis is still based on the direct observation and analysis of particles' images. In the 1990s, with the development of both computer technology and the Cryo-EM technology, scientists began to realize the potential value of the Fourier transform in image processing and analysis and are beginning to apply it to the processing and reconstruction of single-particle Cryo-EM data, then subsequently developed a number of Cryo-EM data processing technologies[86][92]. An example of FT application is shown in figure 2.10.

And now, applying the Fourier transform in single-particle Cryo-EM are mainly focused on image processing, image classification, 3D reconstruction, structure analysis, etc.

As above discussed the 2D image classification and alignment are two of main tasks in single-particle Cryo-EM.



(a) Image rotational alignment (b) Image translational alignment

Figure 2.11: An example of image rotation alignment.[17]

Methods like Chen's[17] were developed based on the searching for the peak cross-correlation and rotation-invariant property of the Fourier transform, this enables us to determine the paired in-plane rotation angle between two input projection images. There's a series of variations of this method all based on the same idea[86].

2.7 Applications of Non-uniform discrete Fourier transform

Non-uniform discrete Fourier transform is a type of discrete Fourier transform. Different from conventional Fourier transform, its input signal or/and output signal is not sampled equally spaced, and NUDFT allows one to obtain frequency domain information of a finite length signal at any frequency[35].

The formula for non-uniform discrete Fourier transform (NUDFT) can be expressed as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi p_n f_k}, \quad 0 \leq k \leq N-1, \tag{2.1}$$

where $p_0, \ldots, p_{N-1} \in [0, 1]$ *are sample points,*
and $f_0, \ldots, f_{N-1} \in [0, N]$ *are frequencies.*
X_k *is the non-uniform discrete Fourier transform of* x_n*.*

There are three types of NUDFT. Type 1 NUDFT transforms uniformly sampled signals to the non-uniformly sampled frequency domain, while Type 2 NUDFT transforms non-uniformly sampled signals to the uniform frequency domain, and Type 3 NUDFT transforms non-uniformly sampled signals to the non-uniformly sampled frequency domain. Note that these types may not be universal, different authors will refer to different numbers of types but these three variants' ideas are the same.

Different types of NUDFT have different applications in the image processing field, as they can handle differently sampled input image and different needs of output in frequency domain. In MRI, type-II NUDFT is a widely used tool as the

data collected from the imaging process is often in polar coordinate format and showing nonuniformly distributed in Cartesian coordinates. All those different NUDFTs were designed to fit different distributions of input and output signals, and corresponding inverse transforms also have been well developed, but due to the computational cost, most people would not drop interest in them.

Using non-uniform discrete Fourier transform one can directly compute frequency of image in polar coordinates useful in rotation alignment. Without interpretation the NUDFT can generate more precise information compared to the conventional Fourier transform. With the appropriate design of the non-uniformly sampling frequency domain, this method have potential to be widely used in single-particle Cryo-EM. Interestingly, it is not widely used, only few publications can be found in the literature[39, 18, 101].

On the other hand, NUDFT is computationally costly, so the non-uniform fast Fourier transform (NUFFT) algorithm is developed to approximate it with little loss of quality [35]. Another limitation of NUDFT is described in [16]. Using NUDFT or NUFFT can obtain the Doppler spectrum of echo signal in condition of non-uniform data rate. However, the spectrum obtained by them in condition of non-uniform data rate will still appear obvious defocusing compared with the spectrum obtained by using FFT in condition of uniform data rate. [16] concluded that NUDFT is not suitable to directly obtain the signal's Doppler spectrum. However, Cryo-EM images are typically acquired as uniformly sampled signals. In this thesis, NUDFT is employed to transforms uniformly sampled signals to the non-uniformly sampled frequency domain.

2.8 Applications of Spectral Clustering

Spectral clustering is a clustering algorithm rooted in graph theory and linear algebra, it holds significant importance and has extensive applications in the field of image processing. Different from traditional clustering algorithms that typically rely on distance or density metrics to define similarity between input samples, spectral clustering treats input data as nodes in a graph. It constructs a similarity graph (weighted undirected graph) among samples and then performs spectral decomposition on this graph to compute the final clustering results.

From a mathematical perspective, spectral clustering can be viewed as a dimensionality reduction-based clustering method, which is achieved through the eigendecomposition of the affinity matrix[91]. This characteristic underscores one of spectral clustering's advantages: it requires only an affinity matrix as input to complete the whole clustering task. The flexibility afforded by the construction of the affinity matrix grants spectral clustering significant adaptability in classification tasks, as it allows for the use of any given definition to measure similarity (or distance) between input images. Consequently, it can optimize the definition of similarity based on the features of input data, yielding a more discriminate affinity matrix for better final clustering performance.

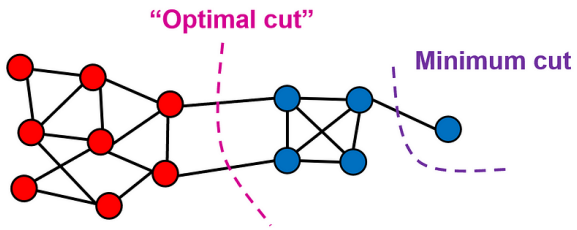


Figure 2.12: A sample of spectral clustering using the idea of graph theory.

The core clustering operation in spectral clustering applies after obtaining the eigenmatrix of the affinity matrix. At this stage, the eigenvectors are used to map the original input images into a low-dimensional space. The optimal segmentation (shown in figure 2.12) obtained after segmenting these low-dimensional nodes represents the result of the spectral clustering. These characteristics endow spectral clustering with several advantages and features which traditional methods do not have[64]. Moreover, as a non-iterative clustering method, spectral clustering does not require setting any thresholds or final cluster numbers to constrain iterative computations. Its final clustering result is achieved through spectral decomposition (i.e.the eigen decomposition), and its inherent dimensionality reduction property makes it have more noise resistance, which results in more robustness to noise.

In single-particle Cryo-EM, 2D image classification requires distinguishing and classification based on rotation-invariant features of input projection images as we described in the previous sections, thus similarity (or distance) based

rotation-invariant features should be used to construct the affinity matrix for spectral clustering. Several methods have been proposed in the context of Cryo-EM, proving its effectiveness in 2D classification[94, 93], but still, its current application numbers is far from its applicability and well performance in this field.

Chapter 3

NUDFT based 2D classification

3.1 Introduction

In this chapter we describe an NUDFT based method to classify 2D images in single-particle Cryo-EM and compare its performance with existing classical FFT based classification methods and other algorithms used in the current Cryo-EM processing software. Different from the classical FT, NUDFT can directly calculate the frequency representation of an image in polar coordinate from its spatial representation in spatial representation in Cartesian coordinates without intermediate step of converting image from Cartesian coordinates to polar coordinates. This unique property of NUDFT motivates us to explore its application in image classification in the single-particle Cryo-EM.

As discussed in the previous chapter, proper classification in single-particle Cryo-EM is required to be rotation-invariant and discriminative, a primary approach to addressing this requirement is extracting rotation-invariant features from images and performing classification using the extracted features[102]. Based on rotation-invariant features, various methods have been developed to classify images with the same projection angle but different in-plane rotations into the same class[104, 81]. However, the images are typically represented in a Cartesian coordinate system, a common approach is to use interpolation to convert frequency information in a Cartesian coordinate system to polar coordinates (for example, see [17]). The interpolation can introduce errors to the extracted

rotation-invariant features or estimate in-plane rotation angles. In this study, we explore an alternative method that does not require interpolation.

To eliminate the need for interpolation and generate the rotation-invariant information directly, we employ the NUDFT, a variant of the conventional Fourier transform. The NUDFT enables the generation of the Fourier transform at a desired and specified frequency point[35]. This capability gives us the flexibility to calculate power spectrum in polar frequency points, thereby converting the rotation in the Cartesian coordinates to 1D translation in the polar frequency coordinates, which is a fundamental step in developing our 2D classification algorithm. We expect this analysis will provide a fresh perspective on the 2D classification in Cryo-EM.

3.2 Methodology

3.2.1 Non-uniform discrete Fourier transform

The Fourier transform is a mathematical tool that transforms a signal from the time domain to the frequency domain and decomposes the signal into a sum of sine and cosine functions. It typically assumes that the signal is obtained under uniform sampling, i.e. with a fixed sampling interval. In practical applications, however, non-uniform sampling may occur, and this is where the non-uniform discrete Fourier transform (NUDFT) comes into play.

The NUDFT is a variant of the Fourier transform that can handle non-uniformly sampled data and generate non-uniformly sampled frequency information[34, 31].

NUDFT comes in three types: Type 1 NUDFT transforms uniformly sampled signals to the non-uniformly sampled frequency domain, while Type 2 NUDFT transforms non-uniformly sampled signals to the uniform frequency domain, and Type 3 NUDFT transforms non-uniformly sampled signals to the non-uniformly sampled frequency domain.

The 2D Type 1 NUDFT is a variant of the Type 1 NUDFT that is used to convert a two-dimensional signal or image from uniformly sampled data to a non-uniformly sampled frequency domain.

The 2D Type 1 NUDFT of an image is defined as:

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) \cdot e^{-i2\pi\left(\frac{ux}{N}+\frac{vy}{M}\right)} \tag{3.1}$$

Where $F(u,v)$ represents the image in the frequency domain (two-dimensional Fourier domain), where (u,v) is the frequency domain coordinate. $f(x,y)$ represents the the image in spatial domain, while (x,y) is the coordinates of the uniformly sampled points. N represents the number of sampling points in the x direction. M represents the number of sampling points in the y direction.

In summary, the key difference between the 2D discrete Fourier transform and 2D Type 1 NUDFT is that the latter is capable of generating non-uniformly sampled frequency information while the former requires equidistant distribution.

3.2.2 Rotation-invariant information

Rotation-invariant information of the image can be obtained from the frequency information generated from the Fourier transform. Fourier transform is a widely used mathematical tool in signal processing that allows us to decompose a signal into its frequency components. When applied to an image, the Fourier transform converts the image from the spatial domain to the frequency domain, providing us with information about the intensity and phase of different frequency components of the image. The power spectrum of Fourier transform is a term used to describe the distribution of signal power as a function of frequency, it represents the amount of energy present in each frequency component of a signal.

The power spectrum is particularly useful in extracting rotation-invariant information in image analysis[54]. As demonstrated in the following sections,the power spectrum of the Fourier transform in polar coordinates does not contain information about the orientation or rotation of the image. Thus, we can use the power spectrum to extract rotation-invariant features (information), which can be used for further image analysis and pattern recognition tasks.

Considering an image f as a two-dimensional discrete signal. And the image g is the in-plane rotated version of f with rotation angle θ_0. Under Cartesian coordinates we can have:

$$G(u', v') = F(u\cos(\theta_0) - v\sin(\theta_0), u\sin(\theta_0) + v\cos(\theta_0)) \tag{3.2}$$

This implies the Fourier transform of the in-plane rotated version $G(u, v)$ and the Fourier transform of the original $F(u, v)$ have the same rotation relationship between $f(x, y)$ and $g(x, y)$ with in-plane rotation angle θ_0.

According to the NUFT property, we can have u,v set in the format converted from ω and φ (which is in polar format), and the original Cartesian coordinate system in the original image f now have corresponding r,θ in the polar coordinate system. Then we have:

$$\begin{cases} x = r \cdot \cos(\theta) \\ y = r \cdot \sin(\theta) \end{cases} \quad \begin{cases} u = \omega \cdot \cos(\varphi) \\ v = \omega \cdot \sin(\varphi) \end{cases} \tag{3.3}$$

Now, the NUDFT formula can be written as:

$$F(\omega, \varphi) = \sum_{r=0}^{N-1} \sum_{\theta=0}^{2\pi} f(r, \theta) \cdot e^{-i \cdot 2\pi \cdot (\omega \cdot r \cdot \cos(\varphi) \cdot \cos(\theta) + \omega \cdot r \cdot \sin(\varphi) \cdot \sin(\theta))} \tag{3.4}$$

$$F(\omega, \varphi) = \sum_{r=0}^{N-1} \sum_{\theta=0}^{2\pi} f(r, \theta) \cdot e^{-i \cdot 2\pi \cdot \omega \cdot r \cdot \cos(\theta - \varphi)} \tag{3.5}$$

And x', y' in image g now convert into r', θ', also, u', v' have its corresponding ω', φ' in polar coordinate system as well. Thus we can have:

$$\begin{cases} \theta' = \theta + \theta_0 \\ \theta = \theta' - \theta_0 \end{cases} \tag{3.6}$$

And the NUDFT of g can be written as:

$$G(\omega, \varphi') = \mathcal{F}\{g(r', \theta')\} = \mathcal{F}\{f(r, \theta + \theta_0)\} \tag{3.7}$$

$$G(\omega, \varphi') = \sum_{r=0}^{N-1} \sum_{\theta=0}^{2\pi} f(r, \theta + \theta_0) \cdot e^{-i \cdot 2\pi \cdot \omega \cdot r \cdot \cos(\theta - \varphi)} \tag{3.8}$$

$$G(\omega,\varphi') = \sum_{r=0}^{N-1}\sum_{\theta'=0}^{2\pi} f(r,\theta')\cdot e^{-i\cdot 2\pi\cdot\omega\cdot r\cdot\cos(\theta'-\theta_0-\varphi)} \tag{3.9}$$

$$G(\omega,\varphi') = \sum_{r=0}^{N-1}\sum_{\theta'=0}^{2\pi} f(r,\theta')e^{-i\cdot 2\pi\cdot\omega\cdot r\cdot\cos[\theta'-(\varphi+\theta_0)]} \tag{3.10}$$

Then we can have:

$$G(\omega,\varphi') = F(\omega,\varphi+\theta_0) \tag{3.11}$$

Thus, by setting the frequency points in polar system format during NUDFT, the rotation relationship between images f and g now converts into a translation relationship among angular frequency. And the corresponding amplitude's (or power spectrum) translation on angular frequency under one single radial frequency is actually a periodic signal with different time shifts, so its amplitude (or power spectrum) should be the same.

Now consider applying a 1D Fourier transform on all angular frequency components at each radial frequency. Considering $F(s)$ and $G(s)$ as the Fourier transform of Amplitude obtained from prior NUDFT, then we can have:

$$\begin{aligned} F_j(s) &= \mathcal{F}_{1D}\{|F(\omega_j,\varphi_t)|\} \\ &= \sum_{t=0}^{N-1} |F(\omega_j,\varphi_t)|\cdot e^{-i\cdot 2\pi\cdot s\cdot t} \end{aligned} \tag{3.12}$$

$$G_j(s) = \mathcal{F}_{1D}\{|F(\omega_j,\varphi_t+\theta_0)|\} \tag{3.13}$$

Now considering the translation on angular frequency due to θ_0 is τ, then the equation (3.13) can be written as:

$$\begin{aligned} G_j(s) &= \sum_{t=0}^{N-1} |F(\omega_j,\varphi_t)|\cdot e^{-i\cdot 2\pi\cdot(s+\tau)\cdot t} \\ &= \sum_{t=0}^{N-1} |F(\omega_j,\varphi_t)|\cdot e^{-i\cdot 2\pi\cdot s\cdot t}\cdot e^{-i\cdot 2\pi\cdot\tau\cdot t} \end{aligned} \tag{3.14}$$

And the amplitude of these Fourier transforms can have the relationship as:

$$\begin{aligned} |G_j(s)| &= \left| \sum_{t=0}^{N-1} |F(\omega_j, \varphi_t)| \cdot e^{-i \cdot 2\pi \cdot s \cdot t} \cdot e^{-i \cdot 2\pi \cdot t \cdot t} \right| \\ &= | \sum_{t=0}^{N-1} |F(\omega_j, \varphi_t)| \cdot e^{-i \cdot 2\pi \cdot s \cdot t}| \cdot \left| e^{-i \cdot 2\pi \cdot t \cdot t} \right| \end{aligned} \tag{3.15}$$

As $|e^{-i \cdot 2\pi \cdot t \cdot t}| = 1$, then we can have:

$$\begin{aligned} |G_j(s)| &= \left| \sum_{i=1}^{M} \sum_{t=0}^{N-1} |F(\omega_j, \varphi_t)| \cdot e^{-i \cdot 2\pi \cdot s \cdot t} \right| \\ &= |F_j(s)| \end{aligned} \tag{3.16}$$

Thus we can recognize this amplitude (or further compute the power spectrum) as the rotation-invariant information of the images.

Based on (3.16) we proposed a rotation-invariant feature extraction algorithm shown in Algorithm 1. Prior to applying the Fourier transform, it is important to select a region of interest within the projection images. In general, the structure of molecules is concentrated in the centre of the image, while the four corners contain only noise information. Thus, we choose a circular region with the diameter set as the width of the square projection image.

After obtaining a circular mask of the original projection image, we apply the NUDFT. By considering the centre of the circle as the origin of the polar coordinate system, we can set all frequency coordinates generated in NUDFT in a polar format, which facilitates further computation. During the Fourier transform, we represent the original information under the frequency domain.

After applying the non-uniform discrete Fourier transform (our previous proposed design in Chapter 3) to the projection images, we obtain the frequency information for subsequent computation. However, the frequency information obtained still contains a rotation factor, albeit limited to column transformation, which means the rotation relationship between images now converts into the translation relationship among angular directions. Since we set the frequency points in a polar coordinate system during NUDFT, now we apply a 1D Fast

Fourier transform to the frequency information (which is generated from the NUDFT) along the angle direction to offset the rotation factor. Consequently, the power spectrum (and the amplitude) from this FFT yields rotation-invariant information, serving as a novel representation obtained from the original projection images.

Algorithm 1: Non-uniform discrete Fourier transform based rotation-invariant feature extraction algorithm

Input: Projection images I_i
Output: Rotation-invariant representation P_i' of the original input projection images

1 Select circle region R of the projection image with the diameter same as the image's width;
2 Apply non-uniform discrete Fourier transform on I_i to get F_i;
3 Apply 1D Fast Fourier transform along the angle direction of F_i to get P_i;
4 Compute the power spectrum by $P_i' = |P_i|^2$;

In summary, amplitude (and the power spectrum) is a fundamental concept in the Fourier transform that is useful in addressing the problem of rotation-invariant feature extraction in image analysis[11]. By analyzing the amplitude of an image and combining the unique property of the NUDFT, we can obtain rotation information of the given images and thus serving as a satisfied computational basis and facilitating further analysis.

3.2.3 Principle component analysis

In the previous section we have shown that the amplitudes (equivalently power spectrum) of NUDFT of an image in polar coordinates are rotation-invariant features of the image. The rotation-invariant features of an image consist of the same size as the image with row being radial frequency and column angular frequency and have inherent noise from the image. Dimension reduction becomes crucial to simplify the computational complexity and mitigate noise effects. Principal Component Analysis (PCA) is a highly effective technique for dimension reduction[50]. We perform PCA on each image's feature matrix, considering rows (radial frequencies) as observation units and columns (angular frequencies) as variables

using the R function prcomp(). This application of PCA is not similar to its previous use for dimension reduction in Cryo-EM, where images are considered as observation units[88, 41]. The number of PC is selected based on the cumulative proportion of variance from each PC. Obtained PC scores are used in further analysis such as K-means clustering.

The rationale for employing PCA as the preferred dimension reduction technique in this context stems from its suitability in handling computational frequency information. The NUDFT yields polar-distributed frequency information rather than the original image intensities. As a result, the feature matrix now represents the original image in the frequency domain. Known that frequency information often exhibits significant contributions in the lower frequency range while minor contributions in other frequencies, and rotation-invariant decisions are partly based on these discrepancies, we aim to preserve this relationship as much as possible during the dimension reduction process. Moreover, the computation process of PCA possesses the ability to effectively reduce the noise present in the information as well as eliminate redundant information pertaining to certain aspects of the frequency information[80]. This characteristic contributes to the enhancement of performance in subsequent cluster processes.

3.2.4 K-means clustering

The rotation-invariant features calculated by the NUDFT serve the purpose of 2D classification. Our objective is to perform this classification task solely based on the features derived from projection images. Hence, an unsupervised clustering method is needed here.

K-means Clustering emerges as a suitable approach due to its simplicity and efficiency. K-means clustering operates by iteratively assigning data points to the cluster with the nearest mean, and then updating the cluster means based on the newly assigned points[44]. In our case, we create a distance matrix by defining the distance between images as an Euclidean distance between corresponding dimension reduced features by PCA. The created distance matrix is used as input to K-means algorithm, and the generated cluster means represent the centroids of the rotation-invariant features. By employing K-means clustering, we can

automatically group the rotation-invariant features into distinct clusters without any prior knowledge of the class labels. This allows us to perform the classification solely based on the similarities and dissimilarities between the rotation-invariant features.

Thus, K-means clustering is well-suited for this classification task as it leverages the mean value calculation to distinguish different frequency information, aligning with the features obtained from the NUDFT.

3.2.5 Experimental studies

We conduct experimental studies to evaluate the classification performance of combining K-means with our NUDFT based feature extraction (Algorithm 1) by comparing with K-means using rotation features extracted from conventional 2DFFT and CL2D used in the Xmipp software for the classification[83]. CL2D is a modified K-means method and uses correntropy as a measurement of similarity of input images. Our goal was to evaluate their performance by conducting the experiments under identical conditions while the only difference is the computation basis. This approach allowed us to observe and quantify the ability and advantages of NUDFT directly. We anticipate discovering intriguing results and gaining insights specifically from the NUDFT analysis.

Furthermore, following the acquisition of rotation-invariant information from the projection images through different FT processes, we apply PCA to perform dimension reduction as we explained prior. Subsequently, we apply the K-means clustering method to classify the rotation-invariant information after dimension reduction and generate distinct classes from the entirety of the projection images. This approach is more conducive to the analysis of the results derived from both Fourier transform processes.

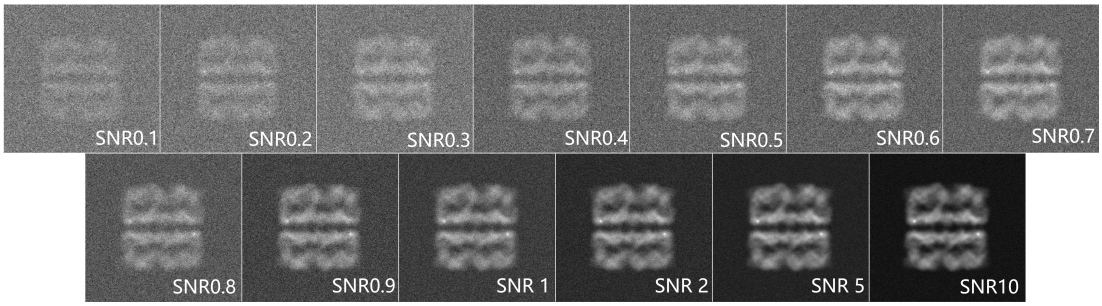


Figure 3.1: Different signal-to-noise ratios.

3.3 Experimental result

3.3.1 Synthetic data generation

To ensure a controlled computational environment and facilitate accurate comparisons, synthetic data was employed in this study. We generated projection images from EMDB 1080[56], a public real-world GroEL dataset, with size 256 $\times$ 256 (slice projection image shown in figure 3.2). 8 distinct class centres were thoughtfully selected to represent different projection angles, and within each class, 32 images were constructed to simulate the in-plane rotation by $\frac{\pi}{16}$ (different class centres are shown in figure 3.4). As a result, a comprehensive set of 256 projection images was successfully constructed, effectively simulating the clustering problem basis before applying image computational processes.

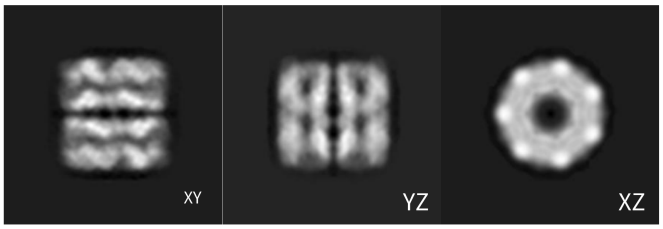


Figure 3.2: EMDB 1080

We simulate the noise environment by adding noise to the clean image generated prior. To be closer to reality, our SNR is defined based on variance:

$$SNR = \frac{Var(Signal)}{Var(Noise)} \tag{3.17}$$

In addition, unlike traditional noise that only uses Gaussian noise, different types of noise my appear during the imaging process, like auto-correlated noise[73]. Thus, we use auto-correlated noise to simulate different type of distributed noise that may appear in the imaging process, and use this to test how the NUDFT algorithm handles this type of noise. In our case, we use AR (2) sequence to generate auto-correlated noise:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \tag{3.18}$$

Where ε_t is white Gaussian noise.

We determine ϕ_1 and ϕ_2 values in the noise model (3.18) by tuning parameter ϕ_1 and ϕ_2 values such that the textural appearance of the experimental and simulated image can be matched quite well (figure 3.3), using the same idea in [26].

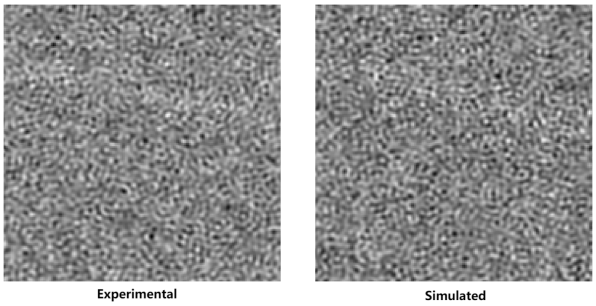


Figure 3.3: Constructed noise is similar with real noise.

To determine the noise variance, we first calculate the variance of the clean image as the signal variance. Then, according to the required SNR, we set the noise variance as $\frac{Var(Signal)}{Var(SNR)}$.

Finally, using the calculated noise variance and ϕ_1 and ϕ_2 values we generate correlated noises according to model (3.18), and add them to the clean projection image to finally generate noisy images with different SNR levels.

We generate noisy images under different SNR ranging from 0.1 to 10 (shown in figure 3.2).

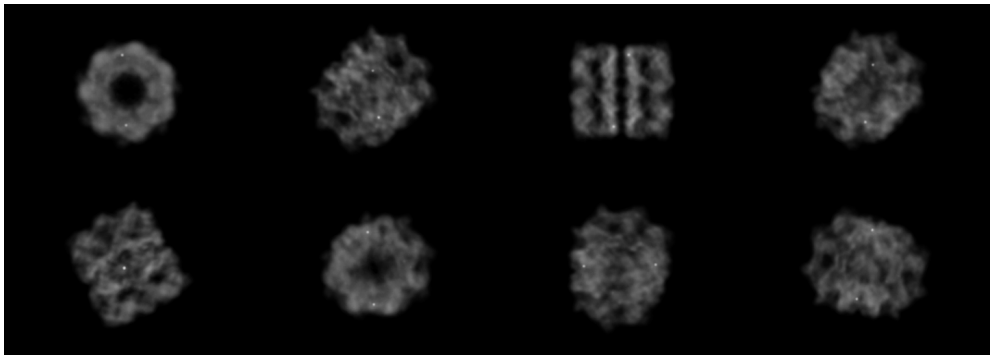


Figure 3.4: Different classes of projection images are specified by different projection angles.

3.3.2 PCA and K-means classification result examination

The computed rotation-invariant representations derived from the NUDFT method still maintain a relatively high dimensionality, leading to substantial computational overhead and potential confusion during subsequent classification procedures. To address these challenges, we employed PCA to perform dimensionality reduction on the data. In our case, the frequency information generated from the NUDFT algorithm has dimensions of 80×256, corresponding to 80 radial components and 256 angular components in the NUDFT. The number of the radial frequencies and angular frequencies used to calculate the NUDFT mainly depends on the input image dimensions and noise level. Normally, the number of frequencies corresponds to the number of pixels in the spatial domain image, i.e. the image in the spatial and frequency domain are of the same size as the Fourier transform represents the image in the frequency domain, and the representation can be transformed back to the original spatial domain image. In our data analysis here, we could calculate the NUDFT of 256 radial frequencies and 256 angular frequencies since the spatial domain images used have dimensions of 256 x 256. The SNR in the input image may also influence the optimal number of radial and angular frequencies- higher SNR more frequencies can be used. Lower SNR less frequencies should be used. We did experiments on different number of radial and angular frequencies and found that 80 radial components and 256 angular components work well in the range of SNR considered in this thesis. The computed rotation-invariant representations derived from the NUDFT method have

dimensions of 80x256, which still maintain a relatively high, leading to substantial computational overhead and potential confusion during subsequent classification procedures.

In PCA, dimension reduction is achieved by selecting the principal components according to the magnitude of their variances. A commonly used strategy is to choose the number of PCs for which the cumulative explained variance exceeds a threshold, e.g., 0.90 (90%). In this thesis, we use the 90% threshold to select the number of PCs. Thus the principle components generated from the PCA has dimensions of 80×80. This reduction aimed to streamline the dataset's dimensions, resulting in improved computational efficiency. As shown in Fig.3.5, in the condition of SNR = 0.1 (the highest noise level in our experiment studies), the cumulative explained variance by the first two PCs exceeds 90% of the variance of the generated rotation-invariant features. By the 90% threshold, we keep only the first two PCs.

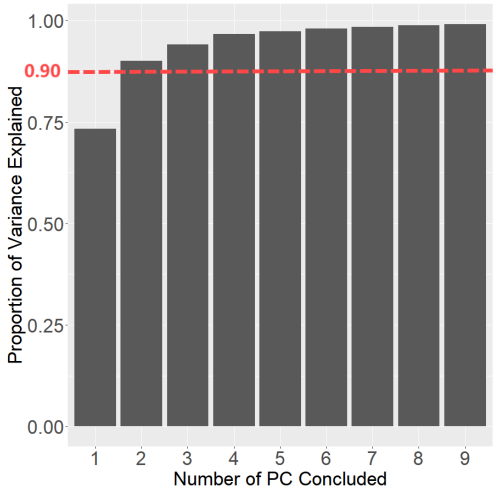


Figure 3.5: The cumulative explained variance by each PCs, the figure shows one case under SNR=0.1

Subsequent to the application of dimension reduction through PCA, an unsupervised classification approach was implemented, utilizing the k-means clustering algorithm. The primary objective of employing k-means clustering was to partition the dataset's data points into well-defined clusters, guided by the resemblance patterns extracted from the reduced-dimensional representation acquired

through PCA. In our case, we concluded the first two components (which has dimensions of 2×80) of each image in the K-means process, and we regarded these PCs as a long vector, thus each image's rotation-invariant features concluded in the K-means clustering has dimensions of 1×160, finally we use Euclidean distance to compute the difference between each images' rotation-invariant features (now we have transformed the original 80×256 frequency information into 1×160 length vector). The distance metrics we used in K-means clustering is defined as:

$$\text{EucDist}(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^\top (\boldsymbol{x} - \boldsymbol{y})} = \sqrt{\sum_{i=1}^{D} (\boldsymbol{x_i} - \boldsymbol{y_i})^2} \tag{3.19}$$

While x,y is each image's rotation-invariant features, in our case, its a vector.

In this unsupervised clustering task, we were also mindful of the potential influence of noise factors, which introduced a level of uncertainty regarding the appropriate number of clusters at the commencement of the computational process.Consequently, we resorted to the elbow point method as a principled technique to ascertain the optimal number of clusters for the k-means algorithm. This method entails an assessment of the within-cluster sum of squares (shown in figure 3.6) across a range of candidate cluster numbers. The pivotal 'elbow point' signifies a juncture where the pace of reduction in the sum of squares experiences a notable deceleration. This inflexion point is indicative of a potential optimal cluster count that attains a balance between model complexity and clustering fidelity. As delineated in the figure 3.6, the number of clusters suggested by the 'elbow point' harmonizes notably well with the ground truth number of clusters. Consequently, we embarked on a subsequent phase of rigorous scrutiny, assessing the classification accuracy vis-à-vis the identified clusters. Furthermore, we conducted a comparison of classification performance by leveraging the rotation-invariant representations computed via these distinct methods.

This multifaceted analytical exploration encompasses a comprehensive examination of the underlying structure, noise resistance, and information retention achieved through the synergy of dimension reduction and unsupervised clustering. The comparison of the 'elbow point' indication and the true cluster count substantiates the efficacy of the elbow method within the context of unsuper-

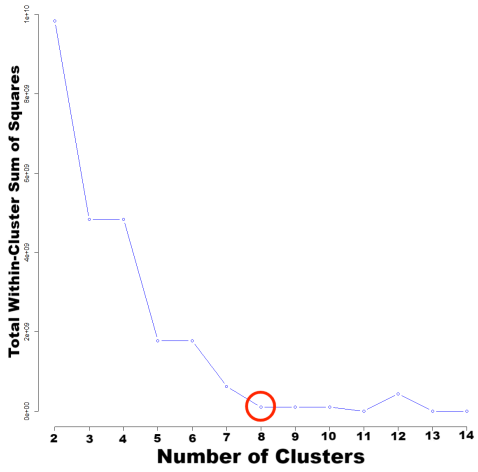


Figure 3.6: Within-Cluster sum of squares under different numbers of clusters, the figure shows one case under SNR=0.1

vised classification scenarios, thereby enhancing our confidence in the subsequent classification assessments conducted.

3.3.3 Final result comparison

In order to assess and compare the efficacy and robustness of noise of the developed algorithms on NUDFT processes, we conducted a rigorous evaluation of a real-world particle with different signal-to-noise ratios. This evaluation involved subjecting the algorithms to constructed datasets featuring different levels of noise, simulating SNRs ranging from 10 to 0.1 (shown in figure 3.1).

As we construct projection images from a pre-known 3D structure, we know each images correct class (or called category) initially, thus when we compute the clustering accuracy we can use a more precise and more realistic way to define the correctly classified images. In our case, the current category of the final clustering result classes is the initial cluster with the largest number of the input images in this class (because we know the correct class of the input images). Therefore, when the number of clusters is the same as the input data, the category with the largest number in each cluster is the current category (i.e. the input images are five images from each of the two categories, cluster 1 has four images from

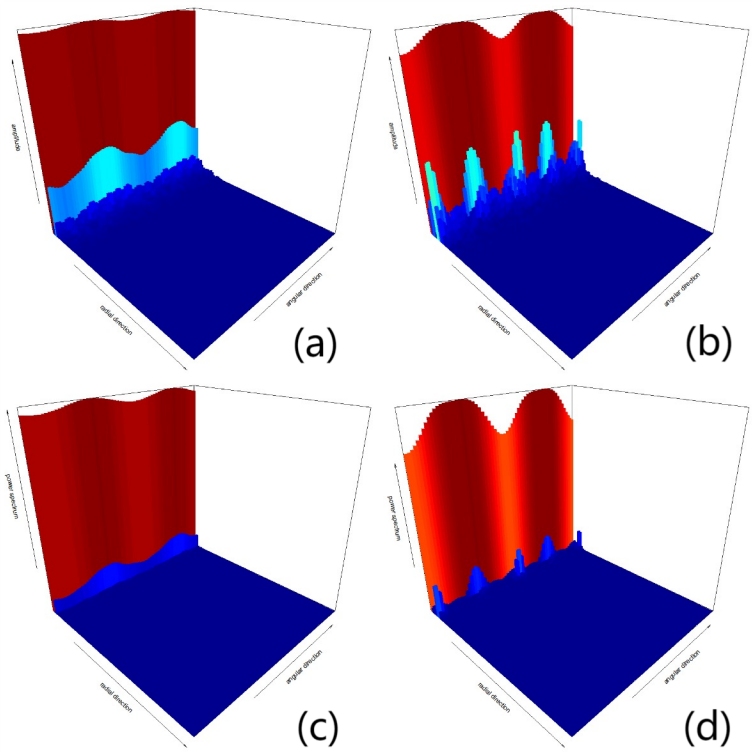


Figure 3.7: Power Spectrum generated from NUDFT
The figure shows one case under SNR=0.1 : (a) and (c) is the amplitude and power spectrum of one image while (b) and (d) are generated from a different class, (a) (b) and (c) (d) both show the clear difference between the two images, which means the rotation-invariant information generated by amplitude and power spectrum both valid.

class 1 and one image from class 2, and cluster 2 has one image from class 1 and four images from class 2, then the accuracy is (4+4)/10); when the number of clusters is different from the input data, the category with the largest number in each cluster is taken as the category of the current cluster (i.e. the input images are five images from each of two categories, cluster 1 has three images from class 1 and one image from class 2, cluster 2 has one image from class 1 and three images from class2, cluster 3 has one image from class 1 and one image from class 2, then the accuracy is (3+3)/10). We use this definition to define the correctly clustered images so that the computed accuracy are more in line with the actual application situation. In simple terms, the accuracy can be written as:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Samples}}{\text{Total Number of Samples}} \tag{3.20}$$

By conducting such a comprehensive evaluation (result shown in figure 3.8 and table 3.1), we sought to draw conclusive insights regarding the comparative performance of the NUDFT processes.

SNR	Xmipp	NUDFT	2DFFT	2DFFT-ZP
10	0.934 (0.930,0.938)	0.982 (0.977,0.984)	0.948 (0.941,0.949)	0.983 (0.977,0.984)
5	0.898 (0.895,0.906)	0.977 (0.972,0.981)	0.944 (0.934,0.945)	0.971 (0.965,0.973)
2	0.887 (0.879,0.891)	0.967 (0.957,0.973)	0.940 (0.926,0.941)	0.967 (0.957,0.969)
1	0.875 (0.863,0.879)	0.964 (0.953,0.969)	0.932 (0.918,0.934)	0.960 (0.945,0.961)
0.9	0.859 (0.845,0.871)	0.961 (0.949,0.965)	0.925 (0.906,0.930)	0.944 (0.926,0.945)
0.8	0.848 (0.828,0.856)	0.949 (0.934,0.957)	0.905 (0.883,0.910)	0.938 (0.922,0.941)
0.7	0.836 (0.809,0.852)	0.933 (0.914,0.949)	0.890 (0.863,0.902)	0.924 (0.898,0.934)
0.6	0.824 (0.789,0.8478	0.930 (0.902,0.945)	0.878 (0.844,0.891)	0.905 (0.867,0.918)
0.5	0.813 (0.770,0.844)	0.923 (0.883,0.941)	0.846 (0.805,0.875)	0.889 (0.844,0.906)
0.4	0.777 (0.691,0.828)	0.902 (0.859,0.934)	0.799 (0.754,0.836)	0.874 (0.828,0.898)
0.3	0.703 (0.648,0.762)	0.887 (0.836,0.928)	0.780 (0.727,0.820)	0.843 (0.789,0.875)
0.2	0.660 (0.609,0.738)	0.883 (0.830,0.926)	0.764 (0.703,0.812)	0.815 (0.754,0.852)
0.1	0.629 (0.559,0.703)	0.872 (0.816,0.922)	0.732 (0.668,0.789)	0.779 (0.715,0.824)

Table 3.1: Mean classification accuracy of each method generated from equation 3.20, in the bracket is the 95% confidence interval from 100 simulations.

Based on the experimental results presented in the table 3.1, it is evident that the NUDFT algorithm outperforms the 2DFFT and Xmipp algorithms in terms of overall signal-to-noise ratios. Although the performance of each method is

acceptable at high SNR conditions, the accuracy of Xmipp has a great influence as the SNR becomes lower. And the result proved that zero-padding can improve the accuracy of the Fourier transform as it benefits the interpolation process, but NUDFT still shows better capability in the overall performance while it does not conclude the zero-padding process. The NUDFT algorithm demonstrates enhanced robustness against noise and exhibits greater stability across varying levels of SNRs. Notably, in cases of higher SNR, the NUDFT algorithm exhibits a remarkable superiority in classification accuracy.

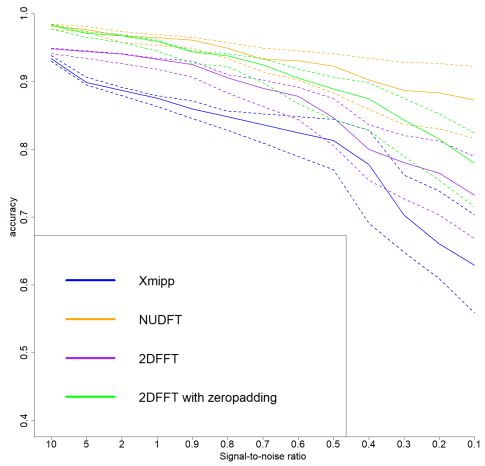


Figure 3.8: Classification accuracy of different methods
The blue line shows the accuracy of the NUDFT method, the red line shows Xmipp's method, the purple line shows the 2DFFT and the green line shows the 2DFFT with zero-padding enhancing its interpolation process.

It is worth noting that our experimental study is limited to a specific set of classes and a fixed number of overall projection images. The primary focus was to explore and compare the distinct behaviours and performances of the NUDFT processes.

For further investigations, the dataset could be expanded by including additional projection angle classes and varying levels of noise. The present comparison results provide valuable insights into the performance disparities arising from the inherent properties of the respective Fourier transform algorithms. These findings contribute to a better understanding of the strengths and limitations of NUDFT in handling noisy data and image classification tasks.

3.4 Summary and discussion

As discussed in the previous section, the NUDFT has demonstrated superior performance in the overall comparison, displaying enhanced noise resistance, particularly in low signal-to-noise ratio environments. These enhanced capabilities are attributed to its inherent property of deriving frequency information at any specified interest frequency point, allowing it to generate frequency information and extract rotation information directly, bypassing the interpolation step which avoids additional errors, which result in its robustness to noise and overall superior performance. Furthermore, it provides the flexibility to generate more frequency information for any given frequency domain and its unique property also grants it superior performance in handling auto-correlated noise.

Although our comparison is based on a limited number of projection images, the results clearly highlight the differences in the NUDFT process. We aim to emphasize the significance of NUDFT and its potential in addressing Cryo-EM 2D classification. By renewing the interest in NUDFT, we can further explore its capabilities and contributions to Cryo-EM research.

Chapter 4

Spectral clustering based on NUDFT features

4.1 Introduction

In this chapter we describe a method for classification of projection images in single-particle Cryo-EM by combining spectral clustering with our NUDFT based invariant feature extraction described in Chapter 3. We designed a novel distance metric to better describe the difference between images, it used both power spectrum and amplitude of the frequency information generated from the NUDFT, by using the unique property of the spectral clustering, this combination can have better classification performance. We build a complete algorithm including all these features and compared its performance with the algorithm we proposed in the Chapter 3, the result proved the effectiveness of this combination, our algorithm can have better classification accuracy in 2D classification tasks in single-particle Cryo-EM, and also demonstrates the unique properties of spectral clustering that facilitate the presentation and interpretation of the result.

In single-particel Cryo-EM, following the computation of rotation-invariant features, clustering methods are typically employed to group data. This involves categorizing images with similar projection angles but different in-plane rotations into distinct classes. This categorization essentially distinguishes all 2D images based on their initial three-dimensional spatial projection angles. The concept

of spectral clustering seamlessly aligns with this methodology. Diverging from traditional k-means clustering, spectral clustering leverages graph theory to define image similarity and construct an adjacency matrix, and thus providing a more fitting, direct, and intuitive approach to resolving 2D classification compared to conventional clustering methodologies, yielding superior classification outcomes.

We have amalgamated the advantages of the non-uniform discrete Fourier transform with spectral clustering to formulate an enhanced algorithm. This novel algorithm introduces the utilization of the non-uniform discrete Fourier transform, different to the conventional Fourier transform, for the representation of input projection images. Furthermore, we have established a more robust distance metric for characterizing the paired similarity between images' rotation-invariant features. This refined metric facilitates the incorporation of all input images into spectral clustering for data grouping, deviating from the traditional clustering methodology.

The envisioned benefits encompass heightened robustness in the presence of noise, an improved description of the representation of original projection images in the frequency domain, and a more explicit and precise articulation of spatial distances between images. The collaborative integration of the non-uniform discrete Fourier transform and spectral clustering shows a potential enhancement in this algorithmic framework's performance. This amalgamation promises to provide a refined and intricate methodology for representing images' relationship and clustering, particularly within the realms of Cryo-EM 2D projection image classification.

4.2 Methodology

4.2.1 Amplitude and power spectrum

In the pursuit of classifying projection images into distinct categories based on projection angles, significant challenges arise from variations not only in the projection angles but also in the in-plane rotation angles of the images. Consequently, it becomes imperative to devise methodologies for computing and extracting rotation-invariant features to effectively distinguish these images.

In Chapter 3, we describe a NUDFT based method for extraction of rotation-invariant features from images. The procedure involves applying a non-uniform discrete Fourier transform to the images, thereby representing them in the frequency domain. Within this frequency domain, a comprehensive analysis and computation of frequency information are conducted, leading to the derivation of rotation-invariant features. The frequency information obtained exhibit the unique property of remaining invariant under rotation.

Subsequent to this, a further modification is introduced through the application of another 1D Fourier transform in the angular direction. This additional step allows for the derivation of rotation-invariant information for the original images. This adjustment ensures that images with the same projection angle but different in-plane rotation angles produce identical invariant features under such computations. Consequently, the methodology facilitates the effective distinction and classification of all images solely based on their projection angles while preserving other relevant features.

Taking into account the noise factor inherent in Cryo-EM's projection process, typically following a Gaussian distribution, thus the frequency information of noise after the Fourier transform tends to concentrate in the high-frequency area. Simultaneously, the primary molecular structure is often more pronounced in the low-frequency part. This intrinsic characteristic enables the Fourier transform to effectively separate and mitigate the influence of noise in subsequent computational processes.

4.2.2 Spectral clustering

In the realm of 2D projection images in Cryo-EM, the application of spectral clustering is of paramount significance in tackling classification challenges[78]. Spectral clustering, grounded in graph theory, portrays images as nodes within a graph and constructs the graph's adjacency matrix based on image similarities, facilitating effective image clustering. This method boasts several advantages and characteristics that render it particularly well-suited for addressing the complexities of this problem.

In spectral clustering, data points (images) are treated as nodes on a graph,

and their pairwise similarities as edge weights (paired similarity or distance). Construct the mutual relationship of all images (nodes) into a matrix and use it to represent the connection relationship (edge weights or distance) between all nodes, which is the adjacency matrix. While each m_{ij} in adjacency matrix describe the distance between node i and node j. Then we can generate the corresponding degree matrix D by:

$$d_i = \sum_{j=1}^{n} m_{i,j} \tag{4.1}$$

While d_i represent the degree of ith node.

The Degree matrix is a diagonal matrix with the main diagonal being (d_1,...d_n):

$$\mathbf{D} = \begin{pmatrix} d_1 & \dots & \dots \\ \dots & d_2 & \dots \\ \vdots & \vdots & \ddots \\ \dots & \dots & d_n \end{pmatrix} \tag{4.2}$$

Then the corresponding Laplacian matrix L is defined as:

$$L = D - M \tag{4.3}$$

Where M is the adjacency matrix.

Then we can have the unique property of the Laplacian matrix as:

$$\begin{aligned} f^T L f &= f^T D f - f^T M f \\ &= \sum_{i=1}^{n} d_i f_i^2 - \sum_{i,j=1}^{n} m_{i,j} f_i f_j \\ &= \frac{1}{2}\left(\sum_{i=1}^{n} d_i f_i^2 - 2\sum_{i,j=1}^{n} m_{i,j} f_i f_j + \sum_{j=1}^{n} d_j f_j^2\right) \\ &= \frac{1}{2}\sum_{i,j=1}^{n} m_{i,j}(f_i - f_j)^2 \end{aligned} \tag{4.4}$$

Where f is the eigenvectors of the Laplacian matrix.

Thus, for the following clustering process, using the normalized cut criteria[28],

our objective is to find K clusters of images, $(A_1, \ldots, A_K)$ by minimizing the following normalized cut:

$$NCut(A_1, A_2, \ldots A_k) = \frac{1}{2}\sum_{i=1}^{k}\frac{W(A_i, \overline{A}_i)}{vol(A_i)} \tag{4.5}$$

Where $\mathrm{vol}(A_i) = \sum_{j\in A_i} d_j$*, which means the sum of the degrees of all nodes in the cluster* A_i*; and* $W(A_i, \overline{A}_i) = \sum_{i\in A_i, j\in\overline{A}_i} m_{i,j}$.

Then, introduce indicator vector $\mathbf{h}_j \in \{h_1, h_2, \ldots, h_k\},\ j = 1, 2, \ldots, k.$

$$h_{ij} = \begin{cases} 0 & , \text{node } i \notin A_j \\ \frac{1}{\sqrt{\mathrm{vol}(A_j)}} & , \text{node } i \in A_j \end{cases} \tag{4.6}$$

The optimization of indicator vectors as strictly 0-1 is highly complex and computationally intensive. Spectral clustering addresses this by introducing the Laplacian matrix, relaxing the optimization problem of indicator vectors into a continuous optimization problem. As a result, the indicator vectors are relaxed to real-valued vectors instead of strictly 0-1.

Thus, for $h_i^T L h_i$, we have:

$$h_i^T L h_i = h_i^T D h_i - h_i^T M h_i \tag{4.7}$$

And then subtract the relationship in equation (4.4), we can have:

$$h_i^T L h_i = \frac{1}{2}\sum_{k=1}^{n}\sum_{n=1}^{n} m_{k,n}(h_{ik} - h_{in})^2 \tag{4.8}$$

And then replace the $h_i k$ and $h_i n$ by the indicator vector in equation (4.6), we can have:

$$\begin{aligned} h_i^T L h_i = & \frac{1}{2}\left(\sum_{k\in A_i, n\notin A_i} m_{k,n}\left(\frac{1}{\sqrt{\mathrm{vol}(A_i)}} - 0\right)^2\right. \\ & \left. + \sum_{k\notin A_i, n\in A_i} m_{k,n}\left(0 - \frac{1}{\sqrt{\mathrm{vol}(A_i)}}\right)^2\right) \end{aligned} \tag{4.9}$$

Continuing to simplify the formula, we can get:

$$\begin{aligned} h_i^T L h_i =& \frac{1}{2}\Big(\sum_{k\in A_i, n\notin A_i} m_{k,n}\frac{1}{vol(A_i)} + \sum_{k\notin A_i, n\in A_i} m_{k,n}\frac{1}{vol(A_i)}\Big) \\ =& \frac{1}{2}\Big(cut(A_i,\overline{A}_i)\frac{1}{vol(A_i)} + cut(\overline{A}_i, A_i)\frac{1}{vol(A_i)}\Big) \\ =& \frac{cut(A_i,\overline{A}_i)}{vol(A_i)} \end{aligned} \tag{4.10}$$

Thus, our objective is to minimize the formula:

$$NCut(A_1, A_2, \dots A_k) = \sum_{i=1}^{k} h_i^T L h_i = \sum_{i=1}^{k} (H^T L H)_{ii} = tr(H^T L H) \tag{4.11}$$

However, in this scenario, $H^T H \neq I$, but if we consider:

$$h_i^T D h_i = \sum_{j=1}^{n} h_{ij}^2 d_j = \frac{1}{vol(A_i)} \sum_{j\in A_i} d_j = \frac{1}{vol(A_i)} vol(A_i) = 1 \tag{4.12}$$

Then, we can have: $H^T D H = I$

Then, our objective formula becomes:

$$\underbrace{arg\ min}_{H}\ tr(H^T L H)\ s.t.\ H^T D H = I \tag{4.13}$$

But, still, the indicator vector h in H is not an orthonormal basis, so we set:

$$H = D^{-1/2} F \tag{4.14}$$

Then, we can have:

$$H^T L H = F^T D^{-1/2} L D^{-1/2} F \tag{4.15}$$

$$H^T D H = F^T F = I \tag{4.16}$$

Subtract the relationship in equation (4.15) and (4.16) into equation (4.13),

we can have our final objective formula in spectral clustering:

$$\underbrace{arg\ min}tr(F^T D^{-1/2} L D^{-1/2} F)\ s.t.\ F^T F = I \tag{4.17}$$

It is worth noting that indicator vectors are based on discrete clustering, and the eigenvectors in spectral clustering lack direct interpretability. Their direct meaning is simply the projection positions in a lower-dimensional space rather than serving as the final classification results. However, by applying additional processing steps (such as clustering methods) can generate classification result from indicator vectors, but such process is different from the graph cut idea of the spectral clustering, which may lead to different result from spectral clustering and influence the final accuracy, this idea still need further investigate.

Primarily, spectral clustering enables the quantification and definition of image similarity through the structure of the graph. In Cryo-EM, 2D projection images exhibit variations not only in projection angles but also in in-plane rotation angles. Spectral clustering becomes instrumental in grouping images with diverse properties into distinct clusters, thereby differentiating their projection angles by using their structures and features.

Secondarily, as projection images are typically represented in high-dimensional spaces, spectral clustering plays a pivotal role in reducing these high-dimensional datasets to more manageable low-dimensional spaces. This reduction not only boosts computational efficiency but also facilitates an intuitive observation and representation of relationships between images. This, in turn, provides a more robust foundation for subsequent analyses and computations.

Moreover, spectral clustering demonstrates a noteworthy robustness to noise. In Cryo-EM, projection images often contend with substantial noise interference, particularly in scenarios characterized by low signal-to-noise ratios where conventional methods of directly computing image distances prove ineffective. The unique similarity measurement and dimension reduction properties of spectral clustering contribute to effective image classification, even in conditions with low signal-to-noise ratios. This underscores the resilience of spectral clustering as a valuable tool for image analysis in Cryo-EM, especially when confronted with challenging noise environments.

4.2.3 Algorithms

While the non-uniform discrete Fourier transform establishes a more accurate foundation for subsequent computational tasks involved in the classification and analysis of input projection images, our algorithm relies on the frequency information extracted through NUDFT for addressing the classification. Leveraging the principles of spectral clustering, we present an algorithm based on spectral clustering and NUDFT and directly compare its performance with K-means clustering in handling frequency information within non-supervised clustering tasks. Our objective is to demonstrate the exceptional adaptability and superior efficacy of spectral clustering compared to K-means clustering, highlighting the algorithm's overall proficiency in addressing classification under low signal-to-noise ratio conditions.

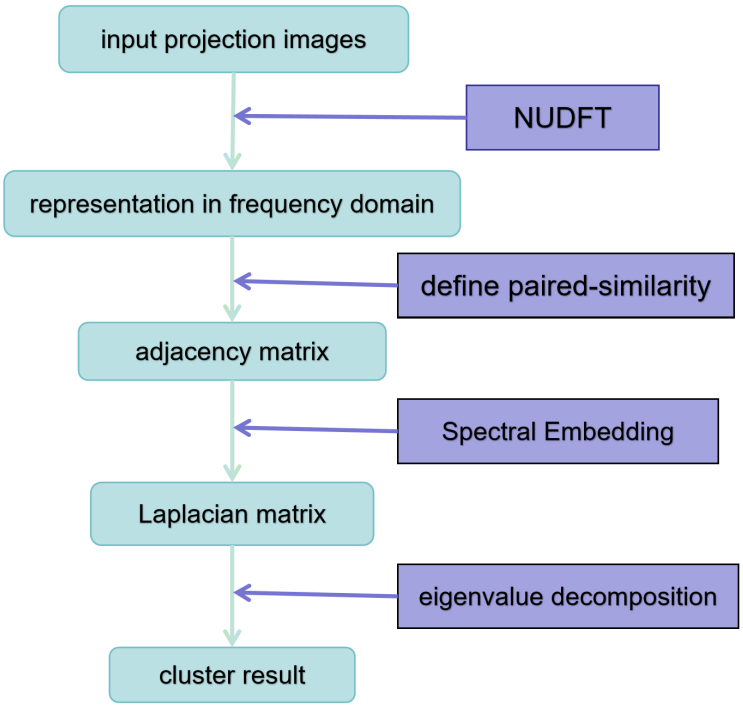


Figure 4.1: Proposed algorithm pipeline

First, we developed an algorithm combining spectral clustering (the full pipeline is shown in figure 4.1). Prior to applying the Fourier transform, it is important to select a region of interest within the projection images. In general, the structure

of molecules is concentrated in the centre of the image, while the four corners contain only noise information. Thus, we choose a circular region with a diameter consistent with the width of the square projection image.

After obtaining a circular mask of the original projection image, we apply the non-uniform discrete Fourier transform. By considering the centre of the circle as the origin of the polar coordinate system, we can set all frequency points of interest in a polar format, which facilitates further computation. During the Fourier transform, we represent the original information under the frequency domain without adding or reducing any data. If the number of data points differs from the original input data, this is simply an interpolation artefact in the frequency domain. Therefore, we set the total number of data points in the frequency domain to be close to the original input data.

After applying the non-uniform discrete Fourier transform to the projection images, we obtain the frequency information for subsequent computation. However, the frequency information obtained still contains a rotation factor, albeit limited to column transformation. Since the frequency points are in a polar coordinate system, we apply the Fast Fourier transform to the frequency information (generated from the NUDFT) along the angle direction to offset the rotation factor. Consequently, the power spectrum from this FFT yields rotation-invariant information, serving as a novel representation obtained from the original projection images.

Unlike other methods may use Gaussian kernel function, pixel intensity or 2D correlation coefficient to describe the similarity to construct the adjacency matrix[57, 93, 94]. As we obtained the rotation-invariant information from the original input projection images, we then can define the distance between images to construct a similarity matrix which follows the idea in spectral clustering. We defined the distance by using both the power spectrum and the amplitude to describe a more precise relationship between images under low SNR conditions, this is due to the property of both intensities. As power spectrum is typically used in these tasks, but its mathematical property causes it to have less sensitivity to small differences in the main structure compared to the large volume of the noise factor, such as small changes in projection angles, this may lose the ability in capture and distinguish two images with similar but different projection

angles or sometimes when the molecules rotated but show similar input intensity in projection images. While amplitude can preserve more information and differences under low SNR conditions, but it has less performance in describing large differences compared to the power spectrum. We will show this difference in later sections. Thus we take into consideration both power spectrum and amplitude in defining distance between images.

Considering the Fourier transform of the image among data sets is $F(u, v)$, then the amplitude and power spectrum of the Fourier transform is:

$$Amplitude : A(F) = |F(u, v)| \tag{4.18}$$

$$Powerspectrum : P(F) = |F(u, v)|^2 \tag{4.19}$$

Then we can have our paired similarity S_{ij} between two image i and j as:

$$D_1 = \sum_{i,j=1}^{N} \sqrt{(A_i(F) - A_j(F))^2} \tag{4.20}$$

$$D_2 = \sum_{i,j=1}^{N} \sqrt{(P_i(F) - P_j(F))^2} \tag{4.21}$$

$$S_{i,j} = \sum_{i,j=1}^{N} \frac{D_1}{\max(D_1)} + \frac{D_2}{\max(D_2)} \tag{4.22}$$

After obtaining the affinity matrix (also can be regard as similarity matrix or distance matrix) followed by ideas in spectral clustering, we then defined the adjacency matrix by using a fully connected approach, thus in our case, there's no any threshold to modify the affinity matrix (similarity matrix), then the adjacency matrix is same as the affinity matrix.

Thus, in our case, the adjacency matrix can be written as:

$$W_{i,j} = S_{i,j} = \sum_{i,j=1}^{N} \frac{D_1}{\max(D_1)} + \frac{D_2}{\max(D_2)} \tag{4.23}$$

And the degree matrix can be written as:

$$D_i = \sum_{j=1}^{N} W_{i,j} \tag{4.24}$$

Following we can compute the degree matrix to finally obtain the Laplacian matrix. And apply eigenvalue decomposition to the Laplacian matrix to generate its eigenvectors.

Different from choosing the first N eigenvectors specified by eigenvalues in the clustering process, we take all eigenvectors into further computations to preserve as much information as possible, and this still can provide satisfy result. Finally, we apply K-means clustering to the Laplacian matrix's eigenvectors to get the final cluster result.

Algorithm 2: Non-uniform discrete Fourier transform based spectral clustering featured 2D image classification algorithm

Input: numbered projection images I_i
Output: classified clusters Ci of the projection images

1. Apply algorithm 1 ;
2. Compute the corresponding power spectrum PP_i' and the amplitude PA_i' from algorithm 1;
3. Construct the similarity matrix by using the defined paired-similarity formula with PP_i' and PA_i' to get mS;
4. Construct the adjacency matrix by using fully-connected approach on mS to get mW;
5. Construct the corresponding degree matrix with mW to get mD;
6. Generate the corresponding Laplacian matrix mL;
7. Apply eigenvalue decomposition on mL to get eigenvectors FE;
8. Use the normalized cut criteria to generate the final clustering Ci;

According to the above description, we can achieve classify the projection images according to different projection angles. While this algorithm uses spectral clustering followed by the idea of graph theory to modify and analyse the frequency information obtained from the NUDFT, we conduct a direct comparison using K-means clustering to handle the rotation-invariant information obtained from the Fourier transform. This algorithm has the same steps in obtaining the

frequency information but only different in the clustering method, it first applies a principle component analysis to the rotation-invariant information obtained from the NUDFT and then applies the K-means clustering application to generate the cluster result. This approach imitates the process outlined in spectral clustering, where PCA also projects high-dimensional data into a lower-dimensional space and the following clustering result is obtained from data after the dimension reduction. We expect to see the superior performance of spectral clustering in handling the non-linear and unintuitive relationships between images' rotation-invariant information.

Algorithm 3: Non-uniform discrete Fourier transform based K-means clustering featured 2D image classification algorithm

Input: numbered projection images I_i
Output: classified clusters Ci of the projection images
1 Apply algorithm 1;
2 Apply principle component analysis and take the first two PCs on P_i' to get DP_i;
3 Apply K-means clustering and generate the final clustering result on DP_i to get Ci;

4.3 Experimental result

4.3.1 Synthetic data generation

To ensure a controlled computational environment and facilitate accurate comparisons, synthetic data was employed in this study. A 3D structure was meticulously designed to emulate the characteristics and properties of the actual particles encountered in Cryo-EM projection processes, and the formula we used to define the 3D structure is adjusted and improved on the basis of Anoshina's work[6], we apply some modification to the formula to make it more realistic. Subsequently, 2D projection images were generated from the constructed structure (shown in figure 4.2) to serve as the basis for further modifications. The synthetic data is defined in cylindrical coordinates (r,θ,h) by the following formula:

First, we replace r by $\hat{r}$ with:

$$\hat{r} = r + 3|\sin(3\theta)| - 0.8; \tag{4.25}$$

$$r \in [0, 128], \quad \theta \in [0, 2\pi]$$

Next, we define the intensity function with:

$$I(\hat{r}, \theta, h) = \begin{cases} f(\theta) \cdot g(\hat{r}, h), & \text{if } \hat{r} < R \\ 0, & \text{otherwise} \end{cases} \tag{4.26}$$

where:

$$f(\theta) = \eta + 20|\sin(1.5\theta)| + 3 \tag{4.27}$$

$$g(\hat{r}, h) = \begin{cases} \frac{9\hat{r}^2}{R^2 - h^2}, & h^2 + 9\hat{r}^2 \in [0, \frac{R}{3}) \\ 1, & h^2 + 9\hat{r}^2 \in [\frac{R}{3}, R] \\ 0, & h^2 + 9\hat{r}^2 > R \end{cases} \tag{4.28}$$

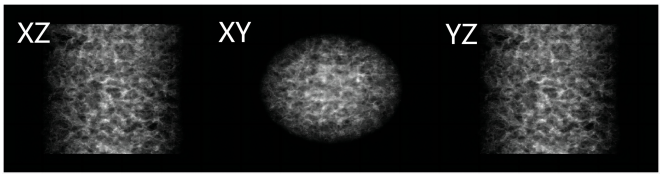


Figure 4.2: Phantom constructed

4.3.2 Projection images generation

Projection images were generated from the previously constructed 3D structure, incorporating 3D rotation to yield distinct projection angles. We thoughtfully selected ten distinct projection angles to generate corresponding projection images which resulted in 10 different classes in later computation, then a bi-linear interpolation method was employed to simulate in-plane rotations by pi/ 16. Furthermore, the previously generated noise was incorporated into the projection

images. We take different levels of noise to construct different signal-to-noise ratio scenarios.

We simulate the noise environment by adding noise to the clean image generated prior. Since we already know the 3D structure, we can calculate the variance of the structure to create and add noise according to different SNR requirements. To be closer to reality, our SNR is defined based on variance:

$$SNR = \frac{Var(Signal)}{Var(Noise)} \tag{4.29}$$

In addition, unlike traditional noise that only uses Gaussian noise, different types of noise my appear during the imaging process, like auto-correlated noise[73, 38]. Thus, we use auto-correlated noise to simulate different type of distributed noise that may appear in the imaging process, and use this to test how the NUDFT algorithm handles this type of noise. In our case, we use AR (2) sequence to generate auto-correlated noise:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \tag{4.30}$$

Thus, we set the noise variance according to the required SNR and then add them to the clean projection image to finally generate noisy images with different SNR levels.

As a result, in each SNR condition, a comprehensive set of 1600 projection images with size 256×256 was successfully constructed (10 classes, 32 in-plane rotations, 5 repeats under each angle), effectively simulating the clustering problem

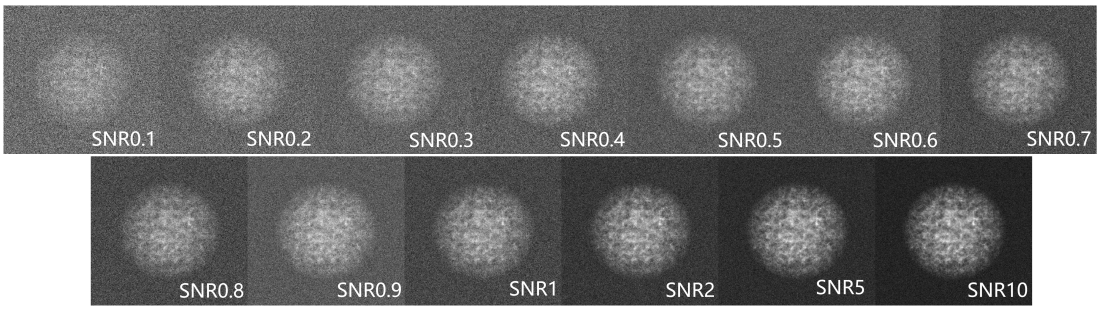


Figure 4.3: Different signal-to-noise ratios

basis before applying any further computation or clustering methods. (different classes of projection images constructed shown in figure 4.4)

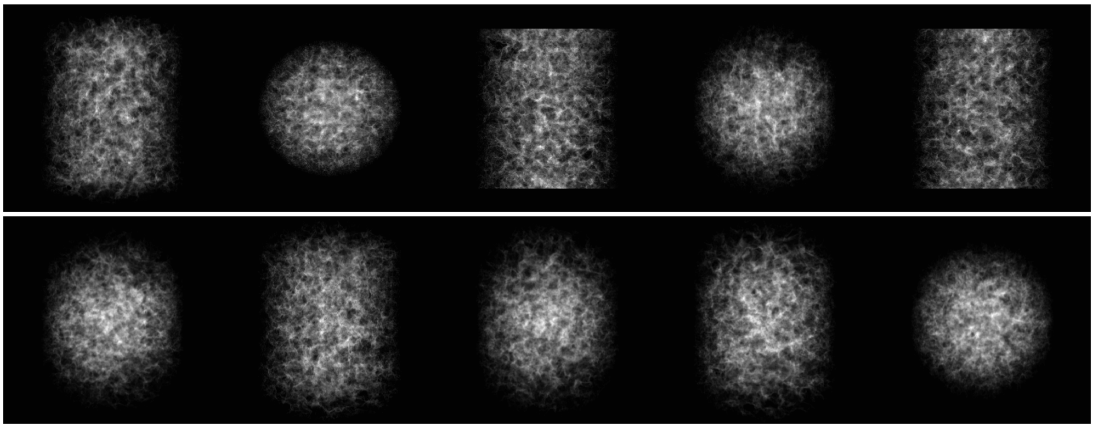


Figure 4.4: Different classes of projection images are specified by different projection angles.

4.3.3 Details in applying spectral clustering method

Given that the non-uniform Fourier transform has represented the original image in a polar coordinate system, we are now poised to compute the power spectrum and amplitude for further computations. Using the defined paired-similarity we can construct the adjacency matrix followed by spectral clustering's method, and our novel definition for paired-similarity shows more obvious differences in the corresponding distance matrix, which also shows in the overall cluster result. In our case, the frequency information obtained from the NUDFT process has dimensions of 80 $\times$ 256, which corresponding to 80 radial components and 256 angular components in total, and the distance metrics we used was computed from the amplitude and power spectrum (shown in equation (4.20), (4.21) and (4.23)), the distance matrix (or the similarity matrix) now has dimensions of 1600 $\times$ 1600, corresponding to 1600 images concluded in our experiment. Then we compute the following degree matrix and finally the corresponding Laplacian matrix, which also has dimensions of 1600 $\times$ 1600. We then apply eigen decomposition to the Laplacian matrix and we use the first two eigen vectors to show the complex internal structural relationships among all input projection images. Through

the visual display of the eigenvectors, we can find that the images are distributed into different clusters under this lower dimensional space. Then we applied the K-means clustering method to all Laplacian matrix's eigenvectors, thus we obtained the clustering result. Its worth noting that, while the eigenvectors obtained from the eigen decomposition has dimensions of 1600 $\times$ 1600, we can use first two eigenvectors (indicate by the two largest eigenvalues) to show the relationship of initial input 1600 images under 2D, and then we apply K-means clustering to all eigenvectors, by only using Euclidean distance in K-means clustering, the result is actually equivalent to normalized cut.

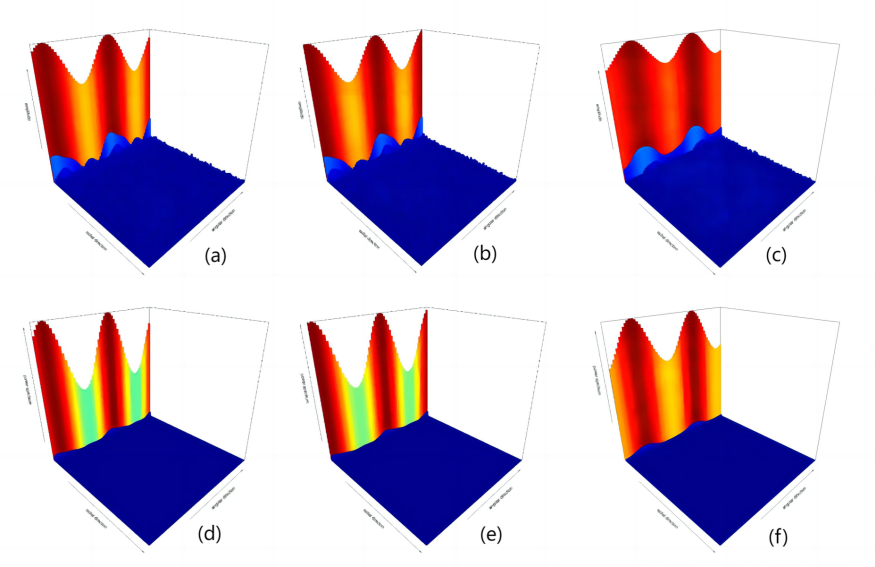


Figure 4.5: Power spectrum and amplitude generated from NUDFT. The figure shows one case under SNR=0.1 : Amplitude(a,b,c) and power spectrum(d,e,f) were generated from NUDFT, (a)(d),(b)(e) and (d)(f) were generated from three different images while (a)(d) and (b)(e) are two images with same projection angle but different in-plane rotation angles, and (d)(f) is from different projection angle.
(a) and (b) shows the amplitude has translation in angular direction while (d) and (e) show power spectrum has the same translation as well, and (c) and (f) both show different distributions with others.

As we construct projection images from a pre-known 3D structure, we know each images correct class (or called category) initially, thus when we compute the clustering accuracy we can use a more precise and more realistic way to define the

correctly classified images. In our case, the current category of the final clustering result classes is the initial cluster with the largest number of the input images in this class (because we know the correct class of the input images). Therefore, when the number of clusters is the same as the input data, the category with the largest number in each cluster is the current category (i.e. the input images are five images from each of the two categories, cluster 1 has four images from class 1 and one image from class 2, and cluster 2 has one image from class 1 and four images from class 2, then the accuracy is (4+4)/10); when the number of clusters is different from the input data, the category with the largest number in each cluster is taken as the category of the current cluster (i.e. the input images are five images from each of two categories, cluster 1 has three images from class 1 and one image from class 2, cluster 2 has one image from class 1 and three images from class2, cluster 3 has one image from class 1 and one image from class 2, then the accuracy is (3+3)/10). We use this definition to define the correctly clustered images so that the computed accuracy are more in line with the actual application situation. In simple terms, the accuracy can be written as:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Samples}}{\text{Total Number of Samples}} \tag{4.31}$$

4.3.4 Spectral clustering and K-means overall result examination

In order to assess and compare the overall performance of the two algorithms, we computed the clustering accuracy under different SNR conditions, ranging from 10 to 0.1. (shown in figure 4.3) Notably, the datasets consisted of 1600 projection images in each SNR category, and within the same class, only in-plane rotations were considered without any additional shifts or translations. The classification accuracy of the algorithms was determined by checking the clustering result with the original version of each class generated from the synthetic data. By conducting such a comprehensive evaluation, we sought to draw conclusive insights regarding the comparative performance of the combination of spectral clustering and NUDFT.

Based on the presented experimental results in figure 4.5, 4.7, 4.6, 4.8 and ta-

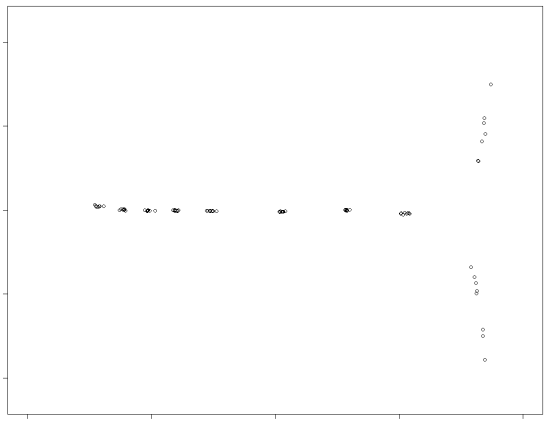


Figure 4.6: Eigenvertors of the images after spectral embedding. The figure shows one case under SNR=0.1 : the groups of data points show the distinction of images with different projection angles after the computation, which proves the performance of the algorithm.

ble 4.1, it is evident that the spectral clustering method outperforms the K-means method across various signal-to-noise ratios. The spectral clustering method exhibits enhanced robustness against noise, demonstrating greater stability and significantly higher classification accuracy, especially in scenarios with higher SNR levels.

Conversely, the K-means method shows a notable decline in accuracy as the noise level increases, indicating its susceptibility to noise influence. It's important to acknowledge that our experimental study is constrained to a specific set of classes and a fixed number of overall projection images. The primary aim was to explore and compare the distinct behaviours and performances of the two clustering processes based on non-uniform discrete Fourier transform's frequency information.

4.4 Conclusion and discussion

As described in the previous section, the spectral clustering has demonstrated an overall superior performance in classifying 2D projection images under different levels of signal-to-noise ratios, while combining the frequency information obtained from the non-uniform discrete Fourier transform. This implies that the

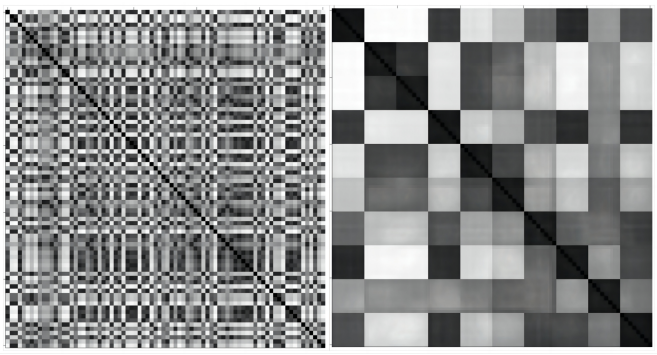


Figure 4.7: Distance matrix of the data.
(The figure shows one case under SNR=0.1 and only part of the data is selected for the simplicity of the picture, or it will be very hard to read.)
the left image shows the paired similarity computed from the defined formula of the original input images, and the right image shows the similarity after the classification, the distance matrix was sorted and shown in order, this also proves the performance of the classification.

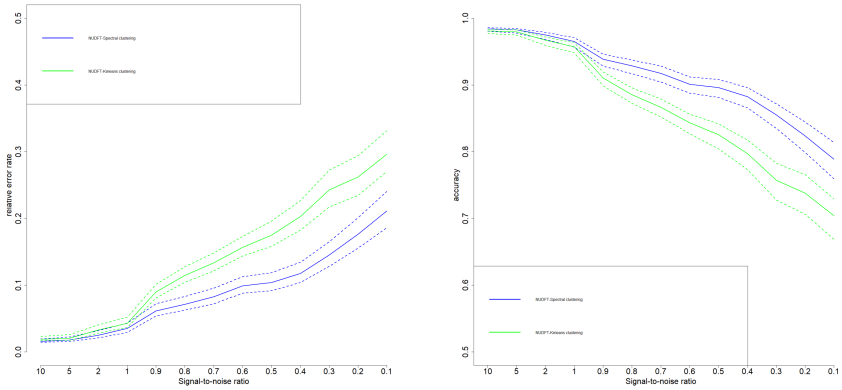


Figure 4.8: Left: error rate of two algorithms under each SNR. Right: accuracy of two algorithms under each SNR.
The green line is the algorithm using spectral clustering and the blue line is the algorithm using K-means clustering.

SNR	NUDFT-Spectral clustering	NUDFT-Kmeans
10	0.9842 (0.9806,0.9856)	0.9809 (0.9769,0.9831)
5	0.9824 (0.9775,0.9844)	0.9796 (0.9744,0.9825)
2	0.9749 (0.9681,0.9788)	0.9671 (0.9594,0.9719)
1	0.9649 (0.9569,0.9707)	0.9571 (0.9481,0.9631)
0.9	0.9386 (0.9281,0.9462)	0.9103 (0.8988,0.9194)
0.8	0.9286 (0.9169,0.9375)	0.8853 (0.8725,0.8956)
0.7	0.9174 (0.9044,0.9281)	0.8665 (0.8519,0.8788)
0.6	0.9011 (0.8875,0.9119)	0.8434 (0.8269,0.8562)
0.5	0.8961 (0.8812,0.9081)	0.8253 (0.8044,0.8419)
0.4	0.8824 (0.8656,0.8962)	0.7971 (0.7731,0.8175)
0.3	0.8549 (0.8350,0.8719)	0.7571 (0.7275,0.7825)
0.2	0.8236 (0.7988,0.8444)	0.7378 (0.7062,0.7656)
0.1	0.7886 (0.7594,0.8138)	0.7039 (0.6688,0.7294)

Table 4.1: Mean classification accuracy of each method generated from equation 4.31, in the bracket is the 95% confidence interval from 100 simulations.

unique property of spectral clustering which follows the ideas in graph theory has very good adaptability and effectiveness for 2D classification in Cryo-EM, by defining a more precise distance to describe the differences between images, we can use spectral clustering to show the spatial relationship of images intuitively, this can much improve the final classification accuracy compared to conventional method only use direct euclidean distance to describe the similarity. While the computation of the Laplacian matrix and eigen decomposition have similar ideas with principle component analysis's dimension reduction process, but a well-constructed Laplacian matrix helps provide a better basis for projecting onto low-dimensional data. Also, by combining the NUDFT, it leveraged and absorbed the advantages of the NUDFT which allows it to better detect and distinguish the rotation-invariant features of the input projection images, because of a more accurate description basis of rotation-invariant information of input projection images, our defined paired-similarity thus can be effective in the following computation. Also, this combination makes it have great robustness to noise and better computation efficiency, by using the NUDFT, we can facilitate the property of the Fourier transform to remove some of the noise effects through frequency separation, as the low SNR conditions is the biggest limitation

applies to all Cryo-EM computation tasks, thus this is a very important essential element.

Although our comparison is based on a limited number of projection images, the results clearly highlight the differences between the two different clustering processes under different levels of signal-to-noise ratios. Also, we use a fully connected method in defining neighbours in constructing an adjacency matrix in applying spectral clustering, we expect to in future we can use some method to obtain a part of the prior information of the data in advance, and then a more accurate threshold can be set, and a more accurate adjacency matrix can be set using the one-way connection neighbour distance threshold. However, in this paper, we aim to emphasize the significance and effectiveness of the combination of spectral clustering and the non-uniform discrete Fourier transform, and its potential in handling Cryo-EM 2D projection images. This method has great potential in 2D classification and benefits a more precise basis for further computation and analysis in Cryo-EM.

Chapter 5

NUDFT based 2D in-plane rotation angle estimation

5.1 Introduction

In this chapter we explore the applicability of the NUDFT in determining the 2D in-plane rotation angle. Different from the traditional FT that needs to apply interpolation for calculation of FT in polar frequency domain, NUDFT can calculate FT of an image in polar frequency domain directly, thus convert the rotation relationship between two images into the translation relationship in one step. Using the NUDFT in Chapter 2, we build an algorithm for estimation of rotation between a pair of images and perform experimental studies to examine its performance by comparing its estimation accuracy with the traditional 2DFT .

Image alignment is the process of matching two or more images to align them in a specific position. Consider two images $I_1(x, y)$ and $I_2(x, y)$ with size of $M \times M$ that are related by rotation θ_0:

$$\begin{aligned} I_2(x, y) =& I_1((x\cos\theta_0 + y\sin\theta_0) \\ & - \Delta x, (-x\sin\theta_0 + y\cos\theta_0) - \Delta y) \end{aligned} \tag{5.1}$$

In general, image alignment aims to estimate (Δx,Δy) and θ_0 by maximizing some similarity metrics between I_1 and I_2, and overlay the two images using

the estimated rotation and translation. Image alignment has a wide range of applications, including cryo-electron microscopy (Cryo-EM) and medical imaging [11, 67]. Intuitively, estimating (Δx,Δy) and rotation θ_0 can be performed by minimizing the following objection function:

$$\begin{aligned} L(\Delta x, \Delta y, \theta_0) = \sum_{x,y} (I_1((x\cos\theta_0 + y\sin\theta_0) \\ - \Delta x, (-x\sin\theta_0 + y\cos\theta_0) - \Delta y) \\ - I_2(x, y))^2 \end{aligned} \tag{5.2}$$

However, solving nonlinear minimization of equation (5.2) can be computationally expensive when the large number of pairs of images need to be aligned such as in single-particle Cryo-EM[40],

To address this problem, based on the idea of combining polar transform and Fourier transform, various algorithms have been proposed to estimate the rotation and translation shift parameters[72, 49, 100, 36, 40].

In this chapter we focus on estimation of rotation between pairs of images. Furthermore, our rotation estimation is based on converting rotation in spatial domain into translation polar frequency domain and then estimating translation (cf. Algorithm 4). Our method can be easily adapted to estimate translation between pairs of images.

In polar coordinates, the rotation by θ_0 is reduced to a shift over θ_0 along the angle coordinate. Thus, estimation of the rotation θ_0 can be solved as estimating 1D translational shift between two images.

We have developed two algorithms for rotation estimation, one is based on the non-uniform discrete Fourier transform and another based on the conventional Fourier transforms. We conducted a meticulous experimental setup, ensuring controlled data conditions and carefully considering the distinctions between these two algorithms. Our primary aim was to underscore the performance differences arising from the distinct Fourier transform processes. By doing so, we sought to demonstrate the applicability of the non-uniform discrete Fourier transform in 2D alignment and elucidate the unique characteristics and advantages it offers to single-particle Cryo-EM 2D alignment.

5.2 Methodology

5.2.1 Fourier alignment

In Chapter 2 we have demonstrated that by applying our NUDFT to an image and its rotated copy by θ_0, the spectrums (power spectrums) of two images are translated by θ along the angle direction in polar frequency domain (cf. equation (3.11)). Thus, rotation of pared images in spatial domain can be estimated by estimation of translation between their corresponding spectrums. Since the peak of cross-correlation directly related to the translation between the two images[7]. This property enables the utilization of cross correlation to estimate the translation between two images. Hence, we propose to estimate rotation between two images by searching for the peak of the cross correlation of their corresponding power plectrums obtained by our NUDFT.

By harnessing the frequency information extracted through the Fourier transform, the proposed method facilitates a more accurate estimation of the in-plane rotation angle, and it proves to be more robust against noise-induced factors. The ability to effectively analyze images in the frequency domain enhances the alignment process, leading to improved image registration and subsequent three-dimensional reconstruction in Cryo-EM studies.

5.2.2 Algorithms

We conducted a comparison between NUDFT and 2DFFT for the 2D alignment. Our goal was to evaluate their performance and differences by developing algorithms for both methods and conducting the experiments under identical conditions while the only difference is the Fourier transform process. This approach allowed us to directly observe and quantify the differences between NUDFT and 2DFFT. We anticipate discovering intriguing results and gaining insights specifically from the NUDFT analysis.

First, we developed an algorithm for NUDFT. Prior to applying the Fourier transform, it is important to select a region of interest within the projection images. In general, the structure of molecules is concentrated in the centre of the image, while the four corners contain only noise information. Thus, we choose a

circular region with a diameter consistent with the width of the square projection image.

After obtaining a circular mask of the original projection image, we apply the non-uniform discrete Fourier transform. By considering the centre of the circle as the origin of the coordinate system, we can set all frequency points of interest in a polar coordinate system, and then use the polar coordinate system which facilitates further computation. During the Fourier transform, we represent the original information under the frequency domain without adding or reducing any data. If the number of data points differs from the original input data, this is simply an interpolation artifact in the frequency domain. Therefore, we set the total number of data points in the frequency domain to be close to the original input data.

After applying the NUDFT to the projection images, we obtain the power spectrum for subsequent computation. According to the 2D Fourier shift theorem, the analysis of power spectrum information derived from the Fourier transform process enables the determination of in-plane rotation angles.

Algorithm 4: Non-uniform discrete Fourier transform based power spectrum featured in-plane rotation angle estimation algorithm

Input: Images I_1 and I_2
Output: Rotation angle θ between I_1 and I_2

1. Select circle region R of the projection image with the diameter same as the image's width;
2. Apply Non-uniform discrete Fourier transform on I_1 and $I_2$to get F_1 and F_2;
3. Compute the power spectrum to get P_1 and P_2 by $P_i = |F_i|^2$;
4. Find the displacement at which the maximum correlation in the paired power spectrum map P_1 and P_2 then return the paired displacement difference to get C_1 and C_2;
5. Convert C_1 and C_2 from angular frequency direction's translation into rotation angles to get paired in-plane rotation angle θ;

In the context of the conventional 2D Fourier transform methodology, an analogous algorithm has been devised. This approach also initiates with the selection of a region of interest, delineated by a circular area with a diameter matching that of the image. By subjecting the original image to a 2D Fourier

transform, its information is represented in the frequency domain. However, the coordinates representing frequencies in this frequency domain, as generated by the traditional Fourier transform, are uniformly distributed and conform to a Cartesian coordinate format.

In order to distil the in-plane rotation angles from the frequency domain information, a series of steps is executed. This encompasses the computation of the power spectrum, followed by its transformation into a polar coordinate system through the utilization of bi-linear interpolation techniques. This transformation is pivotal in aligning the outcomes with the non-uniform discrete Fourier transform algorithm previously expounded upon. It is important to highlight that the fundamental distinction between these approaches resides in the Fourier transform procedure itself.

Algorithm 5: Conventional 2D Fourier transform based power spectrum featured in-plane rotation angle estimation algorithm

Input: Images I_1 and I_2
Output: Rotation angle θ between I_1 and I_2
1 Select circle region R of the projection image with the diameter same as the image's width;
2 Apply 2D Fourier transform on I_1 and I_2 to get CF_1 and CF_2;
3 Apply FFT shift on CF_1 and CF_2 to get SF_1 and SF_2;
4 Compute the power spectrum to get CP_1 and CP_2 by $CP_i = |SF_i|^2$;
5 Apply interpolation to perform a polar transformation on CP_1 and CP_2 and to get CP_1' and CP_2';
6 Find the displacement at which the maximum correlation in the paired power spectrum map CP_1' and CP_2' then return the paired displacement difference to get P_1 and P_2;
7 Convert P_1 and P_2 from angular frequency direction's translation into rotation angles to get paired in-plane rotation angle θ;

Typically, the core structural details are encompassed within the initial 5-10 percent of the low-frequency domain. Given the notably diminished SNR inherent in projection images, despite our employment of the Fourier transform for noise mitigation, the prevailing environment remains inundated with substantial noise. Consequently, we deduce an approximate 5 percent for further computational considerations, focusing on the lower frequency spectrum.

Employing a translation shift on the power spectrum of the selected frequencies, we conduct paired correlation computations between the power spectra of distinct images. This process facilitates the determination of shift disparities between images, with the maximal correlation being the criterion of choice. Notably, the shift disparity, occurring within the angular direction of the Polar format frequency domain, can be directly translated into in-plane rotation angles of the original projection images. This transition assumes that the resolution of angular frequency comprehensively encompasses all feasible in-plane rotation angles, thus yielding the extraction of in-plane rotation angles.

Subsequent to determining the paired in-plane rotation angles, we assess the algorithm's accuracy. Employing a bi-linear interpolation algorithm, we revert all projection images with identical projection angles to their original orientations. This produces a class-average image with enhanced SNR, thereby facilitating an evaluation of the class-average image's performance.

5.3 Experimental result

5.3.1 Synthetic data generation

To ensure a controlled computational environment and facilitate accurate comparisons, synthetic data was employed in this study. A 3D structure was meticulously designed to emulate the characteristics and properties of the actual particles encountered in Cryo-EM projection processes, and the formula we used to define the 3D structure is adjusted and improved on the basis of Anoshina's work[6], we apply some modification to the formula to make it more realistic. Subsequently, 2D projection images were generated from the constructed structure to serve as the basis for further modifications. The synthetic data is defined in cylindrical coordinates (r,θ,h) by the following formula:

First, we replace r by $\hat{r}$ with:

$$\hat{r} = r + 3|\sin(3\theta)| - 0.8; \tag{5.3}$$

$$r \in [0, 128], \quad \theta \in [0, 2\pi]$$

Next, we define the intensity function with:

$$I(\hat{r},\theta,h) = \begin{cases} f(\theta)\cdot g(\hat{r},h), & \text{if } \hat{r} < R \\ 0, & \text{otherwise} \end{cases} \tag{5.4}$$

where:

$$f(\theta) = \eta + 20|\sin(1.5\theta)| + 3$$

$$g(\hat{r},h) = \begin{cases} \frac{9\hat{r}^2}{R^2-h^2}, & h^2+9\hat{r}^2 \in [0,\frac{R}{3}) \\ 1, & h^2+9\hat{r}^2 \in [\frac{R}{3},R] \\ 0, & h^2+9\hat{r}^2 > R \end{cases} \tag{5.5}$$

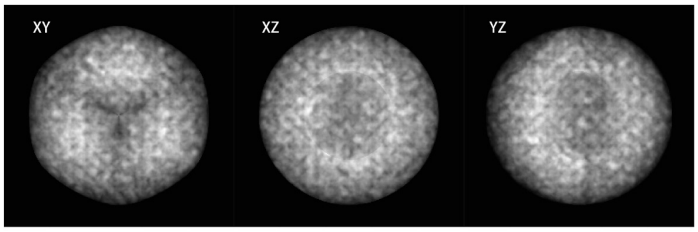


Figure 5.1: Phantom constructed

5.3.2 Projection images generation

Projection images were generated from the previously constructed 3D structure, incorporating 3D rotation to yield distinct projection angles, and a bi-linear interpolation method was employed to simulate in-plane rotations by $\frac{\pi}{16}$. Furthermore, the previously generated noise was incorporated into the projection images. As a result, a comprehensive set of 320 projection images with size 256×256 was successfully constructed (32 in-plane rotation angles, 10 repeats under each angle), effectively simulating the clustering problem basis before applying both Fourier Transform processes.

Given that the non-uniform discrete Fourier transform has represented the original image in a Polar format frequency representation, we are now poised to compute the power spectrum and ascertain paired in-plane rotation angles. In our case, we have 80 radial components and 256 angular components, thus for each image, the frequency information generated from the NUDFT has dimensions of 80×256.For the subsequent computational steps, we select the first four radial frequencies ($\frac{1}{Nr}$, $\frac{2}{Nr}$, $\frac{3}{Nr}$, $\frac{4}{Nr}$), amounting to approximately 5% of the entire radial frequency range, which is 4×256. These chosen power spectra exhibit robust energy content, contributing significantly to resultant intensities in contrast to noise elements, thus possessing elevated noise resistance.

Given the sole consideration of in-plane rotation and operating within the Polar format frequency domain, the in-plane rotation of the original image manifests as a displacement of periodic signals within the frequency domain. Consequently, the anticipation lies in observing distinctive displacements in the selected power spectra. In short, as we set angular frequency information in 256 components, the displacement of 1 unit now corresponding to $\frac{\pi}{256}$. As the power spectra of paired projection images establish a coherent displacement relationship of periodic signals. Subsequently, by identifying the lag discrepancy via maximum correlation, we derive the paired in-plane rotation angles. Furthermore, we uniformly weigh all shift translations of chosen low-frequency power spectra to determine the final in-plane rotation angles collectively. This approach facilitates the direct conversion of displacements into in-plane rotation angles.

The outcomes derived from the non-uniform discrete Fourier transform method exhibit a consistent pattern with those of the power spectrum pertaining to the selected frequencies. This consistent pattern resembles the translation properties of periodic signals showing their differences in angular direction frequencies.

5.3.3 2D alignment performance comparison result

In order to thoroughly assess and compare the performance of two Fourier transform processes and highlight their differences, we conducted a comprehensive comparison across various Signal-to-Noise Ratios ranging from 0.1 to 1, as well as 2, 5 and 10. (shown in figure 5.2) For each SNR level, we meticulously collected

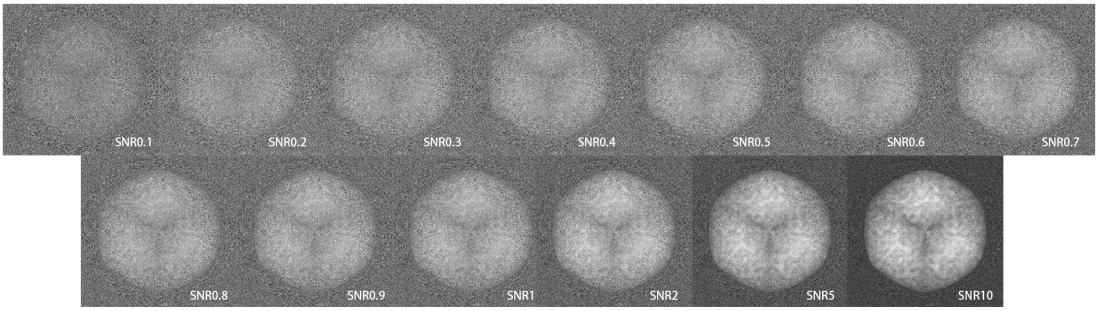


Figure 5.2: Different signal-to-noise ratios.

and processed a total of 320 projection images.

We simulate the noise environment by adding noise to the clean image generated prior. Since we already know the 3D structure, we can calculate the variance of the structure to create and add noise according to different SNR requirements. To be closer to reality, our SNR is defined based on variance:

$$SNR = \frac{Var(Signal)}{Var(Noise)} \tag{5.6}$$

In addition, unlike traditional noise that only uses Gaussian noise, different types of noise my appear during the imaging process, like auto-correlated noise[73, 38]. Thus, we use auto-correlated noise to simulate different type of distributed noise that may appear in the imaging process, and use this to test how the NUDFT algorithm handles this type of noise. In our case, we use AR (2) sequence to generate auto-correlated noise:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \tag{5.7}$$

Thus, we set the noise variance according to the required SNR and then add them to the clean projection image to finally generate noisy images with different SNR levels.

We then meticulously applied both algorithms (a result of NUDFT is shown in figure 5.3), ensuring that the entire computational process was carefully controlled, with minor variations introduced solely to emphasize the distinctions between the two Fourier transform methods.

After determining the in-plane rotation angle, we proceeded to revert all

projection images within the same class through bi-linear interpolation. Subsequently, we computed the class-average image using all the rotated projection images. The accuracy of the in-plane rotation angle estimation was assessed by the relative error rate:

$$\text{Relative Error} = \left| \frac{\text{Estimated in-plane rotation angle} - \text{True in-plane rotation angle}}{\text{True in-plane rotation angle}} \right| \tag{5.8}$$

As we have obtained the clear projection image while constructing images through synthetic data, we now can compute the correntropy between class-averaged image and clear projection image to examine the effectiveness of the estimated in-plane rotation angles and the performance of the previous computations:

$$\text{Correntropy with Gaussian Kernel} = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\|x_{i,j} - y_{i,j}\|^2}{2\sigma^2} \right) \tag{5.9}$$

Where σ is the size of the kernel.

As demonstrated in the results (shown in figure 5.4), the accuracy of in-plane rotation angle estimation is quite satisfactory. Notably, the algorithm employing the non-uniform discrete Fourier transform exhibits superior accuracy across various SNR levels compared to the traditional Fourier transform method. This enhancement can be attributed to its distinct application within the Fourier transform process.

In contrast, the algorithm based on conventional 2D Fourier transform yields more vulnerability, particularly in low SNR conditions, primarily due to its different computational approach.

It's worth noting that the performance of correntropy may also be influenced by the interpolation method used. However, the results from correntropy further underline the distinctions between these two methods. The algorithm employing the non-uniform discrete Fourier transform consistently demonstrates improved performance in class-averaged images and exhibits greater resilience to noise,

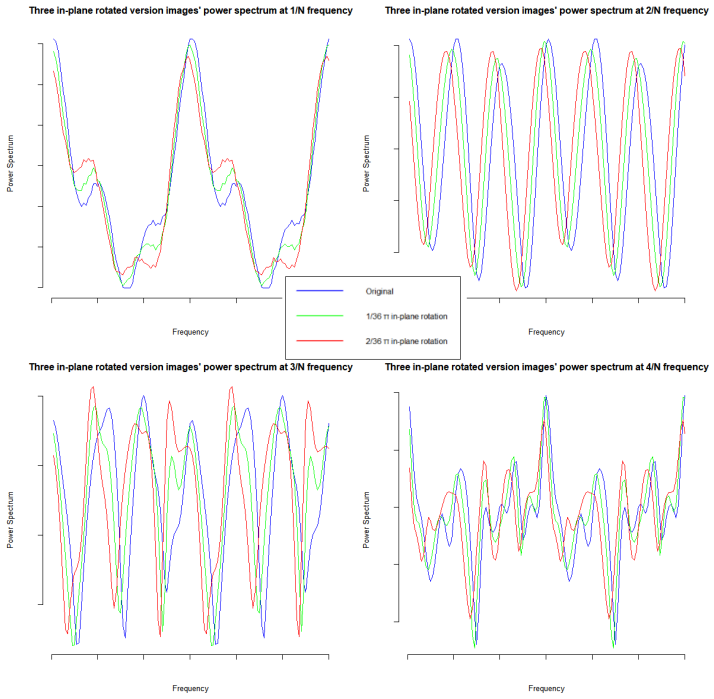


Figure 5.3: Power spectrum generated by the NUDFT.
The figure shows one case under SNR=0.1 : the translation on the angular direction of the frequency information is highly correlated with the in-plane rotations.

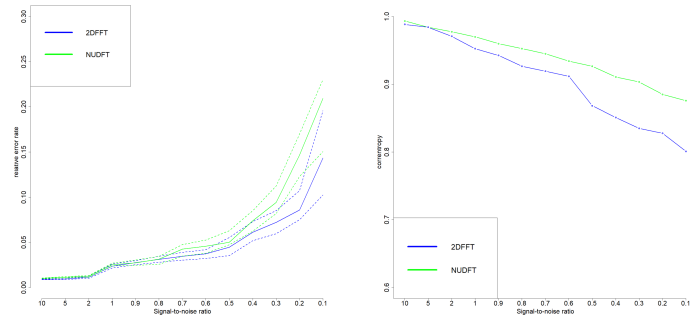


Figure 5.4: Left: relative error rate of two algorithms. Right: correntropy calculated from two algorithms

SNR	NUDFT	2DFFT
10	0.0091 (0.00872,0.00941)	0.0101 (0.00983,0.01047)
5	0.0095 (0.00863,0.01033)	0.0113 (0.01034,0.01203)
2	0.0113 (0.01021,0.01245)	0.0126 (0.01127,0.01316)
1	0.0237 (0.02151,0.02602)	0.0253 (0.02356,0.02683)
0.9	0.0274 (0.02518,0.02976)	0.0273 (0.02458,0.03038)
0.8	0.0312 (0.02795,0.03462)	0.0314 (0.02561,0.03443)
0.7	0.0345 (0.03023,0.03897)	0.0426 (0.03483,0.04759)
0.6	0.0374 (0.03211,0.04191)	0.0457 (0.03791,0.05265)
0.5	0.0446 (0.03534,0.05527)	0.0501 (0.04737,0.06278)
0.4	0.0612 (0.05178,0.07309)	0.0739 (0.06234,0.08505)
0.3	0.0721 (0.05958,0.08514)	0.0941 (0.08181,0.11239)
0.2	0.0857 (0.07512,0.10713)	0.1467 (0.12292,0.17011)
0.1	0.1433 (0.10248,0.19568)	0.2093 (0.15033,0.23035)

Table 5.1: Mean relative error rate of each method generated from equation 5.8, in the bracket is the 95% confidence interval from 100 simulations.

aligning with the conclusions drawn from the relative error rate analysis.

By employing a more suitable interpolation method, we can anticipate even more promising outcomes when rotating back the projection images after obtaining the in-plane rotation angle. This, in turn, should result in superior class-averaged images.

5.4 Summary and discussion

As discussed in the previous section, our proposed algorithm successfully use NUDFT to computed the in-plane rotation angle between single-particle Cryo-EM projection images, it translate the rotation transformation into angular frequency's displacement transformation by setting frequency domain in polar format, our result proved the effectiveness of the NUDFT. Also, the algorithm using the non-uniform discrete Fourier transform has better accuracy and noise resistance compared to the conventional 2D Fourier transform also results in better performance of produced class-averaged images. This is because 2DFFT requires interpolation when mapping the power spectrum information into a polar coordinate system to extract the rotation information. The unique property of the

non-uniform discrete Fourier transform allows it to extract the rotation information without any interpolation, this can avoid bringing more error into the analysis, which allows us to compute the in-plane rotation angle directly.

While our comparison is grounded in a limited set of projection images, the outcomes conspicuously underscore the disparities between the two Fourier transform methodologies. Our intent is to accentuate the importance of the non-uniform discrete Fourier transform and its potential in effectively tackling rotation problems in Cryo-EM. By rekindling interest in NUDFT, we can delve deeper into its capabilities and role in advancing Cryo-EM research.

Chapter 6

Spectral clustering featured NUDFT based 2D in-plane rotation angle estimation

6.1 Introduction

In Chapter 5 we develop a NUDFT based algorithm for estimation rotation between pairs of images. In this chapter, we consider improving its estimation accuracy for 2D images alignment within a class, which is a critical step in calculation of the class-average in single particle Cryo-EM. We propose an innovative method to align images in the same class. The proposed method begins with using spectral clustering to cluster the images to groups, followed by estimating rotation between images and the template image. Finally, we use the mode of all the rotation estimates in a group as the rotation estimate for all the images in the group. Unlike spectral clustering of images in Chapter 4, wherein clustering is based on a rotation-invariant distance metric between images, in this chapter we cluster images based on a distance metric between images, which is dependent on rotations between images as well. Also, we apply Z-score method to remove outlier images in each cluster obtained from the spectral clustering. Furthermore, based on the characteristics of single-particle Cryo-EM projection images, a novel distance metric which uses both power spectrum and amplitude generated

from the NUDFT was designed and used in clustering. We build an complete algorithm to include all these features, and compared its performance with the algorithm only use the NUDFT directly, the result proved the effectiveness of our algorithm.

6.2 Methodology

6.2.1 Introducing spectral clustering

Spectral Clustering is a clustering algorithm rooted in graph theory and linear algebra. It employs graph theory principles to map complex high-dimensional relationships by using eigenvectors of the Laplacian matrix, such as the internal structure of data in high-dimensional space. The subsequent display and processing occur in the low-dimensional space represented by these eigenvectors, enabling a more intuitive expression and processing of intricate data relationships. Mathematical operations akin to dimensionality reduction further contribute to computational efficiency. Subsequent clustering operations conceptualize the original high-dimensional data as nodes in an undirected graph, with the results of unsupervised clustering aligning precisely with optimal graph cut.

The implementation process involves obtaining the corresponding similarity matrix based on defined distances and calculating the corresponding adjacency matrix and degree matrix, then construct corresponding Laplacian matrix. Subsequently, eigendecomposition is performed on the Laplacian matrix to obtain eigenvectors. These eigenvectors construct the low-dimensional space for mapping, and the final step involves partitioning the optimal cut graph, completing the clustering operation. Despite being an unsupervised clustering algorithm, Spectral Clustering's ability to process complex data structures and internal relationships, coupled with its mapping to low-dimensional space, makes it particularly well-suited for addressing the in-plane rotation angle estimation in Cryo-EM's 2D projection images. Furthermore, the flexibility in setting paired similarities within the similarity matrix allows for versatile construction and description of complex relationships among input images.

When calculating paired rotation angles, usually the algorithm follows the

idea of selecting pairs of images with the closest distance or similarity. These calculations is very similar with the operational logic of spectral clustering, involving eigendecomposition of the Laplacian matrix representing mutual relationships between data. This mapping to low-dimensional space can effectively deconstructs complex relationships.

In subsequent algorithm implementation, the algorithm intuitively transforms the rotation relationships of images into two-dimensional transformation relationships, corresponding to their rotation angles. When facing a large number of input images for rotation angle calculation, the algorithm efficiently calculates paired rotation angles by pre-clustering data using the non-uniform discrete Fourier transform. It classifies processed images into classes with consistent or very similar rotation angles based on defined distances and then calculates paired in-plane rotation angles for each class by comparing projection images with selected templates. Further more, the frequency information generated from the NUDFT can be used in both clustering and the following angle estimation, enhancing both efficiency and accuracy.

In summary, the Spectral Clustering algorithm, with combination of NUDFT, proves effective in this quantitative analysis calculation problem. The mapping of data to low-dimensional space by Spectral Clustering aligns well with subsequent processing. This combination results in a logically consistent and harmonious approach, allowing the retained characteristics of each algorithm to contribute to an excellent overall result.

6.2.2 Introducing Z-score outlier detection

Z-score outlier detection is a commonly used statistical method for assessing the deviation of data points from the mean of a dataset. This method quantifies the extent of deviation for each data point by calculating its Z-score, which represents the number of standard deviations the data point is from the mean. Typically, data points exceeding a certain Z-score threshold are considered outliers. This method finds widespread application in quality control and anomaly detection, effectively identifying data points that deviate from the original distribution and providing valuable insights for further analysis and correction. By applying the

method of Z-score with the clustering result obtained from the spectral clustering, we now can further optimize and improve the clustering result. While our clusters were specified by each in-plane rotation angle, considering the limits of spectral clustering brought by the resolution of images and density of frequency domain distribution, there may be scenarios when some images were classified in the clusters while having slightly different rotation angles thus exceed the detection of the algorithm. Thus we apply Z-score outlier detection to each cluster to detect and exclude the images which bypass the detection of the algorithm. By sacrificing a small number of data we now can guarantee better accuracy in the computation of in-plane rotation angles, and this process also shows satisfactory improved result performance.

6.2.3 Algorithms

While the conventional algorithm chooses to compute the in-plane rotation angle directly, we choose to apply the clustering process prior to obtaining some prior information on the data to accelerate computing and improve overall accuracy. We combined spectral clustering with the non-uniform discrete Fourier transform, and merged them into a new algorithm, we also constructed another algorithm based on the conventional method to build a direct comparison between the two algorithms, and we expect to see an improvement in the overall performance of new algorithm.

First, we developed and applied our new algorithm. Prior to applying the Fourier transform, it is important to select a region of interest within the projection images. In general, the structure of molecules is concentrated in the centre of the image, while the four corners contain only noise information. Thus, we choose a circular region with a diameter consistent with the width of the square projection image.

After obtaining a circular mask of the original projection image, we apply the non-uniform discrete Fourier transform. By considering the centre of the circle as the origin of the polar coordinate system, we can set all frequency points of interest in a polar format, which facilitates further computation. During the Fourier transform, we represent the original information under the frequency

domain without adding or reducing any data. If the number of data points differs from the original input data, this is simply an interpolation artefact in the frequency domain. Therefore, we set the total number of data points in the frequency domain to be close to the original input data.

After applying the non-uniform discrete Fourier transform to the projection images, we obtain the power spectrum and the amplitude for subsequent computation. Although the frequency information obtained can be used to compute each image's corresponding paired rotation angle, due to the extreme noise factor in Cryo-EM's projection images, we choose to bypass this step and continue the preparation analysis. We define a novel formula to describe the distance and similarity between two paired input projection images by using their amplitude and power spectrum obtained from NUDFT. As power spectrum is typically used in these tasks, the reason we combine both amplitude and power spectrum is, in low SNR conditions, the mathematical property of power spectrum causes it to have less sensitivity to small differences of the main structure compared to large volume of noise factor, such as small changes in projection angles, this may lose the ability in capture and distinguish two images' rotation relationship or small differences caused by small in-plane rotation angles. While amplitude can preserve more information and differences under low SNR conditions, but it has less performance in describing large differences compared to the power spectrum. We will show this difference in later sections. Thus we take into consideration both power spectrum and amplitude in defining distance between images.

Considering the Fourier transform of the image among data set is $F(u, v)$, then the amplitude and power spectrum of the Fourier transform is:

$$Amplitude : A(F) = |F(u, v)| \tag{6.1}$$

$$Powerspectrum : P(F) = |F(u, v)|^2 \tag{6.2}$$

Then we can have our paired similarity S_{ij} between two image i and j with:

$$D_1(i, j) = \sum_{i,j=1}^{N} \sqrt{(A_i(F) - A_j(F))^2} \tag{6.3}$$

$$D_2(i,j) = \sum_{i,j=1}^{N} \sqrt{(P_i(F) - P_j(F))^2} \tag{6.4}$$

$$S(i,j) = \sum_{i,j=1}^{N} \frac{1}{2} \cdot \frac{D_1(i,j)}{\max(D_1)} + \frac{D_2(i,j)}{\max(D_2)} \tag{6.5}$$

After obtaining the similarity matrix followed by ideas in spectral clustering, we then defined the adjacency matrix by using a fully connected approach. Following we compute the degree matrix to finally get corresponding the Laplacian matrix. And then apply eigenvalue decomposition to the Laplacian matrix to obtain its corresponding eigenvectors.

Different from choosing the first N eigenvectors specified by eigenvalues in the clustering process, we take all eigenvectors into further computations to preserve as much information as possible, and this still can provide satisfy result. Finally, we apply K-means clustering to the Laplacian matrix's eigenvectors to get the clustering result of input projection images specified by each identical in-plane rotation projection angle.

After we had a rough cluster obtained from the input projection images, we assumed each cluster was specified by different identical in-plane rotation angles. However, there may be scenarios in which small changes were not detected by the spectral clustering, which exceeds the distinguish and identify the ability of the algorithm, but we have the cluster labels and cluster mean obtained from the spectral clustering. Then we apply Z-score density anomaly detection to each cluster, we seek to detect and delete the images with small differences in rotation angle but wrongly assigned to the cluster due to the high noise influence.

After we use the Z-score method to generate a more precise cluster, now we assume each cluster contains input projection images with the same in-plane rotation angles. Then we compute the in-plane rotation angle between each cluster's images by finding maximum correlation followed by amplitude and power spectrum obtained from NUDFT. As we can have an intuitive display and clear distance to describe the relationship between images, we can arrange projection clustered images specified by differences in rotation angles. As we have already computed each neighbour cluster's paired in-plane rotation angles, we now can

randomly select one cluster as a template and compute any images' paired in-plane rotation angle by adding up the links connected by their neighbours.

Algorithm 6: Non-uniform discrete Fourier transform based spectral clustering featured 2D in-plane rotation angle estimation algorithm

Input: Images I_i with random chosen image I_0 as template
Output: Rotation angle R_i between I_i and I_0

1 Select circle region R of the projection image with the diameter same as the image's width;
2 Apply non-uniform discrete Fourier transform on I_i to get F_i;
3 Compute the power spectrum and amplitude by $P_i = |F_i|^2$ and $A_i = |F_i|$;
4 Construct the similarity matrix by using the defined paired-similarity formula with P_i and A_i to get mS;
5 Construct the adjacency matrix by using fully-connected approach on mS to get mW;
6 Construct the corresponding degree matrix with mW to get mD;
7 Generate the corresponding Laplacian matrix mL;
8 Apply eigenvalue decomposition on mL to get eigenvectors FE;
9 Apply the normalized cut method to generate the pre-clustering C_i;
10 Apply the Z-score method within each class to get the accurate cluster result C_i';
11 Compute paired in-plane rotation angle for all images to get R_i;
12 Determine the final paired in-plane rotation angle by taking each class's mode to get R_i';

The whole pipeline is shown in figure 6.1. According to the above description, we can achieve the in-plane rotation angle estimation by this new combined algorithm. While this algorithm used spectral clustering to classify the input images specified by their different in-plane rotation angles, its main focus was to better describe the complex relationship between images under low SNR conditions. However, following this idea, the conventional method still can finish this task by directly using the peak correlation method to analyse the frequency information obtained from NUDFT. Thus we conduct another algorithm for the conventional method to build a direct comparison between these two methods. This algorithm has the same steps in obtaining the frequency information but only different in the following process, it directly computes the paired rotation angle by finding the peak correlation between the power spectrum obtained from the NUDFT.

By randomly selecting one input image as a template, we can generate all paired in-plane rotation angles. Thus it can achieve computing in-plane rotation angle between all input images by only using the frequency information obtained from NUDFT and does not need any other modification.

Algorithm 7: Non-uniform discrete Fourier transform based 2D in-plane rotation angle estimation algorithm

Input: Images I_i with random chosen image I_0 as template
Output: Rotation angle R_i between I_i and I_0

1. Select circle region R of the projection image with the diameter same as the image's width;
2. Apply non-uniform discrete Fourier transform on I_i to get F_i;
3. Compute the power spectrum by $P_i = |F_i|^2$ and $A_i = |F_i|$;
4. Determine the displacement of the power spectrum in angular direction between I_i and I_0 by taking the maximum correlation;
5. Compute paired in-plane rotation angle for all images between I_0 to get R_i;

6.3 Experimental result

6.3.1 Projection images generation

The synthetic data used here has been elaborated in detail in Chapter 4.3.1. (3D structure constructed is shown in figure 6.2), and projection images were generated from the previously constructed 3D structure, incorporating rotation to yield distinct in-plane rotation angles(shown in figure 6.4). We thoughtfully selected a projection image showing the square structure of the molecule, then a bi-linear interpolation method was employed to simulate in-plane rotations by $\frac{\pi}{100}$. And 5 extra in-plane rotation angles with $\frac{\pi}{200}$, $\frac{1\pi}{200}$, $\frac{21\pi}{200}$, $\frac{31\pi}{200}$, $\frac{41\pi}{200}$, $\frac{51\pi}{200}$. Furthermore, the previously generated noise was incorporated into the projection images. We constructed different signal-to-noise ratio scenarios from 10 to 0.1 (shown in figure 6.3), and we constructed 10 projection images for each different in-plane rotation angle.

We simulate the noise environment by adding noise to the clean image generated prior. Since we already know the 3D structure, we can calculate the variance

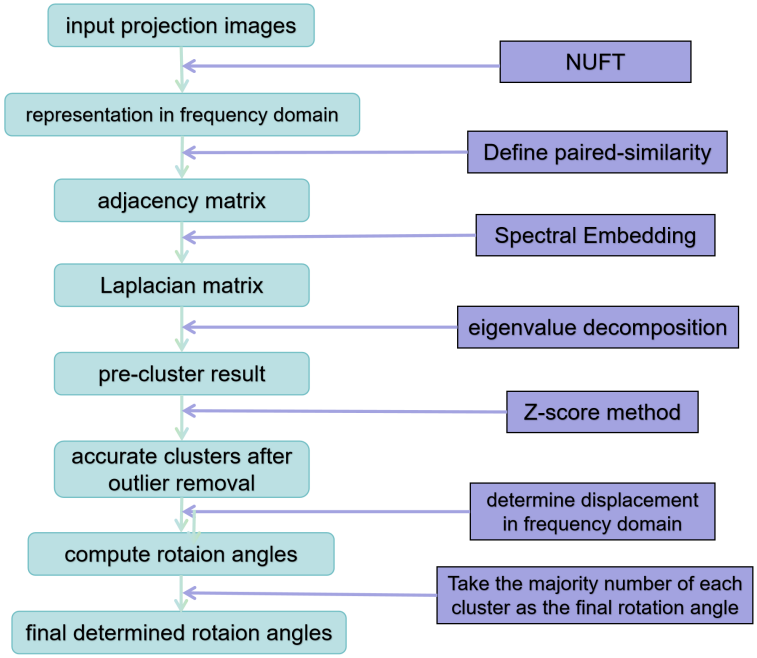


Figure 6.1: Proposed algorithm pipeline.

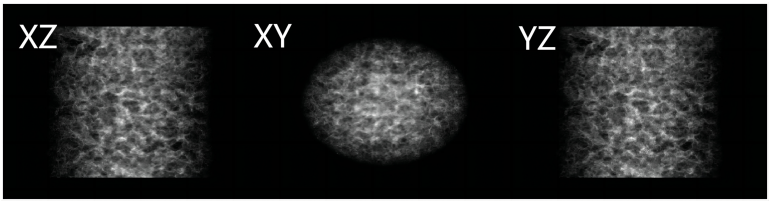


Figure 6.2: Phantom constructed

of the structure to create and add noise according to different SNR requirements. To be closer to reality, our SNR is defined based on variance:

$$SNR = \frac{Var(Signal)}{Var(Noise)} \tag{6.6}$$

In addition, unlike traditional noise that only uses Gaussian noise, different types of noise my appear during the imaging process, like auto-correlated noise[73, 38]. Thus, we use auto-correlated noise to simulate different type of distributed noise that may appear in the imaging process, and use this to test how the NUDFT algorithm handles this type of noise. In our case, we use AR (2) sequence to generate auto-correlated noise:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \tag{6.7}$$

Thus, we set the noise variance according to the required SNR and then add them to the clean projection image to finally generate noisy images with different SNR levels.

As a result, in each SNR condition, a comprehensive set of 1025 projection images with size 256×256 was successfully constructed, total 105 in-plane rotation angles, 10 repeats for those 100 regular images ($\frac{n\pi}{100}$) and 5 repeats for 5 extra angles ($\frac{n\pi}{200}$), effectively simulating the alignment problem basis before applying any further computation or clustering methods.

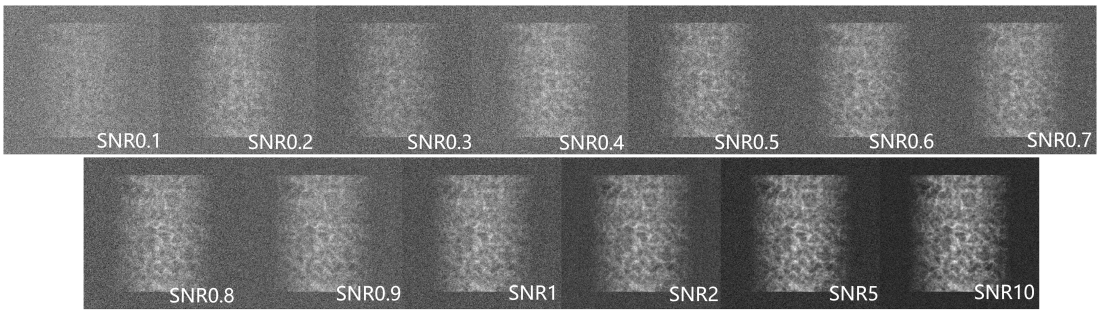


Figure 6.3: Different signal-to-noise ratios

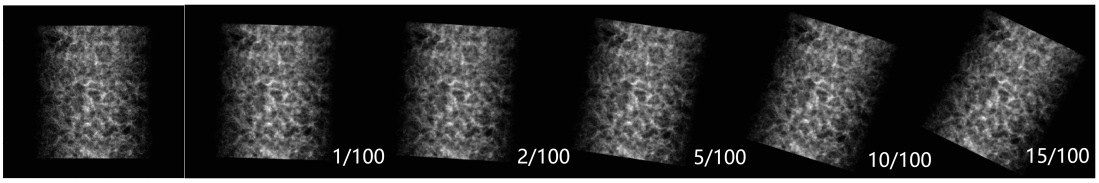


Figure 6.4: Samples of projection images.
A paired in-plane rotation relationship with the chosen template as shown in the plot, $\frac{\pi}{100}, \frac{2\pi}{100}, \frac{5\pi}{100}, \frac{10\pi}{100}, \frac{15\pi}{100}$

6.3.2 Spectral clustering application result

In our case, we set 80 radial components and 200 angular components in the NUDFT process, thus the frequency information entering the clustering process has dimensions of 80 × 400. And the corresponding similarity matrix (also can be called the distance matrix or the affinity matrix) has dimensions of 1025 × 1025, corresponding to 1025 images concluded in the computation. Then the corresponding adjacency matrix was constructed under fully connected method, thus the adjacency matrix is same as the previous affinity matrix (shown in equation (6.3), (6.4), (6.5) and figure 6.6), this can be written as:

$$W(i,j) = S(i,j) = \sum_{i,j=1}^{N} \frac{1}{2} \cdot \frac{D_1(i,j)}{\max(D_1)} + \frac{D_2(i,j)}{\max(D_2)} \tag{6.8}$$

Then, the corresponding degree matrix and the final Laplacian matrix also has dimensions of 1025 × 1025. We then apply eigen decomposition to the Laplacian matrix, the generated eigenvectors has dimensions of 1025 × 1025, and we used the first two eigen vectors (specified by the first two largest eigen values) to show the complex internal relationships of the input images, we then applied the K-means clustering method to all eigen vectors, this result is equivalent to the normalized cut method, thus we obtained the pre-clustering result, and we use first two eigen vectors to show a intuitive outcomes, mapping the in-plane rotation relationships into a two-dimensional space, allowing for intuitive visualization of the optimization results in subsequent processing (shown in figure 6.5).

After applying the spectral clustering, a pre-clustering result was generated, while the images have $\frac{\pi}{100}$ with the chosen template beyond the algorithm's pro-

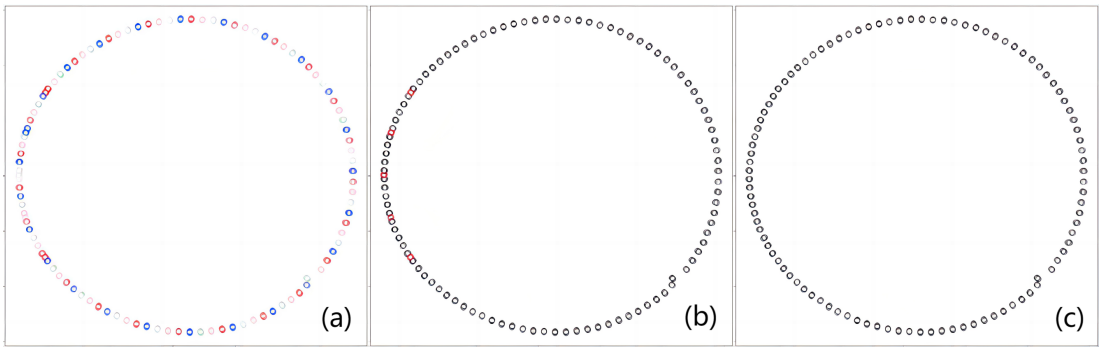


Figure 6.5: Result of spectral clustering and Z-score method. The figure shows one case under SNR=0.1 : (a)(b)(c) show the eigenvectors after spectral embedding, (a) shows the pre-clustering result, while each adjacent different colour represents different classes(we only used a few colours but there are 105 classes same as the in-plane rotations), and the red point in (b) shows the images with $\frac{\pi}{200}$ in-plane rotations, and (c) shows the final result after applying Z-score method, those images with $\frac{\pi}{200}$ in-plane rotations with deleted, showing the effectiveness of the algorithm.

cessing resolution, they will classified into clusters with other images with $\frac{\pi}{100}$ paired in-plane rotation angle. We then calculate the clustering accuracy of those $\frac{\pi}{100}$ sampled images because our main task is to compute those images' rotation angles and enhance the overall angle estimation accuracy of all images. Then, we considered the accuracy of this step by only considering the $\frac{\pi}{100}$ images:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Samples}}{\text{Total Number of Samples}} \tag{6.9}$$

According to the results, the difference between $\frac{\pi}{200}$ is extremely small, it has exceeded the resolution that the algorithm can handle under the influence of noise. However, for the majority of images represented by $\frac{\pi}{100}$, which is our main objective, the vast majority have been correctly classified. This result demonstrates an overall stable and satisfactory accuracy across different levels of signal-to-noise ratios (SNRs), confirming the suitability and robustness of spectral clustering for this task under defined novel paired similarity.

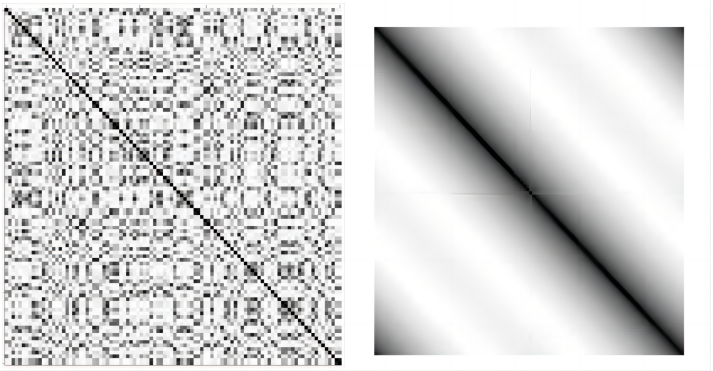


Figure 6.6: Part of the result of the similarity matrix.
The figure shows one case under SNR=0.1 : the left is the original distance matrix generated from the defined paired similarity, and the right image shows the sorted result after applying spectral clustering, which proves the effectiveness of the algorithm.

6.3.3 Final result after Z-score method

While the clustering result obtained from the prior spectral clustering application still contains small differences in images with a $\frac{\pi}{200}$ rotation angle, the Z-score method's objective is to reduce these influences as much as possible. The result accuracy computed now shows the exact accuracy of all images within clusters, this can explain the role of the Z-score method to detect and delete the wrong images within the clusters.

We apply Z-score method on each clusters' image classified from the previous clustering process, and we take images' first 10 radial components' power spectrum generated from the previous NUDFT in the computation, which is 10×400 for each image, corresponding to first 10 radial components and 400 angular components. In our case, while computing the mean value and the standard deviation, we regard the input power spectrum as a long vector, which is 1×4000 length vector, thus the generation of the Z-score is the mean value of the image's power spectrum under those 4000 positions, this can be written as:

$$Z_i = \frac{1}{n} \sum_{j=1}^{n} \frac{X_{i,j} - \mu_j}{\sigma_j} \tag{6.10}$$

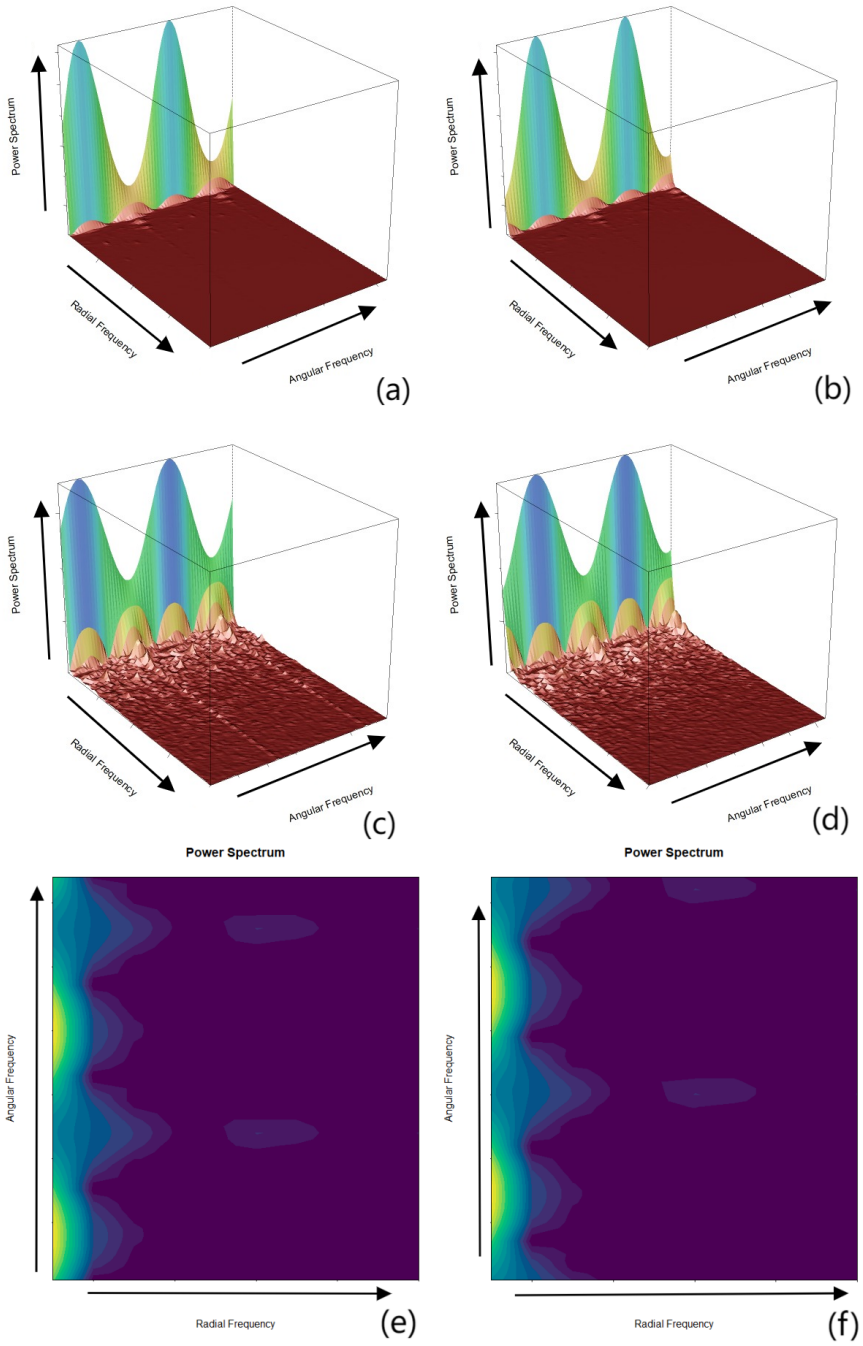


Figure 6.7: Frequency information generated from the NUDFT. The figure shows one case under SNR=0.1 : (a)(b) is the amplitude generated from the NUDFT, and (c)(d) is the power spectrum. (a)(c) and (b)(d) were from two images with an in-plane rotation relationship, and the translation on the angular direction between (a)(b) and (c)(d) showed the effectiveness of the NUDFT in detecting in-plane rotations.(e)(f) is the 2D version of the amplitude, this can also shows the translation relationship directly.

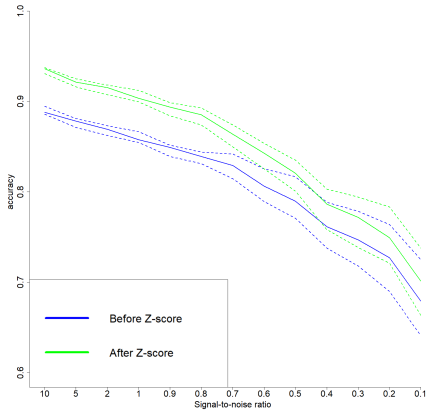


Figure 6.8: Z-score application result.

while,

$$\mu_j = \frac{1}{m}\sum_{i=1}^{m} X_{i,j} \tag{6.11}$$

$$\sigma_j = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(X_{i,j} - \mu_j)^2} \tag{6.12}$$

In our case, we input a combinations of each images' long vector (which can be regard as a matrix) to compute the Z-score, thus i is the image number within current group, j is the position of this image's power spectrum.

As the result shown in figure 6.8 and figure 6.5, the Z-score method well optimized previous cluster results, it successfully detected, distinguished and removed those images with $\frac{\pi}{200}$ in-plane rotation angle within each cluster while they were beyond the processing capability of spectral clustering. The final accuracy in the table 6.2 of this algorithm shows overall better performance than the conventional method. It shows better resistance to noise in low SNR conditions, while the conventional method has a great drop in performance as the SNR becomes extremely low. This result proves this new algorithm has better precision and robustness to noise as well.

SNR	Before Z-score	After Z-score
10	0.8879 (0.8859,0.8946)	0.9361 (0.9307,0.9376)
5	0.8783 (0.8712,0.8810)	0.9213 (0.9161,0.9249)
2	0.8691 (0.8624,0.8732)	0.9152 (0.9073,0.9180)
1	0.8577 (0.8546,0.8663)	0.9034 (0.8995,0.9122)
0.9	0.8491 (0.8390,0.8517)	0.8937 (0.8839,0.8985)
0.8	0.8392 (0.8312,0.8439)	0.8853 (0.8741,0.8927)
0.7	0.8292 (0.8146,0.8420)	0.8636 (0.8498,0.8741)
0.6	0.8063 (0.7893,0.8254)	0.8427 (0.8254,0.8537)
0.5	0.7897 (0.7707,0.8166)	0.8207 (0.8010,0.8351)
0.4	0.7615 (0.7376,0.7883)	0.7864 (0.7580,0.8029)
0.3	0.7469 (0.7180,0.7785)	0.7718 (0.7385,0.7941)
0.2	0.7271 (0.6898,0.7639)	0.7492 (0.7210,0.7834)
0.1	0.6792 (0.6410,0.7249)	0.7012 (0.6634,0.7376)

Table 6.1: Mean accuracy of the classification before and after Z-score application generated from 6.9, in the bracket is the 95% confidence interval from 100 simulations.

6.3.4 Performance comparison

The overall classification accuracy of the pre-clustering result generated from spectral clustering was satisfied, and the result also shows the application of the Z-score method has successfully optimized the previous cluster result, which means the Z-score method can detect and delete the images with $\frac{\pi}{200}$ in-plane rotation relationships while they were beyond the processing capability of spectral clustering and were classified into the wrong classes. Figure 6.7 clearly shows the changes in the NUDFT amplitude (and power spectrum) of images with different in-plane rotation angles only in the angular direction of the frequency domain, Which also show the effectiveness of NUDFT in generating rotation information of the image.

We used correntropy to measure the similarity between the computed final class-average and the clear projection template we used:

$$\text{Correntropy with Gaussian Kernel} = \sum_{i=1}^{M}\sum_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-\|x_{i,j}-y_{i,j}\|^2}{2\sigma^2}\right) \tag{6.13}$$

Where σ is the size of the kernel.

SNR	NUDFT-novel	NUDFT-direct
10	0.9872	0.9832
5	0.9761	0.9753
2	0.9632	0.9455
1	0.9555	0.9133
0.9	0.9251	0.8874
0.8	0.9044	0.8649
0.7	0.8899	0.8519
0.6	0.8735	0.8405
0.5	0.8563	0.8335
0.4	0.8435	0.8193
0.3	0.8289	0.7865
0.2	0.8008	0.7369
0.1	0.7821	0.7099

Table 6.2: Final correntropy of each algorithm

As shown in figure 6.9 and table 6.2, the result of the spectral clustering featured algorithm has an overall better performance than the algorithm that only uses displacement in the angular direction of frequency information to determine the in-plane rotation angles. And it shows better noise resistance in low signal-to-noise as well.

6.4 Conclusion and discussion

The overall performance of our proposed algorithm shows great capability in in-plane rotation angle estimation, in our designed task it proves to have an overall better performance than the conventional method. While the final determination of the paired in-plane rotation angle is based on the calculation of the displacement in the angular direction of the rotation information under the frequency domain of the image, we designed a complete algorithm to make it more accurate, and after combining the application of spectral clustering and Z-score method, the final accuracy was greatly improved. With the application of the Z-core method, unmanageable small angle differences can now remove most of their effects. Also, with the use of the novel paired-similarity we defined, spectral clus-

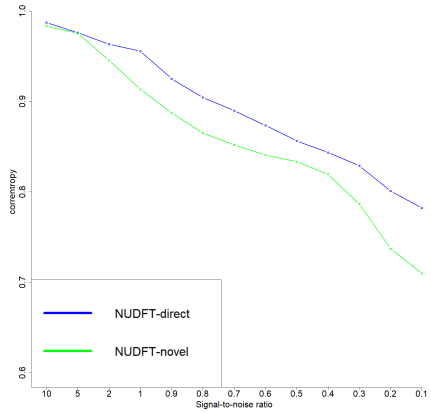


Figure 6.9: Correntropy obtained from two algorithms.
The blue line is the result correntropy obtained from the proposed algorithms under different SNRs, and the green line it the correntropy obtained from the NUDFT direct algorithm.

tering's process can provide an accurate and trusted basis for the final decision of in-plane rotation angle. This algorithm also demonstrates a close fit between the Laplacian matrix and its eigenvectors in spectral clustering and a direct description of the complex relationships between images in paired in-plane rotation angle estimation. This combination has enhanced the ability of the conventional method, and taking the mode number to determine the rotation among classes shows more robustness to noise. We expect this combination can provide another aspect in solving the in-plane rotation angle estimation problem and show the great fit of the Laplacian matrix.

Chapter 7

Discussion, conclusion and future work

In single-particle Cryo-EM data analysis, reconstructing a 3D particle structure from projection images is the most important and extremely challenging task due to the very low SNR of images. To improve the SNR, projection images are classified and aligned to generate the class average images. In this work, we consider development of image classification and alignment algorithms, which can be used to compute class averages in single-particle Cryo-EM data analysis.

First, we discussed and investigated the applicability of NUDFT in calculation of the FT of an image in polar coordinates, innovatively using NUDFT to represent a rotation of an image as translation of its power spectrum in polar coordinates. This enables us to extract rotation invariant features of an image by applying NUDFT to its power spectrum and construct NUDFT-based K-means algorithm. In comparison with K-means based extraction of rotation-invariant features by FFT of an image in polar coordinates, the proposed algorithm shows a good improvement in classification accuracy and robustness to noise, proving NUDFT's efficiency in building rotation-invariant classification models.

Second, we explored the possibility of further improve its efficiency in building 2D classification by combining NUDFT with spectral clustering using both power spectrum and amplitude features of images to calculate similarities between images. Our experiment results demonstrated the efficiency of the combination of

the spectral clustering and the NUDFT in image classification. Using both power spectrum and amplitude features to calculate similarities can improve classification performance. We constructed a NUDFT based spectral clustering algorithm and perform a comparison with the NUDFT-based K-means clustering, as shown by the comparison results, by leveraging spectral clustering and our novel distance metric, the proposed NUDFT based spectral clustering has better noise resistance and classification performance than NUDFT-based K-means.

Third, we examined the applicability of the NUDFT in the 2D image alignment (2D in-plane rotation angle estimation), we constructed an algorithm which leveraged NUDFT for eliminating the need for interpolation compared to classical FFT based rotation estimation. We performed simulation studies to compare the NUDFT based rotation estimation with the one based on FFT, the results demonstrated the efficiency of NUDFT in 2D image rotation alignment.

Last, we designed a complete 2D image alignment algorithm based on the NUDFT, by incorporating the spectral clustering and Z-score method as preprocessing in image rotation estimation. Pre-processing by the Z-score method and the spectral clustering method can provide a more reliable basis of applying NUDFT to image rotation estimation. Comparing with the NUDFT-based image rotation estimation without pre-processing, pre-processing can improve the estimation performance and have better noise resistance.

In single-particle Cryo-EM, image noise can be correlated. We tested all the proposed algorithms on simulated images with the correlated noise, we found that their capabilities of classification and alignment of projection images are not influenced by such type of noise.

To examine the potential of the proposed NUDFT algorithm in classifying real images from single particle Cryo-EM, We performed our classification method on projection images of the extracted particle images of the human facilitates chromatin transcription (FACT) in complex with partially assembled sub-nucleosomes (EMPIAR-10333) [53]. The results are promising and demonstrate the effectiveness of our NUDFT based classification algorithm. This real-world data study has been submitted to the Journal of Structure Biology, now under a second revision.

In this study we focus on applicability of NUDFT in image rotation-invariant

classification and rotation alignment. In single particle Cryo-EM, images from the same projection angle can be rotated and translated around each other. Rotation and translation invariant classification methods and rotation and translation alignment methods have been developed for calculating class-averages in single-particle Cryo-EM. The proposed NUDFT based algorithms solve the rotation problem by converting it to a translation problem. Hence, the proposed algorithms can be easily adapted to rotated and translated images.

For future exploration, we aim to investigate the inverse problem of the NUDFT for class-averaging in single-particle Cryo-EM. As the current class-averaging pipeline must determine in-plane rotation angles and align all images within a class before computing the class-average, we aim to generate the class-average of rotation-invariant features within a class and then apply the inverse NUDFT to obtain the class-average image. However, the inverse NUDFT's accuracy is very dependent on the selection of the kernel function. We aim to construct an algorithm to analyse the images to optimize the kernel function and then improve the accuracy of the inverse NUDFT, thus obtaining the class-average under NUDFT's frequency domain.

Bibliography

[1] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[2] David Agard, Yifan Cheng, Robert M. Glaeser, and Sriram Subramaniam. Chapter two - single-particle cryo-electron microscopy (cryo-em): Progress, challenges, and perspectives for further improvement. *Advances in Imaging and Electron Physics*, 185:113–137, 2014.

[3] Bruno Aiazzi, Stefano Baronti, and Massimo Selva. Image fusion through multiresolution oversampled decompositions. *Image Fusion: Algorithms and Applications*, pages 27–66, 2008.

[4] Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, James D Watson, et al. *Molecular biology of the cell*, volume 3. Garland New York, 1994.

[5] Nadezhda A. Anoshina, Andrey S. Krylov, and Dmitry V. Sorokin. Correlation-based 2d registration method for single particle cryo-em images. In *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6, 2017.

[6] Nadezhda A Anoshina, TB Sagindykov, and Dmitry V Sorokin. A method for generation of synthetic 2d and 3d cryo-em images. *Programming and Computer Software*, 44:240–247, 2018.

[7] Paul E Anuta. Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques. *IEEE transactions on Geoscience Electronics*, 8(4):353–368, 1970.

[8] Xiao-Chen Bai, Greg McMullan, and Sjors HW Scheres. How cryo-em is revolutionizing structural biology. *Trends in biochemical sciences*, 40(1):49–57, 2015.

[9] Lindsay A Baker and John L Rubinstein. Radiation damage in electron cryomicroscopy. In *Methods in enzymology*, volume 481, pages 371–388. Elsevier, 2010.

[10] William T Baxter, Robert A Grassucci, Haixiao Gao, and Joachim Frank. Determination of signal-to-noise ratios and spectral snrs in cryo-em low-dose imaging of molecules. *Journal of structural biology*, 166(2):126–132, 2009.

[11] Tamir Bendory, Alberto Bartesaghi, and Amit Singer. Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities. *IEEE signal processing magazine*, 37(2):58–76, 3 2020.

[12] Xu Benjin and Liu Ling. Developments, applications, and prospects of cryo-electron microscopy. *Protein Science*, 29(4):872–882, 2020.

[13] Lisa Borland and Marin Van Heel. Classification of image data in conjugate representation spaces. *JOSA A*, 7(4):601–610, 1990.

[14] Ronald Bracewell. The projection-slice theorem. In *Fourier Analysis and Imaging*, pages 493–504. Springer, 2003.

[15] E Oran Brigham. *The fast Fourier transform and its applications*. Prentice-Hall, Inc., 1988.

[16] Xuebin Chen, Chunmao Ye, Yong Wang, Yan Dai, and Qingrong Hu. Imaging algorithm for inverse synthetic aperture radar in condition of non-uniform data rate. *IEEE Transactions on Computational Imaging*, 9:396–408, 2023.

[17] Yu-Xuan Chen, Rui Xie, Yang Yang, Lin He, Dagan Feng, and Hong-Bin Shen. Fast cryo-em image alignment algorithm using power spectrum features. *Journal of Chemical Information and Modeling*, 61(9):4795–4806, 2021.

[18] Yuxiang Chen and Friedrich Förster. Iterative reconstruction of cryo-electron tomograms using nonuniform fast fourier transforms. *Journal of structural biology*, 185(3):309–316, 2014.

[19] Anchi Cheng and Yue Yu. Recent advances in data collection for cryo-em methods. *Current Opinion in Structural Biology*, 86:102795, 2024.

[20] Yifan Cheng. Single-particle cryo-em—how did it get here and where will it go. *Science*, 361(6405):876–880, 2018.

[21] Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[22] James G Colsher. Fully-three-dimensional positron emission tomography. *Physics in Medicine & Biology*, 25(1):103, 1980.

[23] Yao Cong, Julio A Kovacs, and Willy Wriggers. 2d fast rotational matching for image processing of biophysical data. *Journal of structural biology*, 144(1-2):51–60, 2003.

[24] Radostin Danev, Haruaki Yanagisawa, and Masahide Kikkawa. Cryo-electron microscopy methodology: current aspects and future directions. *Trends in biochemical sciences*, 44(10):837–848, 2019.

[25] Louis De Broglie. Waves and quanta. *Nature*, 112(2815):540–540, 1923.

[26] Borja Rodríguez de Francisco, Armel Bezault, Xiao-Ping Xu, Dorit Hanein, and Niels Volkmann. Mepsi: A tool for simulating tomograms of membrane-embedded proteins. *Journal of Structural Biology*, 214(4):107921, 2022.

[27] DJ De Rosier and Aaron Klug. Reconstruction of three dimensional structures from electron micrographs. *Nature*, 217(5124):130–134, 1968.

[28] Inderjit S Dhillon, Yuqiang Guan, and Brian Kulis. Kernel k-means: spectral clustering and normalized cuts. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 551–556, 2004.

[29] Ieva Drulyte, Rachel M Johnson, Emma L Hesketh, Daniel L Hurdiss, Charlotte A Scarff, Sebastian A Porav, Neil A Ranson, Stephen P Muench, and Rebecca F Thompson. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallographica Section D: Structural Biology*, 74(6):560–571, 2018.

[30] Jacques Dubochet, Marc Adrian, Jiin-Ju Chang, Jean-Claude Homo, Jean Lepault, Alasdair W McDowall, and Patrick Schultz. Cryo-electron microscopy of vitrified specimens. *Quarterly reviews of biophysics*, 21(2):129–228, 1988.

[31] Alok Dutt and Vladimir Rokhlin. Fast fourier transforms for nonequispaced data. *SIAM Journal on Scientific computing*, 14(6):1368–1393, 1993.

[32] Edward H Egelman. The current revolution in cryo-em. *Biophysical journal*, 110(5):1008–1012, 2016.

[33] Hwa-ping Feng. Early cryo-em work. *Nature Structural Biology*, 7(1):22–22, 2000.

[34] J. A. Fessler. Iterative tomographic image reconstruction using nonuniform fast fourier transforms. *Commun. Signal Process. Lab., Dept. Elect. Eng. Comput. Sci., Univ. Michigan, Ann Arbor, MI, Tech. Rep*, 2001.

[35] Jeffrey A Fessler and Bradley P Sutton. Nonuniform fast fourier transforms using min-max interpolation. *IEEE transactions on signal processing*, 51(2):560–574, 2003.

[36] Hassan Foroosh, Josiane B Zerubia, and Marc Berthod. Extension of phase correlation to subpixel registration. *IEEE transactions on image processing*, 11(3):188–200, 2002.

[37] Joachim Frank. Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nature protocols*, 12(2):209–212, 2017.

[38] Brent Fultz and James M Howe. *Transmission electron microscopy and diffractometry of materials*. Springer Science & Business Media, 2012.

[39] Zhi Geng, Zhun She, Qiang Zhou, Zheng Dong, Fei Zhan, Heng Zhang, Jianhua Xu, Zengqiang Gao, and Yuhui Dong. Nudim: A non-uniform fast fourier transform based dual-space constraint iterative reconstruction method in biological electron tomography. *Journal of Structural Biology*, 213(3):107770, 2021.

[40] Nathaniel R Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of mathematical statistics*, 34(1):152–177, 1963.

[41] Rajkiran Gottumukkal and Vijayan K Asari. An improved face recognition technique based on modular pca approach. *Pattern Recognition Letters*, 25(4):429–436, 2004.

[42] Michael Greenacre. *Correspondence analysis in practice*. chapman and hall/crc, 2017.

[43] Yonghua Han, Bin Ma, and Kaizhong Zhang. Spider: software for protein identification from sequence tags with de novo sequencing error. In *Proceedings. 2004 IEEE Computational Systems Bioinformatics Conference, 2004. CSB 2004.*, pages 206–215. IEEE, 2004.

[44] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series C (applied statistics)*, 28(1):100–108, 1979.

[45] G.T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer, Berlin, 2009.

[46] H Malcolm Hudson and Richard S Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging*, 13(4):601–609, 1994.

[47] Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh. Rotation-invariant multiresolution texture analysis using radon and wavelet transforms. *IEEE transactions on image processing*, 14(6):783–795, 2005.

[48] Laurent Joyeux and Pawel A Penczek. Efficiency of 2d alignment methods. *Ultramicroscopy*, 92(2):33–46, 2002.

[49] Charles D. Kuglin. The phase correlation image alignment method. 1975.

[50] Caterina Labrín and Francisco Urdinez. Principal component analysis. In *R for Political Data Science*, pages 375–393. Chapman and Hall/CRC, 2020.

[51] Kenneth Lange, Richard Carson, et al. Em reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr*, 8(2):306–16, 1984.

[52] Maofu Liao, Erhu Cao, David Julius, and Yifan Cheng. Structure of the trpv1 ion channel determined by electron cryo-microscopy. *Nature*, 504(7478):107–112, 12 2013.

[53] Yang Liu, Keda Zhou, Naifu Zhang, Hui Wei, Yong Zi Tan, Zhening Zhang, Bridget Carragher, Clinton S Potter, Sheena D'Arcy, and Karolin Luger. Fact caught in the act of manipulating the nucleosome. *Nature*, 577(7790):426–431, 2020.

[54] L. Lucchese and G. M. Cortelazzo. A noise-robust frequency domain technique for estimating planar roto-translations. *IEEE Transactions on Signal Processing*, 48(6):1769–1786, 2000.

[55] Steven J Ludtke, Philip R Baldwin, and Wah Chiu. Eman: semiautomated software for high-resolution single-particle reconstructions. *Journal of structural biology*, 128(1):82–97, 1999.

[56] Steven J Ludtke, Joanita Jakana, Jiu-Li Song, David T Chuang, and Wah Chiu. A 11.5 Å single particle reconstruction of groel using eman. *Journal of Molecular Biology*, 314(2):253–262, 2001.

[57] Arif Mahmood, Ajmal Mian, and Robyn Owens. Semi-supervised spectral clustering for image set classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 121–128, 2014.

[58] Duncan E McRee. *Practical protein crystallography*. Elsevier, 1999.

[59] Stefano Michieletto, Francesca Stival, and Enrico Pagello. A probabilistic approach to reconfigurable interactive manufacturing and coil winding for industry 4.0. In *Advances in Mathematics for Industry 4.0*, pages 61–93. Elsevier, 2021.

[60] Jacqueline LS Milne, Mario J Borgnia, Alberto Bartesaghi, Erin EH Tran, Lesley A Earl, David M Schauder, Jeffrey Lengyel, Jason Pierson, Ardan Patwardhan, and Sriram Subramaniam. Cryo-electron microscopy–a primer for the non-microscopist. *The FEBS journal*, 280(1):28–45, 2013.

[61] Jonathan Mitchell, J Beau W Webber, and John H Strange. Nuclear magnetic resonance cryoporometry. *Physics Reports*, 461(1):1–36, 2008.

[62] Azad Mohammed and Avin Abdullah. Scanning electron microscopy (sem): A review. In *Proceedings of the 2018 International Conference on Hydraulics and Pneumatics—HERVEX, Băile Govora, Romania*, volume 2018, pages 7–9, 2018.

[63] K Muneeswaran, L Ganesan, S Arumugam, and K Ruba Soundar. Texture classification with combined rotation and scale invariant wavelet features. *Pattern Recognition*, 38(10):1495–1506, 2005.

[64] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14, 2001.

[65] Eva Nogales. The development of cryo-em into a mainstream structural biology technique. *Nature methods*, 13(1):24–27, 2016.

[66] Melanie Ohi, Ying Li, Yifan Cheng, and Thomas Walz. Negative staining and image classification—powerful tools in modern electron microscopy. *Biological procedures online*, 6:23–34, 2004.

[67] Francisco PM Oliveira and Joao Manuel RS Tavares. Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93, 2014.

[68] Lori A Passmore and Christopher J Russo. Specimen preparation for high-resolution cryo-em. *Methods in enzymology*, 579:51–86, 2016.

[69] R Porter and N Canagarajah. Robust rotation-invariant texture classification: wavelet, gabor filter and gmrf based schemes. *IEE Proceedings-Vision, Image and Signal Processing*, 144(3):180–188, 1997.

[70] M Radermacher, T Wagenknecht, A Verschoor, and J Frank. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50s ribosomal subunit of escherichia coli. *Journal of microscopy*, 146(Pt 2):113–136, May 1987.

[71] Vittoria Raimondi and Alessandro Grinzato. A basic introduction to single particles cryo-electron microscopy. *AIMS Biophysics*, 9(1):1–, Jan 2022.

[72] B Srinivasa Reddy and Biswanath N Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996.

[73] Ludwig Reimer. *Transmission electron microscopy: physics of image formation and microanalysis*, volume 36. Springer, 2013.

[74] Michael Schatz and Marin Van Heel. Invariant classification of molecular views in electron micrographs. *Ultramicroscopy*, 32(3):255–264, 1990.

[75] Sjors HW Scheres. Semi-automated selection of cryo-em particles in relion-1.3. *Journal of structural biology*, 189(2):114–122, 2015.

[76] Sjors HW Scheres, Rafael Núñez-Ramírez, Carlos OS Sorzano, José María Carazo, and Roberto Marabini. Image processing for electron microscopy single-particle analysis using xmipp. *Nature protocols*, 3(6):977–990, 2008.

[77] Gerd Schneider, E Anderson, S Vogt, C Knöchel, D Weiss, M Legros, and C Larabell. Computed tomography of cryogenic cells. *Surface Review and Letters*, 9(01):177–183, 2002.

[78] Thomas Schultz and Gordon L Kindlmann. Open-box spectral clustering: applications to medical image analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2100–2108, 2013.

[79] Lawrence A Shepp and Yehuda Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE transactions on medical imaging*, 1(2):113–122, 1982.

[80] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.

[81] Amit Singer. Mathematics for cryo-electron microscopy. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 3995–4014. World Scientific, 2018.

[82] Georgios Skiniotis and Daniel R Southworth. Single-particle cryo-electron microscopy of macromolecular complexes. *Journal of Electron Microscopy*, 65(1):9–22, 2015.

[83] 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. *Journal of Structural Biology*, 171(2):197–206, 2010.

[84] Lei Sun, Xinzheng Zhang, Song Gao, Prashant Rao, Victor Padilla-Sanchez, Zhenguo Chen, Siyang Sun, Ye Xiang, Sriram Subramaniam, Venigalla Rao, and Michael Rossmann. Cryo-em structure of the bacteriophage t4 portal protein assembly at near-atomic resolution. *Nature communications*, 6:7548, 07 2015.

[85] Kenneth A Taylor and Robert M Glaeser. Electron diffraction of frozen, hydrated protein crystals. *Science*, 186(4168):1036–1037, 1974.

[86] Thomas C Terwilliger, Steven J Ludtke, Randy J Read, Paul D Adams, and Pavel V Afonine. Improvement of cryo-em maps by density modification. *Nature methods*, 17(9):923–927, 2020.

[87] Rebecca F Thompson, Matt Walker, C Alistair Siebert, Stephen P Muench, and Neil A Ranson. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, 100:3–15, 2016.

[88] Md Palash Uddin, Md Al Mamun, and Md Ali Hossain. Pca-based feature reduction for hyperspectral remote sensing image classification. *IETE Technical Review*, 38(4):377–396, 2021.

[89] M Van Heel. Classification of very large electron microscopical image data sets. *Optik (Stuttgart)*, 82(3):114–126, 1989.

[90] Marin van Heel. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy*, 13(1-2):165–183, 1984.

[91] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17:395–416, 2007.

[92] Lanhui Wang, Yoel Shkolnisky, and Amit Singer. A fourier-based approach for iterative 3d reconstruction from cryo-em images. *arXiv preprint arXiv:1307.5824*, 2013.

[93] X. Wang, Y. Lu, and X. Lin. Heterogeneous cryo-em projection image classification using a two-stage spectral clustering based on novel distance measures. *Briefings in Bioinformatics*, 23(3):bbac032, 2022.

[94] Xiangwen Wang, Yonggang Lu, and Jiaxuan Liu. A fast image alignment approach for 2d classification of cryo-em images using spectral clustering. *Current Issues in Molecular Biology*, 43(3):1652–1668, 2021.

[95] David B Williams, C Barry Carter, David B Williams, and C Barry Carter. *The transmission electron microscope*. Springer, 1996.

[96] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[97] Michael M Woolfson. *An introduction to X-ray crystallography*. Cambridge University Press, 1997.

[98] Mengyu Wu and Gabriel C Lander. Present and emerging methodologies in cryo-em single-particle analysis. *Biophysical journal*, 119(7):1281–1289, 2020.

[99] Songshan Xiao and Yongxing Wu. Rotation-invariant texture analysis using radon and fourier transforms. *Chinese Optics Letters*, 5(9):513–515, 2007.

[100] Bian Xiong, Qinghua Zhang, and Vincent Baltazart. On quadratic interpolation of image cross-correlation for subpixel motion extraction. *Sensors*, 22(3):1274, 2022.

[101] Zhengfan Yang and Pawel A Penczek. Cryo-em image alignment based on nonuniform fast fourier transform. *Ultramicroscopy*, 108(9):959–969, 2008.

[102] S. Yin, B. Zhang, Y. Yang, Y. Huang, and H.-B. Shen. Clustering enhancement of noisy cryo-electron microscopy single-particle images with a network structural similarity metric. *Journal of Chemical Information and Modeling*, 59(4):1658–1667, 2019.

[103] Kai Zhang, Y. Zhang, Zhongjun Hu, G. Ji, and Fei Sun. Development and frontier of electron microscopy 3d reconstruction. *ACTA BIOPHYSICA SINICA*, 26:209–15, 01 2010.

[104] Z Hong Zhou. Towards atomic resolution structural determination by single-particle cryo-electron microscopy. *Current opinion in structural biology*, 18(2):218–228, 2008.