Power-Focus,

Robust Aberration Correction in Electron Microscopy

Thesis

Rangoli Saxena

München 2024

Power-Focus,

Robust Aberration Correction in Electron Microscopy

Thesis

Rangoli Saxena

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Rangoli Saxena
aus Kanpur, Indien

München, den 27.02.2024

Erstgutachter: Prof. Dr. Winfried Denk

Zweitgutachter: PD Dr. Torsten Ensslin

Tag der mündlichen Prüfung: 09.04.2024

Contents

Abstract (English) **vii**

Abstract (German) **viii**

1 Introduction **1**

1.1 From a neural tissue to a connectome 5

1.2 Focusing in Scanning Electron Microscopes 7

1.2.1 From the electron beam source to an image in SEM 9

2 Power-Focus **11**

2.1 Introduction 11

2.2 Method 13

2.2.1 Power-Focus workflow 15

2.2.2 Closed-form Solution 18

2.2.3 Sample and Setup details 20

2.3 Results 21

2.3.1 Resilience against phase shifts 21

2.3.2 Advantage of adaptive filtering 22

2.3.3 Recovery limits 23

2.3.4 Convergence experiments with non-biological samples 24

2.3.5 Simulations 25

2.3.6 Calibrating Power-Focus 26

2.3.7 Field Curvature in MultiSEM 28

2.3.8 Time profiling 30

2.4 Discussion 30

3 Data validation and visualization 31

3.1 Introduction 31

3.2 Piezo Stage Calibration 33

3.2.1 Method 33

3.2.2 Result 38

3.3 Detecting gaps within the piezo scan 39

3.3.1 Method 39

3.3.2 Result 41

3.4 Montaging 42

3.4.1 Method 42

3.4.2 Result 44

3.5 Detecting gaps between piezo scans 46

3.5.1 Method 46

3.5.2 Result 49

3.6 Stack alignment 50

3.6.1 Method 51

3.6.2 Result 53

4 Conclusion and Outlook 55

A Tools Used 57

B Acronyms 59

C Supplementary 61

C.1 MTF Derviation 61

Acknowledgement 70

List of Figures

1.1 Connectomics time scales. 2
1.2 Acquisition setup 6
1.3 Wavefront aberrations 8
1.4 General parts of Zeiss, Gemini family, SEM 9

2.1 Ray aberration diagram 14
2.2 Method workflow 15
2.3 Closed-form solution 19
2.4 closed-form timing comparison 19
2.5 Error against Phase Shifts 21
2.6 Convergence from large defocus 22
2.7 Recoverable aberration range 23
2.8 Convergence experiments on multiple samples 24
2.9 Simulation experiments 26
2.10 Power-Focus calibration 27
2.11 Field Curvature MSEM 29
2.12 Time profiling 30

3.1 Imaging stages 32
3.2 Piezo positions reconstruction 34
3.3 Shift vector estimation using CC 35
3.4 Small overlap shift vector estimation 36
3.5 Piezo Stage calibration 38
3.6 Piezo scan gaps estimation 39
3.7 Piezo scan overlap 41
3.8 Piezo gap validation 42
3.9 Piezo scan frames 43

3.10 Piezo montage 44
3.11 Overview montage 45
3.12 FOV gap lookup table 47
3.13 FOV gap validation 49
3.14 stack pairs 51
3.15 Aligned stack 53

C.1 61

Abstract

Our project aims to develop a high-throughput neural volumetric scan acquisition machine capable of scanning brain tissue with an isotropic resolution of 10 nm^3 to enable analysis of neural pathways at synaptic resolution. We utilize a 91-beam Scanning Electron Microscope (SEM) to achieve the required resolution.

A key innovation of our project is the development of a novel and robust focusing routine for electron microscopy, *Power-Focus*, which recovers reliably from large aberrations. The experiments conducted with single-beam SEM demonstrated successful recovery from up to 120 μm in defocus and 75 μm in astigmatism. This method uses multiple images taken at known shifts around the aberration state and then uses the power spectrum of these images to deduce the aberration. The novel contribution of this work is the ability to circumvent the usage of the phase values of the input data by decoupling it from the final solution. As the phase values are insensitive to the system's aberrations, their information offered no gain in the final estimation, and by removing them, the method became resilient towards any potential misalignment between images. We introduced an adaptive filtering technique based on every frequency's signal-to-noise ratio, adjusting the filter with each iteration as the system approaches the correct focus. We finally formulated a closed-form solution for the aberration, previously based on curve optimization, by determining the curve's polynomial expression in terms of its first and second-order derivatives using Taylor expansion and then solving it to find the aberration at the maximum. The computational time is $\sim$270 ms for a 512×512 sized image pair. With parallelization on a multicore system, the processing time for 91 such beam pairs is 1.8 seconds, achieving a 14-fold speedup. We have also developed visualization tools to facilitate detailed examination of a small part of the dataset at high resolution or a large part at low resolution, along with a few validation techniques to ensure gap-free acquisitions. These are crucial for maintaining the continuity of scans as we acquire them in segments by performing stage movements.

Abstrakt

Unser Projekt zielt darauf ab, eine volumetrische Scanmaschine mit hohem Durchsatz zu entwickeln, die in der Lage ist, Hirngewebe mit einer isotropen Auflösung von 10 nm^3 zu scannen, um eine Analyse der neuronalen Bahnen mit synaptischer Auflösung zu ermöglichen. Wir verwenden ein 91-Strahl-Rasterelektronenmikroskop (SEM), um die erforderliche Auflösung zu erreichen.

Eine Schlüsselinnovation unseres Projekts ist die Entwicklung einer neuartigen und robusten Fokussierungsroutine für die Elektronenmikroskopie, *Power-Focus*, die sich zuverlässig von großen Aberrationen erholt. Die mit einem single-beam SEM durchgeführten Experimente haben gezeigt, dass eine Defokussierung von bis zu 120 μm und ein Astigmatismus von 75 μm erfolgreich behoben werden können. Bei dieser Methode werden mehrere Bilder verwendet, die bei bekannten Verschiebungen um den Aberrationszustand herum aufgenommen wurden, und dann wird das Leistungsspektrum dieser Bilder verwendet, um die Aberration abzuleiten. Der neuartige Beitrag dieser Arbeit besteht darin, dass die Verwendung der Phasenwerte der Eingabedaten umgangen werden kann, indem sie von der endgültigen Lösung abgekoppelt wird. Da die Phasenwerte unempfindlich gegenüber den Aberrationen des Systems sind, bieten ihre Informationen keinen Gewinn für die endgültige Schätzung, und durch ihre Entfernung wird die Methode unempfindlich gegenüber potenziellen Fehlausrichtungen zwischen Bildern. Wir haben eine adaptive Filtertechnik eingeführt, die die Frequenzfilterung auf der Grundlage des signal-to-noise ratio jeder Frequenz optimiert und den Filter mit jeder Iteration anpasst, wenn sich das System dem korrekten Fokus nähert. Schließlich formulierten wir eine geschlossene Lösung für die Aberration, die zuvor auf einer Kurvenoptimierung basierte, indem wir den Polynomausdruck der Kurve in Form ihrer Ableitungen erster und zweiter Ordnung mit Hilfe der Taylor-Erweiterung bestimmten und dann lösten, um die Aberration im Maximum zu finden. Die Rechenzeit beträgt $\sim$270 ms für ein 512×512 großes Bildpaar. Mit Parallelisierung auf einem Multicore-System beträgt die Verarbeitungszeit für 91 solcher

Strahlenpaare 1,8 Sekunden, was eine 14-fache Beschleunigung bedeutet. Wir haben auch Visualisierungstools entwickelt, um die detaillierte Untersuchung eines kleinen Teils des Datensatzes mit hoher Auflösung oder eines großen Teils mit niedriger Auflösung zu erleichtern, sowie einige Validierungstechniken, um lückenlose Aufnahmen zu gewährleisten. Diese sind entscheidend für die Aufrechterhaltung der Kontinuität der Scans, da wir sie in Segmenten durch Bühnenbewegungen erfassen.

Chapter 1

Introduction

The series of efforts initiated in the 1960s by Sydney Brenner, a South African-born biologist, with the aim to construct a wiring map of the nervous system of an animal successfully culminated after almost 2 decades with the publication of the wiring map of the brain of a nematode worm, *C. elegans* which consisted of just 302 neurons [1]. The result received global acclaim and adulation from others pursuing similar goals. The wiring map was extensively reproduced and analyzed - one of the early findings being the identification of specific neurons that were selectively responsive to particular odorants, showing that different sensory neurons have specialized roles in detecting distinct chemicals in the environment [2]. Insights such as these led to the emergence of a new field of scientific study which came to be known as *connectomics*, and the wiring map aka. wiring diagram aka. synaptic connectivity matrix became popularly known as *connectome*. Central to the field is the belief that the connectomes will serve as an essential tool for understanding animal behavior, analogous to how the genomes have been critical in facilitating and expediting research in genetics.

In recent years, projects have been underway to produce the connectomes for animals with larger brains - some of the popular picks being the *Drosophila melanogaster* (fruit fly), mouse, and zebra finch - a songbird [3–7] (for a more comprehensive list see Table 1 in the work by Kornfeld *et al.*, [8]). Several teams worldwide are working towards acquiring the connectome of the whole or a fraction of the nervous system of these animals at a resolution of 4-20 nm using volume electron microscopy (VEM) [9]. Recently, the connectome of a large fraction of the brain of the fruit fly, *Drosophila melanogaster*, was presented by the *FlyEM* team at HHMI Janelia in collaboration with Google Research. With around 25k neurons and 20 million synaptic connections between them [10], it boasted as the largest connectome produced until last year when the *FlyWire* team at Princeton University produced the connectome of the whole brain of an adult female fruit fly [11] by using the VEM data acquired by the team led by Davi Bock at HHMI Janelia [5]. Although currently available only as a preprint on bioRxiv, if it undergoes a successful peer review, this work will become the largest connectome mapped to date, encompassing 120k neurons and more than 30 million synaptic connections.

This rapid advancement in tackling large brain sizes can be primarily attributed to improvements in computing power and the implementation of advanced machine learning algorithms, which have significantly enhanced automation capabilities. Still, the *Curse of Time* weighs heavy on connectomics.

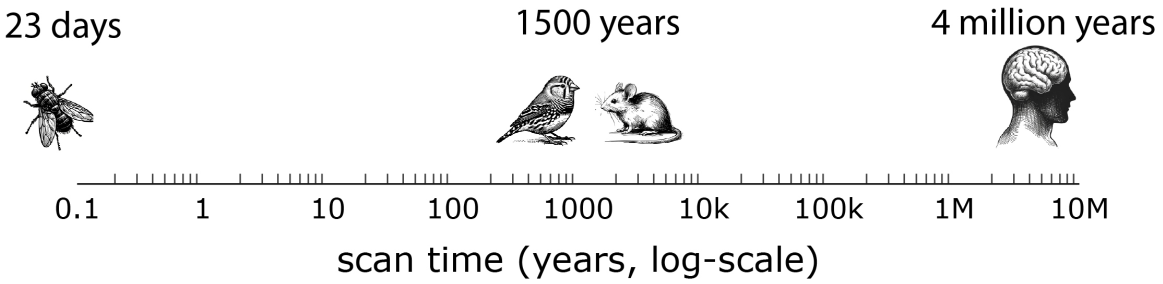


Figure 1.1: *Curse of time in connectomics.* The scan time required to image the whole brains of a fruit fly, a zebra finch, a mouse, and a human using a scan speed of 10 MHz, imaging at a cubic voxel resolution of side 10 nm. (The artistic images of the four animals were created using ChatGPT 4 [12])

Simple calculations show where we stand with the typical current capacities of a single-beam scanning electron microscope (SEM) to acquire whole brain scans at synaptic resolution. The time needed quickly mounts to thousands or even millions of years with the increasing brain volume (Fig 1.1). The pursuit of the connectomes of such large brain volumes has steered the efforts of the field toward developing methods that explore the parallelization of different stages of the connectome acquisition pipeline. The development of MultiSEM technology by Carl Zeiss GmbH, featuring SEMs with multiple electron beams - 61, 91, and an upcoming model with 331 beams [13, 14] - is one such innovation that was aimed at dramatically improving the imaging throughput. To truly utilize the high-throughput multi-beam imaging, however, needs a closely coupled high-throughput tissue-sample milling process as well. The team at HHMI Janelia, [15] developed a wide-area tissue-sample milling method compatible with the MultiSEM - the Gas Cluster Ion

Beam (GCIB). It used ionized argon clusters hitting a rotating sample at a glancing angle of 30° to achieve uniform depth removal within 10 *nm*, giving the voxel an isotropic dimension of 10 *nm*.

The principal project of our lab is centered on utilizing a 91-beam MultiSEM, in combination with a GCIB milling method, to set up a maximally automated high-throughput acquisition machine, targeting volumetric scales significantly beyond the entire brain of a fruit fly. The team is currently working towards acquiring the individual brain nuclei of a zebra finch, with the aspirational undercurrent of wanting to map the whole brain of a mouse or a zebra finch eventually. Over the course of setting up the equipment, a multitude of interesting challenges emerged, one among them being the routine of *focusing* the images which became a major bottleneck, and removing this bottleneck became the primary part of my doctoral work.

Acquiring a large volumetric scan involves prolonged continuous operations - spanning weeks or months, preferably with little human intervention. So, consistently high resolution can only be maintained if the imaging system stays focused and recovers quickly from any aberration abruptly appearing over the period of acquisition process. Furthermore, the current high-throughput acquisition machine is markedly more sophisticated than its single-chambered predecessors [16, 17], (a schematic of our setup: Fig. 1.2) necessitating frequent movements of the sample as part of the acquisition. This complexity demands a focusing routine capable of recovering from aberrations resulting from the uncertainties inherent in these movements, which can sometimes be substantial.

Our initial attempts at automated focusing involved implementing and testing an aberration estimation technique, developed by Paxman *et al.* [18], on our single-beam SEMs. This technique was based on the work done by Robert Gonsalves, RW Gerchberg, and WO Saxton in the 1970s [19, 20], modified derivations of which are used in astronomy even today [21, 22]. This technique was proven to work on SEM data taken for brain samples in the study done by Jonas Binding and Winfried Denk in 2013 [23]. Testing with our imaging parameters and expected window of aberrations on the single-beam SEMs, we found that recovering from relatively large aberrations would often fail, and manual intervention would be usually needed.

In pursuit of expanding this narrow recovery range, one crucial observation was the irrelevance of the object's phase values in determining the transfer function of our optical system. Decoupling the phase and power information in the derived probability equations (revisited in detail in the next chapter) became the core of our research efforts. As a successful consequence of these efforts, the first chapter of this thesis presents a novel focusing method, named *Power-Focus*, capable of consistently correcting large aberrations, with the robustness originating from circumventing the use of phase values of the input data and utilising an adaptive frequency filter based on the signal-to-noise ratio (*SNR*) of each frequency.

The experiments on single-beam SEMs, performed at 10 nm pixel size show that in the absence of astigmatism, defocus as high as $\pm 150\ \mu m$ converges to below 0.1 μm within 5-8 iterations. The range decreases by $\sim$20% when astigmatism is introduced in the system but still surpasses the previously attempted methods. The recoverable area with defocus and astigmatism is (120, 75) μm. The normed aberrations fall below 1 μm within 10±5 iterations. We find that it takes $\sim$270 ms to run the computations for a single pair of frames of size 512×512. In case of MultiSEM with 91 beam pairs, the same computation takes $\sim 1.8s$ by utilising parallelization on a multi-core machine (machine configuration details in Ch. 2) with a speedup of $\sim 14\times$.

The second chapter presents the imaging-related tasks within the pipeline, the necessity of which emerged alongside the progression of the acquisition project. A significant enhancement to the MultiSEM in our project was the addition of a new physical stage, developed in-house by Winfried Denk and Juergen Tritthardt, to the MultiSEM system's in-built mechanical stage. This newly added stage, in principle, should contribute upto a 2x increase in acquisition speed, primarily due to its better stabilization characteristics. I contributed towards calibration of this stage which proved surprisingly challenging, one reason being the observed non-linearities in the stage control parameters. Further, we developed tools to streamline the validation of the completeness of the acquisitions. The total sample area in our setup can reach up to 25 mm^2 and is covered in sections, achieved by a combination of movements between both stages. The precision of these movements can vary depending on the distances moved or the specifics of the stage hardware. Therefore, adequate overlap, between the different segments of data is essential for a gap-free full surface scan. The tools also montage raw data in various resolutions to visually inspect the scans quickly. Moreover, the work towards aligning the surface scans from different layers into a volumetric scan allowed for immediate verification of the milling progress during the acquisition process, as evidenced by observable changes in the neural data with each milling cycle.

1.1 From a neural tissue to a connectome

The complete workflow of creating a connectome from a neural tissue involves several stages, each at a different level of developmental maturity at present. Several teams are simultaneously working on addressing the challenges of each of these stages.

The initial step of the pipeline is to prepare the brain sample, which involves expert laboratory work. I did not work with these methods in my doctoral project; however, a high level description of a typical sample preparation process should be helpful for the reader. The first step in sample preparation is *fixation* of the neural tissue to preserve the cellular structure, which arrests their activity and prevents dehydration. The next step is *staining* since the biological tissue has a low natural contrast for EM where the differential in electron scattering and/or secondary electron emissions provide the contrast. To enhance the contrast, the tissue is impregnated with a heavy metal, a popular pick being osmium. The membranes inside tissue get laden with the heavy nuclei of the metal, which increases the scattering of the incoming electron beam. The final step is *embedding*, usually with a resin. When the tissue is infiltrated with it, a supportive matrix is formed after polymerization, which preserves the structural integrity of the sample and hardens it, which ends up feeling like a rigid plastic block to the touch. Researchers in this field frequently note that the quality of tissue embedding significantly influences the ease of tissue sectioning. The sample is then trimmed to reveal the target area, and sections with a thickness ranging from 50 nanometers to 1 micron are precisely cut using an ultramicrotome fitted with a diamond knife. Methods to collect sections with minimal loss are also an ongoing area of research. In our lab, the choice of surface on which the sections are collected are 2-inch silicon wafers, doped with boron and with a low resistivity of 0.005 Ω/cm. The wafers are exceptionally flat, ensuring the focus variation across the surface is not abrupt. They also have sufficient thermal stability to tolerate the local heating due to the electron beam.

The next step in the pipeline is the raw data acquisition cycle which is an alternating cycle of imaging and milling of the sample surface. In our lab, a 91-beam Carl Zeiss Scanning Electron Microscope (SEM) is used in combination with a GCIB for milling to achieve a high-throughput scan acquisition system. The setup is depicted in Fig 1.2. A typical MultiSEM acquisition with 91 beam frames, referred to as the MultiSEM hexagon, is also shown. Inside the milling chamber, the milling stage is positioned to create a glancing angle of 30° with the ion beam. A gate valve, installed between the two chambers, controls the isolation of the two chambers. A robotic mechanism is also used to transfer the wafer between these chambers. My work in the lab was focused on the development of the necessary methods and instrument control software for the imaging process. Working on the hardware components or the milling process was not part of my doctoral research.

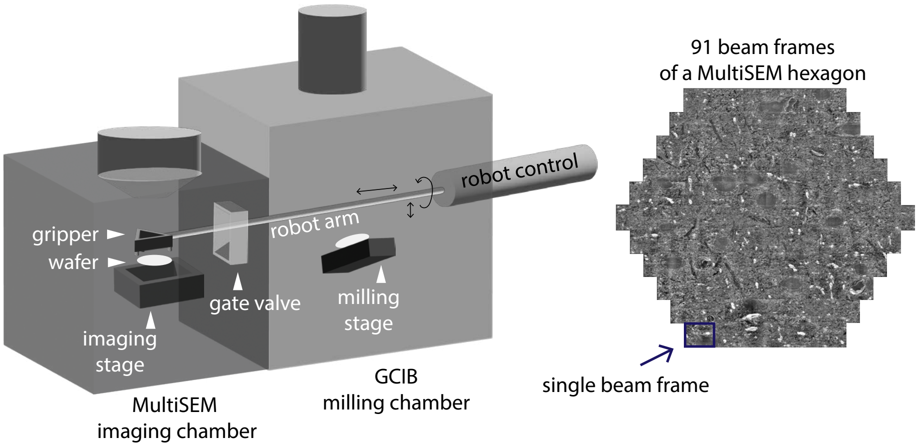


Figure 1.2: *Left:* the acquisition setup. *Right:* single acquisition from the MultiSEM showing 91 beam frames - a single frame shown in the blue rectangle.

At the beginning of an acquisition cycle, two wafers with prepared samples on them are manually placed inside the system - one each on top of the imaging and the milling stage. Then the surfaces of the samples on the wafer inside the imaging chamber are imaged with MultiSEM before the robot swaps the positions of the two wafers. After the switch, the robot retracts, and the gate valve closes isolating the two chambers. After two chambers have been isolated, the imaging of surfaces of the samples on the second wafer starts. Simultaneously, the first wafer undergoes milling inside the milling chamber, removing a tissue layer of 5 - 10 nm thickness to expose a deeper layer of all samples, which novel surfaces are then imaged after the next deisolation-exchange-isolation cycle. The acquisition cycle ends with all the samples on both of the wafers completely milled down to the respective wafer surface. The raw VEM data of the complete neural tissue at a voxel size of 10 nm^3 is then available in computer storage.

The next step in the pipeline is the processing of raw data into an aligned volumetric scan of the neural tissue. Our lab is capable of validation and rough alignment for small subsets of this dataset. However, given its massive scale, which can extend to several petabytes, we collaborate with the connectomics team led by Viren Jain at Google Research, [24]. They employ advanced alignment techniques, notably the Adversarial Image Alignment method based on Convolutional Neural Networks [25] for aligning the scans.

The next step following the alignment is the segmentation of the neural structures. A leading method for segmentation is the flood-filling networks [26]. Segmentation is followed by thorough proofreading. Sophisticated visualization tools like Knossos [27] or Neuroglancer [28] allow researchers to navigate and inspect the segmented volumetric scan across all three planes. Synapses are identified based on their characteristic features, which pave the way for mapping the synaptic connections which then leads to a fully constructed connectome of the neural tissue.

1.2 Focusing in Scanning Electron Microscopes

In SEMs with Field-emission guns (FEG), the electron beam is largely coherent, and the lens aberration in such an optical system is described using wave optical theory as the coherence results in interference upon interacting with the sample, which can amplify or diminish the electron intensity at the area of contact ([29, Ch. 2, p. 34-38]). Ideal focusing occurs when a converging lens operates on the incoming wavefront, adds a phase shift, and transforms it into a spherical wavefront that converges to a single point. Assuming geometric optics, the point spread function (PSF) for a spherical wavefront is a delta function. In the presence of aberrations, the shape of the wavefront deviates from its ideal spherical form, broadening the cross-section of the PSF. This aberration can be interpreted as a superimposing corrupting set of phase shifts (or the aberration wavefront), and breaking it down into a polynomial expression of ρ and ϕ (polar coordinates) simplifies its inversion process. Similar to light optics, a compound of correcting lenses is employed in SEMs, where the different lenses address individual terms in this function. A mathematical framework used often in lens design is the Zernike polynomials. They are a set of orthogonal polynomials that form a complete basis set over the unit circle [30], of which the phase aberrations for defocus and astigmatism are shown in Fig. 1.3.

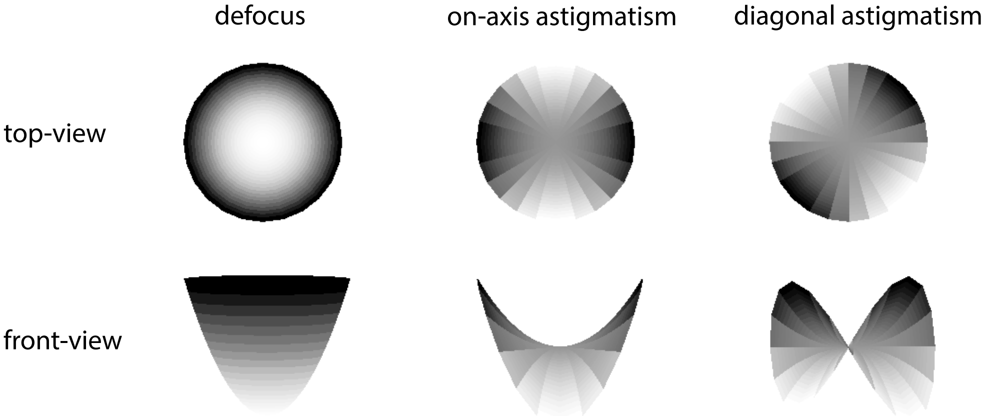


Figure 1.3: The phase shifts added to the spherical wavefront in the case of defocus and astigmatism as seen from along the optical axis (top-view) and normal to the optical axis (front-view).

In modern SEMs, such as the ones offered by Zeiss, aberrations such as spherical, coma, and distortion are effectively minimized to meet the resolution requirements of our project. This achievement is the result of precise engineering and the foundational theoretical contributions made by researchers, with Otto Scherzer, whose theoretical insights laid the groundwork for correcting spherical aberration in magnetic electron lenses, being particularly noteworthy [31]. Defocus and astigmatism, however, are still fickle and can change abruptly to the detriment of the image sharpness. These aberrations can be manually corrected by changing the strength of the current in the stigmator and the objective lens using control dials on the microscope. To automate the process of correction, Power-Focus routine was developed as a script which can be used from programs running on the control computers for either the single-beam SEM or the MultiSEM. This script enables the control software to estimate and correct defocus and astigmatism across a broad, recoverable area, correcting for normed aberrations up to $\sim$140 μm.

The concept of treating aberrations as phase shifts added to the spherical wavefront is further explored in Section 2.2, which delves into the principle of *phase-diversity* (PD). Power-Focus, a technique encompassed within PD methods for focusing, employs known phase shifts to the aberration and captures multiple images as measurements [19]. The premise is that acquiring at least two such images provides enough data to estimate the system's initial aberration state.

1.2.1 From the electron beam source to an image in SEM

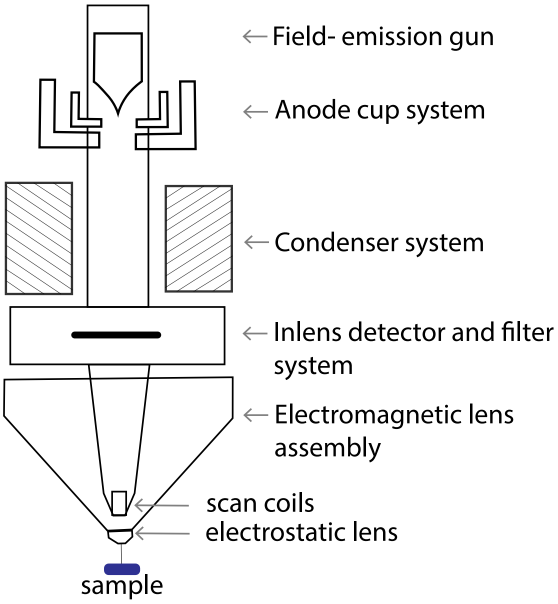


Figure 1.4: General parts of Zeiss, Gemini family, SEM. [32].

The optical principles underlying the MultiSEM are built on the same foundational concepts as those of a traditional single-beam SEM, as illustrated in the basic schematic in Fig. 1.4. A distinctive feature of the MultiSEM is its ability to divide a single beam, generated - in case of Zeiss MultiSEM - by a Schottky Field-Emission Gun (FEG), into multiple beams using a micro aperture array [33]. The beam then traverses a condenser lens system that demagnifies it, typically by a factor of ~5000 [29, 34].

The electromagnetic and electrostatic lens system manipulates the electrons' trajectory, bringing the beam's cross-section to within 10 nm. Similar to the functioning of a confocal light microscope, in an SEM, the electron probe is finely focused on the sample surface and then raster scanned across the field of view by energizing two sets of electromagnetic scan coils in x and y.

The interaction between the electron beam and the sample generates both back-scattered and secondary electrons, with the ratio between these emissions largely dependent on the beam landing energy, the mean atomic number of the sample interaction area, and the angle at which the beam strikes the sample. Low-voltage imaging (with beam landing energy of less than 10 keV) is advantageous for biological samples as it lessens radiation damage and reduces charging effects. So the Zeiss machine which is specifically designed for biological samples by operating within a voltage range of 1.2-1.5 keV features in-lens secondary electron detectors since the signal from secondary electron emissions is stronger in this range of landing energy.

In general the configuration of the detector system varies; it can either be integrated within the optical column or installed externally. In this study, the microscopes are equipped with an in-lens detector system located inside the chamber, designed for secondary electron detection [32].

Chapter 2

Power-Focus

2.1 Introduction

Typically, manual focusing involves adjusting the focal plane incrementally in a specific direction and visually inspecting whether the image becomes sharper or more blurred. So unsurprisingly, a part of focusing efforts branched out in favor of search-based algorithms using a sharpness measure computed in real space [35–39], or frequency space [40]. These approaches notably fall short in addressing astigmatism, with only a few attempting to correct it. Among these attempts, some need up to 50 recordings [41] or rely on theoretical frameworks lacking empirical validation [42]. Furthermore, the burden of deciding multiple operational parameters often falls on the operator. A reason for this failure to address astigmatism adequately could be the substantial expansion of the search space when introducing astigmatism in the system and the sharpness curve becoming multimodal. So, the research trend seems to have pivoted towards leveraging deep learning models to manage this increased complexity. However, this approach also brings challenges, such as data collecting and training periods spanning several days. Furthermore, they report that they fail when there are significant changes in imaging parameters, with attempts at quick recalibration using the same model resulting in diminished performance [43, 44].

So it appears that the focusing efforts in SEM have not effectively leveraged insights from its peers in astronomy, who have managed to exploit the knowledge of the imaging process itself very cleverly [45, 46]. Since the imaging process is fairly well understood, we can predict how the image will change with the perturbation. The sensitivity of the frequency values is tied to its aberration value through the *PSF*, and the ratio of these frequency values for multiple measurements taken at two known distances from the current aberration is also unique for the current aberration (ignoring noise). This technique of recording multiple measurements with known phase shifts to infer the aberration state became known as *phase diversity*.

It was initially introduced in the 1970s through the Gerchberg-Saxton algorithm [19, 20], a variation of which is still being used in the focusing routine of the James Webb Space Telescope ([47]). Around the same time, a similar algorithm was developed by Robert Gonsalves, who also introduced a likelihood function in his work. This was then generalized for an arbitrary number of phase diversities by Paxman *et al.* [18], who also laid the theoretical foundation for systems with translation-invariant *PSF* based on Gaussian and Poisson noise models using a maximum likelihood approach. Its focusing and stigmating capability on an SEM was reported by Binding *et al.* [23], where they referred to the method as Maximum-A-Posteriori Focusing and Stigmation (MAPFoSt), and this is where we began our investigation.

We found that MAPFoSt lacked the robustness needed to fully automate the focusing routine in SEM because recovery from large aberrations (beyond 20 μm for 10 nm pixel size) often failed. As the expected window of the aberrations in our machine was much broader, the advancement of this narrow recovery range was our primary area of investigation, starting at the limits where the method began failing.

The initial key observation was the irrelevance of the spectral phase values of the input measurement in the aberration estimation. MAPFoSt is a frequency-based method, initiating with the Fourier transformation of the multiple measurements, hereafter referred to as phase diverse images. It uses both the phase and the magnitude which makes it sensitive to shifts between the phase diverse images, and aligning them within a single pixel becomes more difficult with increasing aberrations and/or when the object becomes sparser. We successfully decoupled the spectral phase from the power in the course of deriving the final likelihood equation. This removed the alignment constraint. Furthermore, we derived a closed-form solution for the aberration by performing a Taylor expansion of the newly formulated likelihood function around the aberration state $(0, 0, 0)$ up to the second order and then solving for the aberration which maximized the approximated curve. Addressing the effect of noise in the system, several investigators independently reported a strong bias in estimation as the *SNR* is lowered [48–50] and this is what we observed as well. Filtering the frequencies still stands as the simplest and the most effective technique for reducing this bias. Advancing further on this path, we have found that filtering out the frequencies according to their individual *SNR* is superior to a low-pass filter set at a constant cutoff. This adapts the number of frequencies to the remaining aberration as it reduces in successive iterations and checks the retention of high frequencies with high SNR, making the method robust without compromising precision. The experiments on single-beam EMs, performed at 10 nm pixel size, show defocus as high as ± 150 μm converges to below 0.1 μm within 5-10 iterations. The performance drops slightly when astigmatism is introduced in the system the recoverable volume with defocus and astigmatism is (120, 50, 50) μm. The normed aberrations fall below 1 μm within 10±5 iterations. We find that the computation for a single frame of size 512×512 takes $\sim 270ms$ and parallelizing on a multi-core machine in case of 91 beams brings the total time to ∼1.8 s, with a speedup of $\sim 14\times$ (machine configuration in section 2.3.8). We further demonstrate it on SEM using inorganic samples.

2.2 Method

The PSF is usually parametrized by following the cross-section of the electron probe spot as the optical system spans the aberration space. The boundary of this spot is defined by the propagation of the marginal rays at the exit pupil. The ray aberration diagram is shown in Fig. 2.1a. Assuming geometrical optics, if the astigmatism coefficients are negligible, then the spread of the PSF is a circular disk of radius $(\rho/f) * \sqrt{\delta^2}$, where ρ is the pupil radius, f is the focal length, and δ is the defocus as shown as the cross-sections given in Fig 2.1b.

When astigmatism is present in the system, the focal points of the ray fans are not at a single point but are spread out over a distance, η, along the optical path. The focal points at the extremities of this distance belong to a pair of ray fans normal to each and at an arbitrary angle, ϕ, from the scan direction x and y as shown in Fig 2.1a and 2.1c. The PSF is no longer strictly circular in such a system as the imaging plane travels along the optical path. It is a line at the two extreme focal points, a circular disk at the midpoint, and an ellipse in between. The two parameters, (η, ϕ) can be used to parametrize astigmatism. However, the stigmator system is generally a pair of quadrupole electromagnets placed concentric to the optical path and at an angular distance of 45°. By varying the ratio of the current in the electromagnets, the ϕ space could be spanned entirely, without having to rotate a single quadrupole electromagnet mechanically. The normed strength of current in the excitation coil of the two quadrupole magnets control the astigmatic difference, η, in physical space. The mathematical relation between (η, ϕ) and (α, β) has been explained in the Supplementary section C.1. The scaling from μm to microscope units is done by the astigmatism scaling calibration parameter, γ, and has been explained in section 2.3.6.

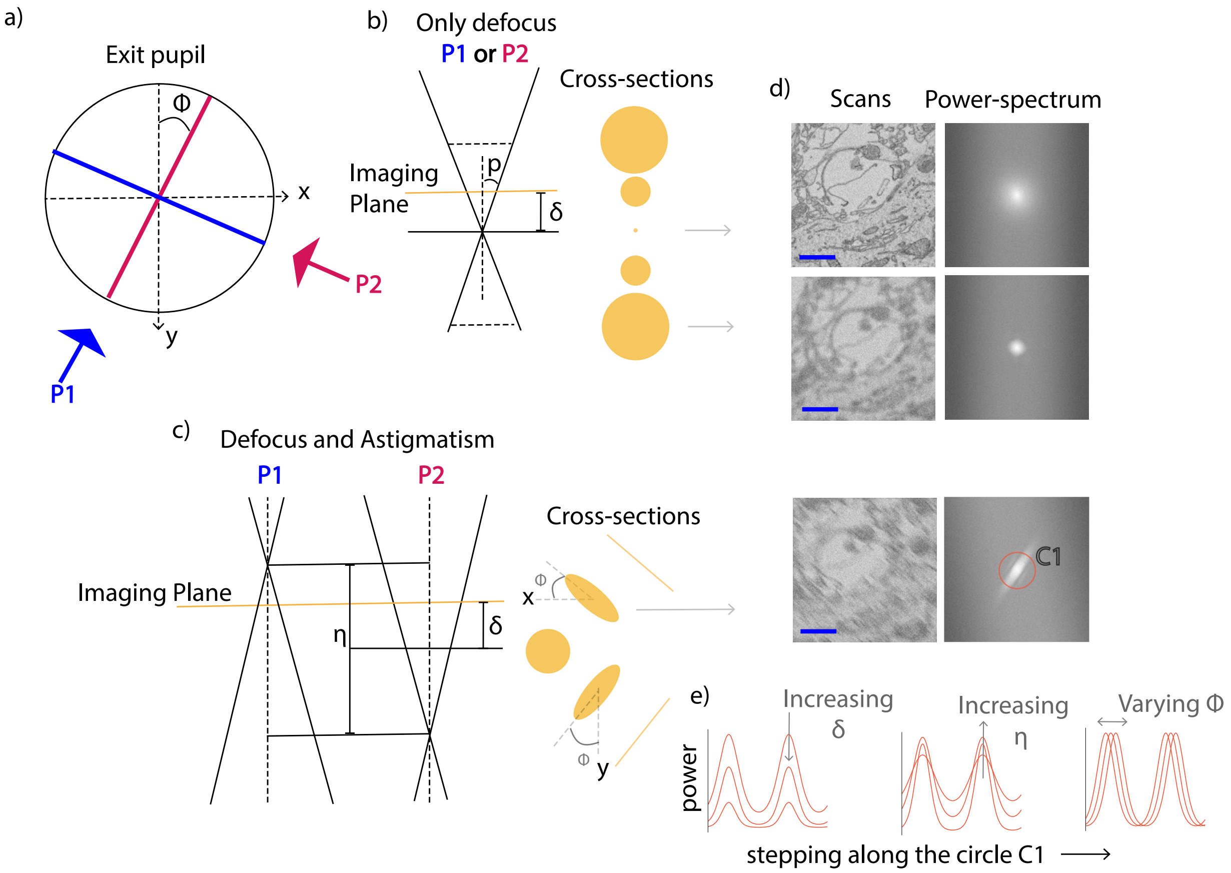

a)
Exit pupil
Φ
x
y
P1
P2
b)
Only defocus
P1 or P2
Cross-sections
Imaging
Plane
p
δ
d)
Scans
Power-spectrum
c)
Defocus and Astigmatism
P1
P2
Cross-sections
Imaging Plane
η
δ
Φ
x
Φ
y
C1
e)
Increasing
δ
Increasing
η
Varying Φ
power
stepping along the circle C1

Figure 2.1: *Ray aberration diagram*: a) Top view: Orthogonal ray fans (in blue and red), at a variable angle from the image scan direction (x, y) as seen from the exit pupil, along the optical path. b) Side view: In the absence of astigmatism, the pair of ray fans share the same focal point. The cross-section of the beam is circular and its diameter varies linearly with the distance between the image plane and focal plane, and is a point when $\delta = 0$. c) In the presence of astigmatism, the shape of the cross-section is no longer constrained to a circle. The two side-views, one from the direction P1 and the other from the direction P2 show the focal points separated by a distance of η. The cross-section of the beam is a line at the two endpoints, a circle at the midpoint, and an ellipse in between. The elliptical tilt varies with ϕ. d) The scans in real space are shown for a focused, defocused, and defocused + astigmatic system (top to bottom). The corresponding zero-centered power spectrums are shown on the right. e) The three figures plot the spectral power (y-axis) against the steps taken along the circle, drawn in red, in the center of the third power spectrum and its variation with increasing δ, η, and ϕ is shown. The scale bars (in blue) are 1 μm.

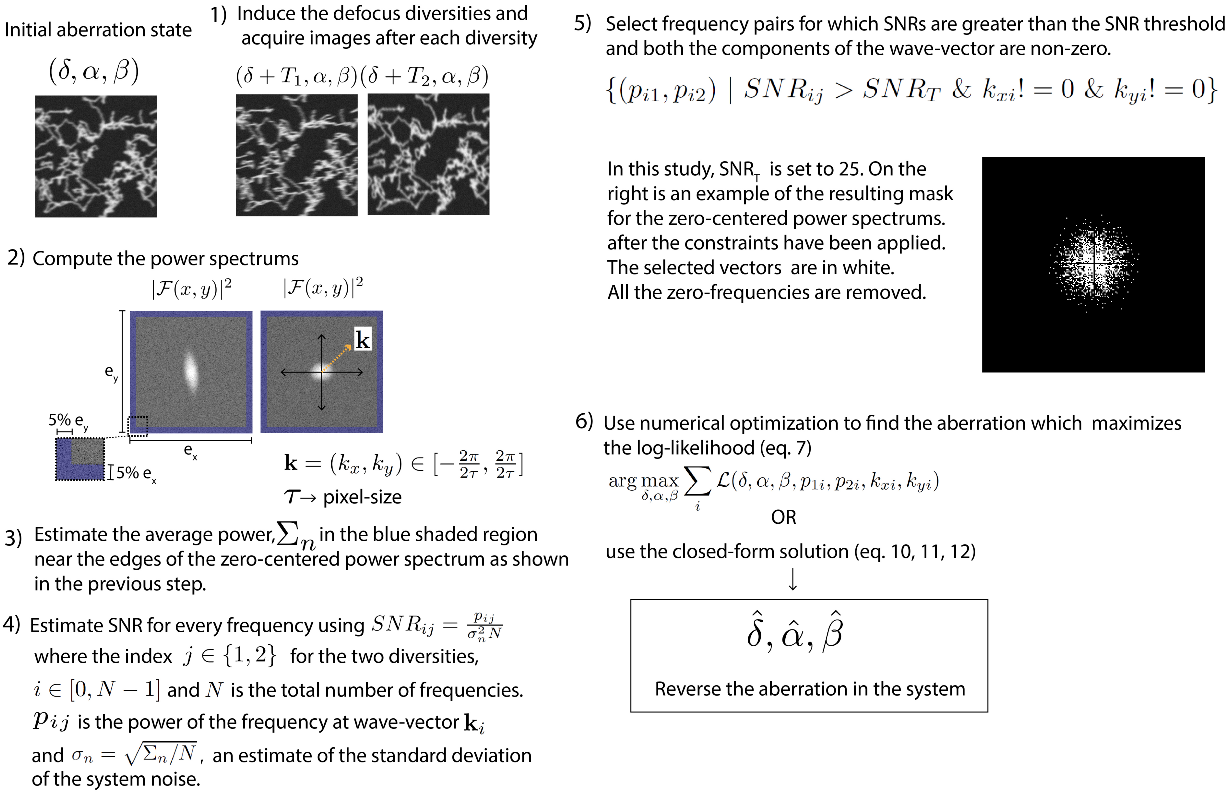
The *MTF*, being the Fourier transformation of the PSF, is a Bessel function of the first kind J_1. The MTF derivation has been given in the Supplementary section C.1. This study uses a Gaussian approximation of this function as proposed by Binding [23]. The numerical aperture, na, is equal to $sin(p)$ in vacuum, where p is the electron probe convergence angle (see Fig 2.1b). It has been approximated as $tan(p)$, for very small values. Estimating the na of the system has been explained in the section 2.3.6.

$$MTF(\delta, \alpha, \beta, k_x, k_y, na) = \exp(-\frac{1}{8} * na^2((k_x^2 + k_y^2)(\delta^2 + \alpha^2 + \beta^2) \\ +2\delta\alpha(k_x^2 - k_y^2) - 4\delta\beta k_x k_y)))$$

Here, k_x, k_y are the two wave vectors.

2.2.1 Power-Focus workflow

With the MTF formula at our disposal, we can proceed to a typical application of the Power-Focus method.



Initial aberration state

(δ, α, β)

1) Induce the defocus diversities and acquire images after each diversity

$(\delta + T_1, \alpha, \beta)(\delta + T_2, \alpha, \beta)$

2) Compute the power spectrums

$|\mathcal{F}(x,y)|^2$ $|\mathcal{F}(x,y)|^2$

$\mathbf{k} = (k_x, k_y) \in [-\frac{2\pi}{2\tau}, \frac{2\pi}{2\tau}]$

$\tau \rightarrow$ pixel-size

3) Estimate the average power, $\sum_n$ in the blue shaded region near the edges of the zero-centered power spectrum as shown in the previous step.

4) Estimate SNR for every frequency using $SNR_{ij} = \frac{p_{ij}}{\sigma_n^2 N}$ where the index $j \in \{1, 2\}$ for the two diversities, $i \in [0, N-1]$ and N is the total number of frequencies. p_{ij} is the power of the frequency at wave-vector $\mathbf{k}_i$ and $\sigma_n = \sqrt{\Sigma_n / N}$, an estimate of the standard deviation of the system noise.

5) Select frequency pairs for which SNRs are greater than the SNR threshold and both the components of the wave-vector are non-zero.

$\{(p_{i1}, p_{i2}) \mid SNR_{ij} > SNR_T \ \& \ k_{xi}! = 0 \ \& \ k_{yi}! = 0\}$

In this study, SNR_T is set to 25. On the right is an example of the resulting mask for the zero-centered power spectrums. after the constraints have been applied. The selected vectors are in white. All the zero-frequencies are removed.

6) Use numerical optimization to find the aberration which maximizes the log-likelihood (eq. 7)

$\arg\max_{\delta,\alpha,\beta} \sum_i \mathcal{L}(\delta, \alpha, \beta, p_{1i}, p_{2i}, k_{xi}, k_{yi})$

OR

use the closed-form solution (eq. 10, 11, 12)

$\downarrow$

$\hat{\delta}, \hat{\alpha}, \hat{\beta}$

Reverse the aberration in the system

Figure 2.2: *Method workflow*

The workflow is shown in Fig. 2.2. The aberration state (δ, α, β) of the system is shifted by two known defocus values $T1, T2$ independently, and images are acquired after each shift. The images are Fourier transformed and their power spectrums are computed. The signal-to-noise ratio (*SNR*) for every frequency is estimated by dividing the powers by total noise power $N\sigma_n^2$, where N is the number of frequencies and σ_n is the standard deviation of the system noise. The noise in the signal in SEM is statistical and follows a Poisson distribution but can be approximated to a Gaussian distribution with a sufficiently high average number of secondary electrons collected per pixel. Some experimental evidence suggests that a reasonable limit of this value is 3.2 electrons per pixel, as presented in Hayes *et al.* [51]. If the σ_n of the system is not known, an estimation of it can be made by plugging the mean power at the four edges (Σ_n) of the zero-centered power spectrum in $\sigma_n = \sqrt{\Sigma_n/N}$. The frequencies lower than the *SNR* threshold are filtered out, and the selected pairs are fed into the likelihood function 2.9. The aberration vector, which maximizes the likelihood function, is then determined using numerical optimization methods. Alternatively, the closed-form solution can be used (eq. 2.13-2.15, sub-section 2.2.2).

Derivation of the likelihood function

We assume that the optical properties of the SEM can be described by a linear *MTF* because the *PSF* is invariant under translation [52]. The image formation can be written as,

$$i(k_x, k_y) = o(k_x, k_y)MTF(\delta, \alpha, \beta, k_x, k_y,) + n(k_x, k_y) \tag{2.1}$$

Here $i(k_x, k_y)$ is the image frequency, and $o(k_x, k_y)$ is the object frequency for the wave-vector (k_x, k_y). The aberration state is (δ, α, β) for defocus, astigmatism x, and astigmatism y, and $n(k_x, k_y)$ is the noise term. The diversity is added by adding t_1, t_2 to the defocus, and the two power spectrums $p_j(k_x, k_y), j \in (1, 2)$ are then computed (eq. 2.2).

$$p_j(k_x, k_y) = |o * MTF(\delta + t_j, \alpha, \beta, k_x, k_y) + n_j(k_x + k_y)|^2, \ j \in (1, 2) \tag{2.2}$$

Power pairs for which $SNR >> 1$, the second-degree noise term in the expansion can be ignored, and the equation 2.2, can then be approximated to,

$$p_j = (o_I m_j)^2 + 2n_{Ij}o_I m_j + (o_R m_j)^2 + 2n_{Rj}o_R m_j \tag{2.3}$$

where, $o = io_I + o_R, n = in_I + n_R$, $o, n \in \mathbb{C}$,
$m_j = MTF(\delta + t_j)$
The probability distribution function (PDF) for p_j, given δ, α, β,and o is

$$f_{P_j}(p_j|o, \delta, \alpha, \beta) = \frac{\exp \frac{-(m_j^2(o_I^2+o_R^2)-p_j)^2}{8m_j^2(o_I^2+o_R^2)\sigma^2}}{\sqrt{8\pi(m_j^2(o_I^2+o_R^2)\sigma^2)}} \tag{2.4}$$

where $\sigma = \frac{\sigma_n}{\sqrt{2}}$, where σ_n is the standard deviation of the Gaussian noise model in real space.

The joint probability can then be written as

$$f_P(p_1, p_2|o, \delta, \alpha, \beta) = \prod_{j=1}^{2} \frac{\exp - \frac{(m_j^2 o_P - p_j)^2}{8m_j^2 o_P \sigma^2}}{\sqrt{8\pi(m_j^2 o_P \sigma^2)}} \tag{2.5}$$

where $o_P = o_I^2 + o_R^2$, o_P being the object power.
Using Bayes' theorem the PDF for δ, α, β, given p_1, p_2, o can be formulated as,

$$f_A(\delta, \alpha, \beta|p_1, p_2, o) = \frac{f_P(p_1, p_2|o, \delta, \alpha, \beta) f_{A|O}(\delta, \alpha, \beta\ , o)}{f_{P|O}(p_1, p_2, o)} \tag{2.6}$$

where the likelihood, f_P can be plugged in from eq. 2.5. So far, the object is an unknown variable in the density functions. To eliminate it we followed the same trick which the Fineup team [18] used in their method, which is to compute the object solution that maximizes the joint probability, f_P. For this, we found the derivative function of the f_P for the object and equated it to 0 (eq. 2.7). The object solution of the resulting equation (eq. 2.8) is the maximal probable object which is then plugged back in f_P.

$$\frac{\partial f_P}{\partial o_P} \overset{!}{=} 0 \tag{2.7}$$

$$\rightarrow o_P = \frac{-8\sigma^2 + \sqrt{-4(m_1^2 + m_2^2)(-p_1^2/m_1^2 - p_2^2/m_2^2) + 64\sigma^4}}{2(m_1^2 + m_2^2)} \tag{2.8}$$

The likelihood is derived by plugging in eq. 2.8 in eq. 2.5 and then the eq. 2.5 in eq. 2.6. The prior $f_{A|O}$ is assumed to be flat because the availability of object information does not offer any evidence of the aberration. The marginal likelihood, $f_{P|O}$ is a constant normalization factor because it is independent of the aberration and is dropped out. Taking the *log* of the resulting function gives us the final log-likelihood, eq. 2.9.

$$\mathcal{L}(\delta,\alpha,\beta)=\sum\frac{p_1+p_2-\sqrt{p_1^2+\frac{m_2^2p_1^2}{m_1^2}+p_2^2+\frac{m_1^2p_2^2}{m_2^2}+16\sigma^4}}{4\sigma^4}+ln(m_1^2+m_2^2)$$

$$-ln(2m_1m_2\pi\sigma^2(-4\sigma^2+\sqrt{(1+\frac{m_2^2}{m_1^2})p_1^2+(1+\frac{m_1^2}{m_2^2})p_2^2+16\sigma^2})) \tag{2.9}$$

The aberration state which maximizes the log-likelihood function 2.9, is then determined by either non-linear numerical optimization or a closed-form solution explained in the next section. In the case of numerical optimizations, we observed that BFGS (or the limited memory version, L-BFGS) [53, 54], performed the best.

2.2.2 Closed-form Solution

Closed-form solutions are derived using Taylor expansion of the log-likelihood function, eq. 2.9, around the aberration state $(0,0,0)$ up to the second order. This gives a polynomial in three dimensions. The first and second-order coefficients, c_{lmn} of this polynomial, (eq. 2.10) are determined by the Jacobian and the Hessian matrix (eq. 2.11). The log-likelihood curve is therefore parametrized using the 12 coefficients, and an approximate paraboloid function is determined (Fig. 2.3a). The aberration state that maximizes this curve is the desired value, (eq. 2.13 - 2.15), which was derived by solving the set of equations 2.12 for (δ,α,β).

$$g(\delta,\alpha,\beta)=\sum_{l,m,n=0}^{1}c_{lmn}\delta^l\alpha^m\beta^n\ ,\quad l+m+n<=2 \tag{2.10}$$

$$=\begin{pmatrix}\delta,\alpha,\beta\end{pmatrix}\begin{pmatrix}\frac{\partial\mathcal{L}}{\partial\delta}\\ \frac{\partial\mathcal{L}}{\partial\alpha}\\ \frac{\partial\mathcal{L}}{\partial\beta}\end{pmatrix}+\begin{pmatrix}\delta,\alpha,\beta\end{pmatrix}\begin{pmatrix}\frac{\partial^2\mathcal{L}}{\partial\delta^2} & \frac{\partial^2\mathcal{L}}{\partial\delta\partial\alpha} & \frac{\partial^2\mathcal{L}}{\partial\delta\partial\beta}\\ \frac{\partial^2\mathcal{L}}{\partial\alpha\partial\delta} & \frac{\partial^2\mathcal{L}}{\partial\alpha^2} & \frac{\partial^2\mathcal{L}}{\partial\alpha\partial\beta}\\ \frac{\partial^2\mathcal{L}}{\partial\beta\partial\delta} & \frac{\partial^2\mathcal{L}}{\partial\beta\partial\alpha} & \frac{\partial^2\mathcal{L}}{\partial\beta^2}\end{pmatrix}\begin{pmatrix}\delta\\ \alpha\\ \beta\end{pmatrix} \tag{2.11}$$

$$\frac{\partial\mathcal{L}}{\partial\delta}=0,\frac{\partial\mathcal{L}}{\partial\alpha}=0,\frac{\partial\mathcal{L}}{\partial\beta}=0 \tag{2.12}$$

$$\delta = \frac{-c_{011}^2 c_{100} + c_{002}c_{020}c_{100} + c_{010}c_{011}c_{101} - c_{001}c_{020}c_{101} - c_{002}c_{010}c_{110} + c_{001}c_{011}c_{110}}{2\left(-2c_{011}c_{101}c_{110} + c_{002}c_{110}^2 + c_{011}^2 c_{200} + c_{020}\left(c_{101}^2 - c_{002}c_{200}\right)\right)} \tag{2.13}$$

$$\alpha = \frac{c_{011}c_{100}c_{101} - c_{010}c_{101}^2 - c_{002}c_{100}c_{110} + c_{001}c_{101}c_{110} + c_{002}c_{010}c_{200} - c_{001}c_{011}c_{200}}{2c_{020}c_{101}^2 - 4c_{011}c_{101}c_{110} + 2c_{002}c_{110}^2 + 2c_{011}^2 c_{200} - 2c_{002}c_{020}c_{200}} \tag{2.14}$$

$$\beta = \frac{c_{110}(c_{011}c_{100} + c_{010}c_{101} - c_{001}c_{110}) - c_{010}c_{011}c_{200} + c_{020}(-c_{100}c_{101} + c_{001}c_{200})}{2\left(-2c_{011}c_{101}c_{110} + c_{002}c_{110}^2 + c_{011}^2 c_{200} + c_{020}\left(c_{101}^2 - c_{002}c_{200}\right)\right)} \tag{2.15}$$

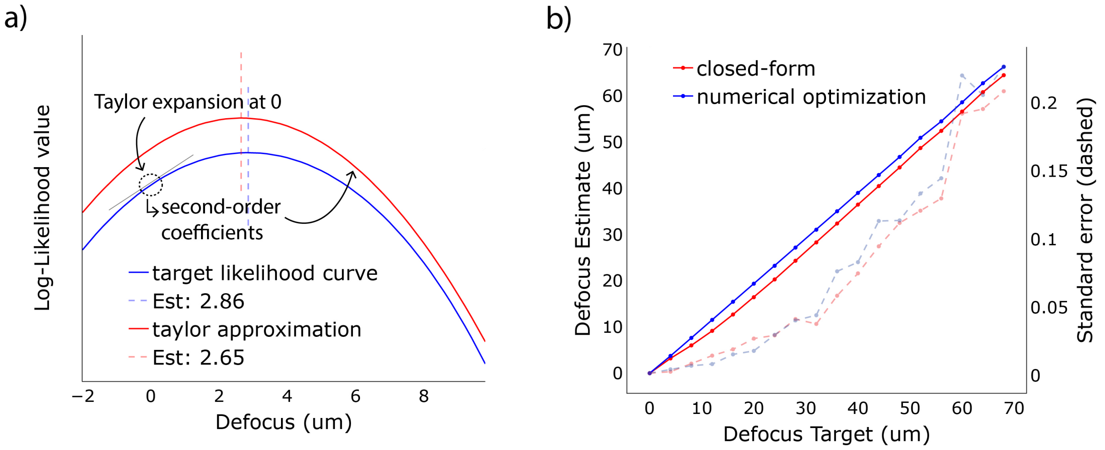


Figure 2.3: a) Log-likelihood curve plotted in blue (eq. 2.9) for values at different defocus from -2 to 8 um (no astigmatism) and an approximation (in red) made by determining the coefficients of the parabola using first and second order derivatives at $(\delta, \alpha, \beta) = (0, 0, 0)$ (eq. 2.11). b) The estimations for a target defocus 0 to 70 μm are shown in red for closed-form and in blue for optimization. The standard error is shown in dashed. The experiments were done using synthetic data (2.2.3), repeated 50 times.

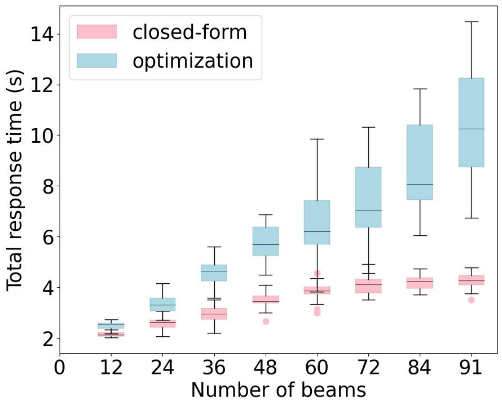


Figure 2.4: Comparing closed-form and optimization in total response time

One advantage of using the closed-form solution is the improvement in computation time. The estimation of derivatives for each data point is an independent process, which can be parallelized over multiple cores. This is not possible when using numerical optimization, where the data points are pooled together, each with its own MTF value. Multiple iterations are run over this pool with updated MTF values each time before the convergence is reached. As can be seen in Fig. 2.4, the total response time in the case of optimization increases with the number of beams, reaching ∼13s while the closed-form solution saturates at ∼4s. These estimations were performed for 91 beam pairs of size 1560x1360 on a cluster with the configuration mentioned in section 2.3.8. Furthermore, the Hessian matrix, a byproduct of this method, can be used as an error metric to flag any improbable estimations as a higher curvature at the estimation point indicates a more confident estimation. However, this study does not include a disciplined investigation of the subject.

2.2.3 Sample and Setup details

The SNR threshold was set at 25 for all the experiments unless specified otherwise.

Setup A

Carl Zeiss Ultra, Field-Emission Scanning Electron Microscope operated at a beam landing energy of 1.5 keV, the beam current at 212 pA, image pixel resolution of 2048×1536, and at a dwell time of 100 ns.

Setup B

Carl Zeiss MultiSEM, Multi-Beam Electron Microscope operated at beam landing energy 1.2-1.5 keV, beam current at 692 pA, and at a dwell time of 200 ns.

Sample A

A neural tissue volume from a mouse brain, stained and embedded in Epoxy, prepared by Maria Kormacheva using the rOPO staining protocol as described in Kormacheva 2023, [55, see Ch. 2].

Sample B

A neural tissue volume from a zebra finch brain, stained and embedded in Epoxy, prepared by Jonas Hemesath using the Song staining protocol [56].

Synthetic data

The synthetic data was generated using a blank 8-bit frame of size 512×512, all valued at 0. 30 seeds were randomly selected on the frame and were then made to walk on the frame in random directions, turning the values of the path they took to 255. The direction taken at every step was also chosen randomly from the 8 different directions with a higher probability of maintaining their direction. Examples are provided in section 2.3.5.

2.3 Results

2.3.1 Resilience against phase shifts

The first claim was resilience against phase shifts. This was verified through simulations conducted on single-frequency 1D and multi-frequency 2D objects (images), as depicted in Fig. 2.5a and b.

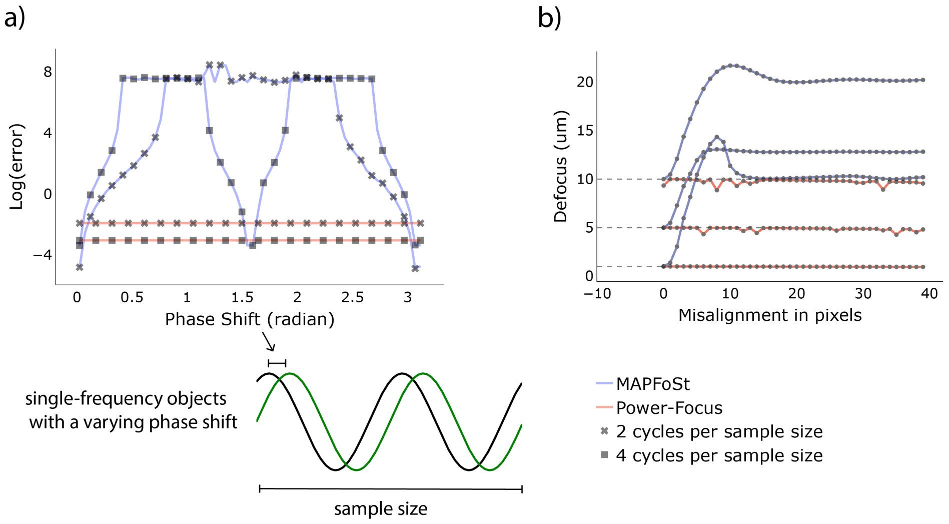


Figure 2.5: *Phase-shift versus error* a) Comparison for single-frequency sinusoidal objects. b) Comparison for simulated 2D objects with synthetic data. The misalignment was induced in the phase diverse images, stepwise in a single direction, and the defocus was estimated using Power-Focus and MAPFoSt for initial target defocus of 1, 5, and 10 μm.

In single-frequency sinusoidal objects, any misalignment in the multiple diversity measurements produces large systematic estimation errors for MAPFoSt, while Power-Focus is insensitive. The errors show dependency on the object frequency. The second set of experiments involved simulated 2D objects (synthetic data in section 2.2.3). Here, it was observed that MAPFoSt's accuracy in estimation deteriorates rapidly with increasing misalignment among phase diverse images, while Power-Focus remains resilient.

2.3.2 Advantage of adaptive filtering

We observed a strong dependence of low-pass filtering of the phase diverse images on the robustness. This was also independently reported by several investigators [48–50]. While noise filtering seems the simplest way to reduce bias, we have found that targeting frequencies based on their individual *SNR* works better than just using a low-pass filter with a fixed cutoff. This clear advantage is shown in Fig 2.6. The filtering method was explained in section 2.2.

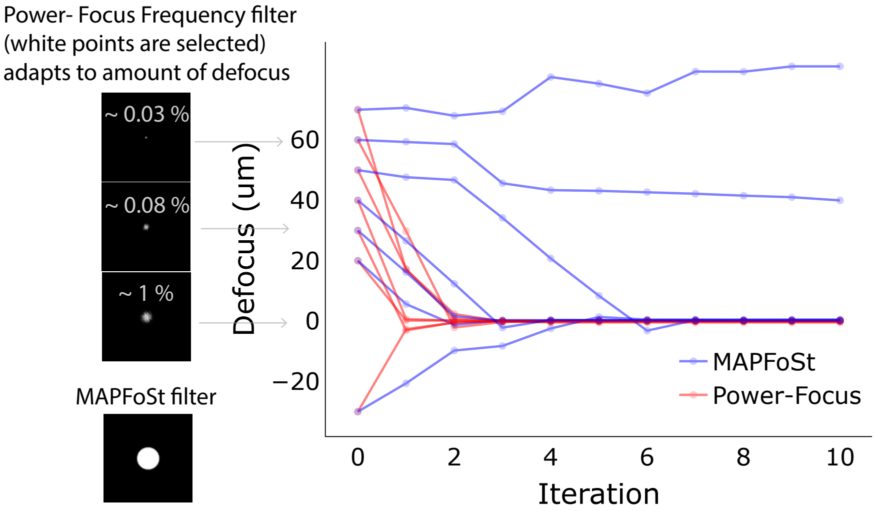


Figure 2.6: *Convergence from large defocus:* Power-Focus was compared to MAPFoSt for large defocus-only aberrations. The frequency filters on the left are for MAPFoSt (bottom) and Power-Focus(top three).

MAPFoSt begins to perform poorly after 40 μm while Power-Focus converges quickly within 3 iterations, even at 70 μm because the *SNR* threshold filtering adapts the number of frequencies to the remaining aberration as it reduces in successive iterations. This keeps in check the retention of high frequencies with high *SNR* making the method robust without compromising on precision. The experiments for this study were done on setup A using sample A at 10 nm pixel size as detailed in section 2.2.3.

2.3.3 Recovery limits

The recoverable range for defocus-only aberrations is found to be +- 150 μm, as shown by the green line in Fig 2.7a. With astigmatism, the recoverable range is (+-120, +-50, +-50) μm for (δ, α, β), as shown by the green shaded area in Fig. 2.7b. The images were taken at 10 nm pixel size for a pixel resolution of 2048×1568. The frame pairs were split into 12 pairs of non-overlapping areas of size 512×512, which were then used for the estimation. It was conducted with setup A, using sample A.

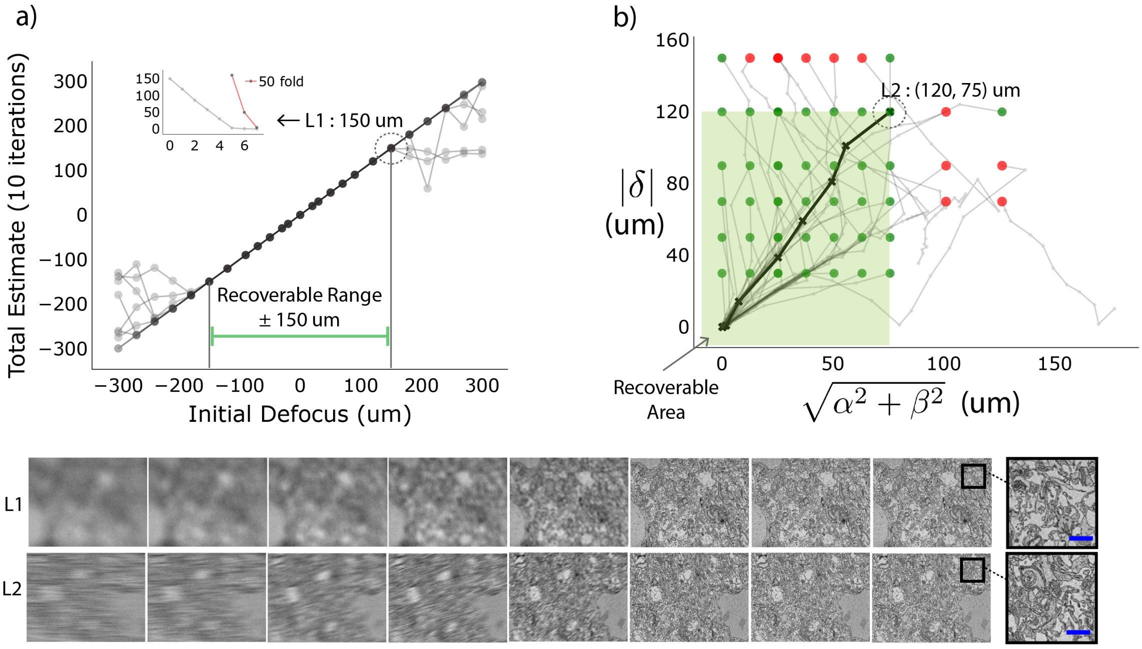


Figure 2.7: *Recoverable aberration range*: a) The method was tested for a range of defocus values within +- 300 μm, repeated at 10 different locations. The recoverable range is shown in green and is around +-150 μm. The convergence for one of the extremities of this range is labeled L1 and the images after every iteration has been shown below the plots. The focusing routine was stopped when the estimations were below 0.1 μm and the maximum iterations allowed were 10 after which the convergence was labeled failed. b) Convergence experiments were performed with astigmatism in the system, keeping $\alpha = \beta$. Successful convergence are dotted in green if it reached a normed residual below 1 μm within 10 iterations, otherwise, it was labeled as failed and dotted in red. The recoverable area is shown in green, the limit of which is (120, 75) μm. The convergence at the limit has been labeled point L2 and the image after every iteration has been shown below the plots. All scale bars (in blue) are 1 μm.

2.3.4 Convergence experiments with non-biological samples

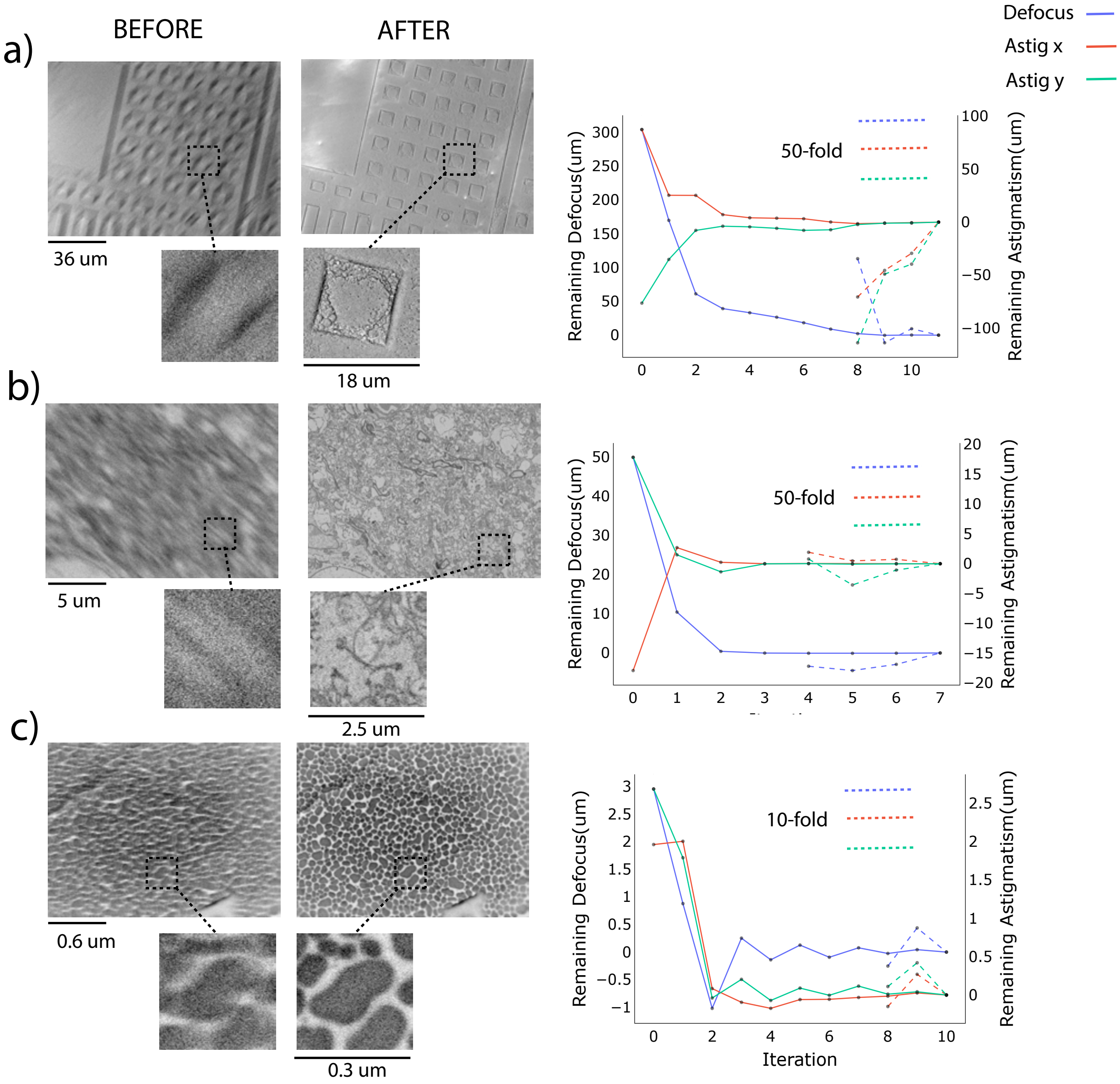


Figure 2.8: Convergence experiments on a.) a decapsulated integrated chip at pixel size 72 *nm* b.) sample A at pixel size 10 *nm* and c.) a gold-on-carbon test specimen from Agar Scientific at pixel size 1.2 *nm*. Multi-fold values are plotted as dashed lines for later iterations to show the fluctuations around zero after convergence. The before and after FOVs are displayed on the left with smaller regions expanded.

We further report that the usability of this method is not limited to biological samples by performing convergence experiments on 2 other samples - a decapsulated integrated chip (prepared with the help of Jonas Hemesath) and a gold-on-carbon test specimen. They were all imaged using setup A but at vastly different magnifications, as shown in Fig. 2.8. The images were taken at a pixel resolution of 2048×1568 and were split into 12 pairs of non-overlapping areas of size 512×512, which were then used for the estimation. The defocus diversities for different magnifications were also linearly scaled with the pixel size, with the value being $\pm 4\ \mu m$ at $10\ nm$.

It should be mentioned that the decapsulated integrated chip began to charge significantly after just a few imaging rounds. Using silver paint around the chip helped decrease this charging to a degree. However, the limited improvement after the second iteration is likely due to ongoing charging issues. This charging effect is visible in the 'after' images as the shiny white areas.

2.3.5 Simulations

We used simulations to evaluate the performance behavior of the method as the noise in the system varies. The results are shown in Fig. 2.9. The desired aberration (δ, α, β) was induced into synthetic objects (section 2.2.3) to produce the image. The images were then made noisy using a Gaussian noise model with the desired standard deviation, σ_n, and a zero mean. Fig 2.9a shows examples of the generated data for varying σ_n and δ. The *SNR* of the images was calculated by eq. 2.16.

$$SNR = \frac{\sum_i p_i - \sigma_n^2 N^2}{\sigma_n^2 N^2} \tag{2.16}$$

where $\sum_i p_i$ is the total power of the input images and N is the sample size which in this experiment is 512×512.

For the first simulation, we generated the input data for a target defocus of $15\ \mu m$ with very little noise variance and made an estimate. The noise variance was then increased gradually until the *SNR* of the input data became very close to 0. We simulated new data for each estimate. This was repeated 200 times, the mean of which has been plotted in Fig 2.9b against the *SNR*. The experiment was repeated for different targets from 0 to $30\ \mu m$ and it was observed that the estimate had an increasing bias towards 0 as the *SNR* was lowered and fell sharply below *SNR* 1.

The precision of the estimate, as shown in Fig 2.9c, also worsens with decreasing *SNR* following an inverse relation.

Lastly, it was found that the SNR_T was critical to the robustness of the method, and increasing it reversed the bias towards 0. However, this came with a trade-off of increasing standard error, which resulted in a lower precision (Fig 2.9d).

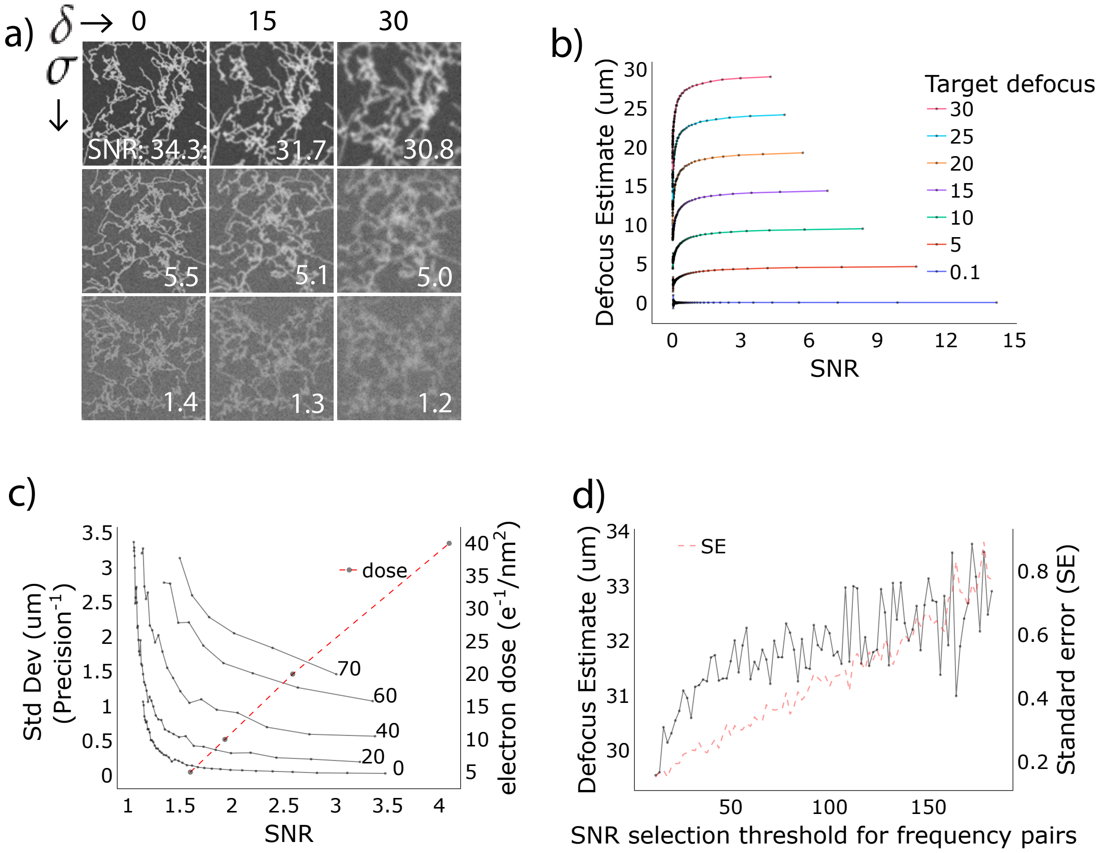


Figure 2.9: *Simulation experiments* a) Examples of synthetic data for varying defocus and noise variance. The SNR is given as an inset. b) The defocus was estimated with increasing noise variance, which varied the SNR of the input data. Every data point in the plot is the mean defocus estimate of 200 measurements. c) The precision of the estimates (std dev of 200 measurements) against falling SNR is plotted for the target defocus of 70, 60, 40, 20, and 0 μm. The dashed line in red is the corresponding electron dose. d) The defocus estimate with varying SNR_T is plotted in black. Every data point is the mean defocus estimate of 200 measurements. The standard error is plotted in red.

2.3.6 Calibrating Power-Focus

A calibration routine should be run before running experiments on a fresh sample and/or new imaging parameters to find the three machine parameters needed to use the focusing model, stigmator rotation (ω), stigmator scaling (γ), and numerical aperture (na). They are essential to calibrate the general optical model to the real-world setup. The na is used in computing the MTF, ω is used to rotationally align the (k_x, k_y) coordinate space of the MTF to the stigmator space which will often be oriented at an angle from the image scan directions, and γ is used to scale the unit of μm of the astigmatism estimation to the machine unit of % which is the control variable on the microscope and denotes the amount of stigmation as compared to the maximum stigmation capacity. The astigmatism in the machine unit % is estimated by the following eq. .

$$\alpha_m = \alpha \times \gamma, \ \beta_m = \beta \times \gamma \tag{2.17}$$

, where α, β are the astigmatism values in μm and α_m, β_m are the astigmatism values in the machine unit %. Results from a typical calibration convergence experiment are plotted in Fig. 2.10.

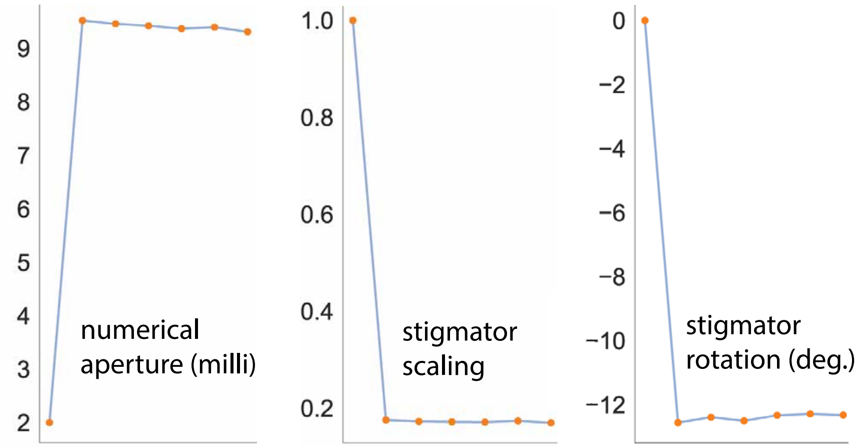


Figure 2.10: Calibration results: the three parameters (*y-axes*) against the update iteration (*x-axes*)

To find these parameters, the focusing is first done manually and a known defocus (δ_K) is introduced. Starting with three initial machine parameters, ($\omega_0 = 0, \gamma_0 = 1, na_0 = 0.002$), a defocus estimate ($\delta_E$) is made using the Power-Focus method, the updated value of na is then determined by the eq. 2.18.

$$na = \sqrt{na_0^2 \frac{\delta_E}{\delta_K}} \tag{2.18}$$

The introduced defocus is then reversed and a known astigmatism (α_K, β_K) is introduced in the system. As before, an astigmatism estimate (α_E, β_E) is made and the remaining two machine parameter values are updated by minimizing the cost function, eq. 2.19 for ω and by eq. 2.20 for γ.

$$cost_\Omega(\omega) = [tan^{-1}(\frac{\alpha_R}{\beta_R}) - tan^{-1}(\frac{\alpha_K}{\beta_K})]^2 \tag{2.19}$$

where,

$$(\alpha_R, \beta_R) = \begin{pmatrix} \cos 2\omega & -\sin 2\omega \\ \sin 2\omega & \cos 2\omega \end{pmatrix} \begin{pmatrix} \alpha_E \\ \beta_E \end{pmatrix}$$

$$\gamma = \gamma_0 \frac{\sqrt{\alpha_K^2 + \beta_K^2}}{\sqrt{\alpha_E^2 + \beta_E^2}} \tag{2.20}$$

The calibration routine is repeated until the values stabilize which was found to be typically around 2-4 update cycles.

2.3.7 Field Curvature in MultiSEM

With the current state of hardware development, the field curvature in the MultiSEM is not negligible, and minimizing it for later generations is an ongoing field of research. In the machine we have, there is a notable difference between the focal points of the central and peripheral beams, and the focal points of the various beams fall on a tilted parabolic surface as shown in Fig. 2.11 (left), while the maximum difference typically ranging from 2.5 to 3.5 μm.

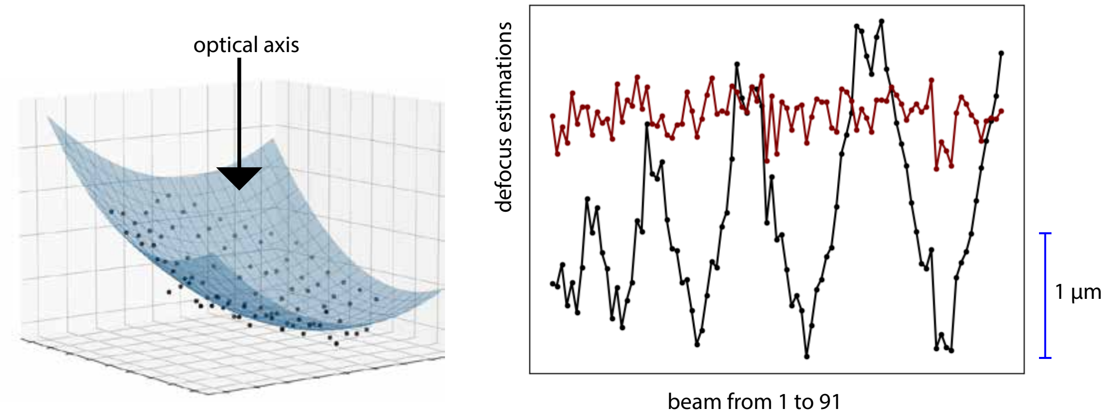


Figure 2.11: *Correcting for field curvature .* *Left:* optimal focal points estimated for the 91 beams plotted across their xy positions. *Right:* the defocus estimations versus the beam number, corrected for the field curvature (in maroon) and without any correction (in black)

The beams, however, share a common working distance from the objective lens because of how the hardware is designed. This results in a varying defocus across the 91 beams. Although the blurring of the beams is so slight that it is challenging to identify by visual inspection and does not pose any danger to future data processing, the modeling of the field curvature is crucial for determining the optimal shared imaging plane.

In scenarios where the MultiSEM FOV only partially contains the brain sample, and the remaining area is blank or lacks contrast, the average focal plane can shift towards an aggregate of the individual optimal focal planes of the beams containing some contrast. Such a shift distorts the surface model of the wafer, which is undesirable for applications like large-scale brain imaging projects. To correct this curvature, focus estimation is initially conducted for each beam individually, and the model of the parabolic surface is recorded. This model serves to adjust the defocus diversity for the individual beams. Instead of the common value of ± 4 μm, a list of 91 defocus diversities is fed, one for each beam pair. These diversities are shifted such that the final imaging plane intersects right through the parabolic surface, and the center and peripheral beams bear the maximum defocus. After applying this correction, the defocus estimation loses its correlation to its beam position, as shown in Fig. 2.11(right), where the corrected estimations are plotted in maroon, and the uncorrected ones are plotted in black.

2.3.8 Time profiling

The computation time for processing a pair of 512x512 frames is about ~270 ms, as detailed in Figure 2.12.

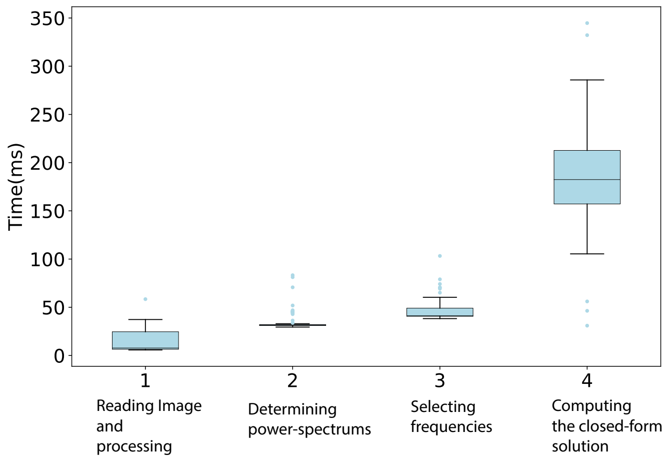


Figure 2.12: *Time profiling:* The breakdown of total computation time across the four main steps, analyzed over 80 beam pairs.

Breaking it down: reading the images is the quickest step at ~10 ms; calculating the power spectrums takes ~30 ms; filtering the frequencies requires ~42 ms; and the most time-consuming part is the solution estimation using the closed-form method, taking ~180 ms. We use a multicore machine and multiprocessing for the MultiSEM to process 91 beam pairs concurrently, reducing the computation time to approximately 1.8 seconds, achieving a speedup of ~14×. The program was run on a cluster equipped with 72 cores, powered by an Intel(R) Xeon(R) CPU E7-8867 v4 @ 2.40GHz, with 1.4 TB of memory. It was written in Python 3.6.8. The images were opened using Python Imaging Library (PIL, [57]), and the image operations were done using the NumPy module(version 1.12.1).

2.4 Discussion

We have found that using a combination of a power-only based, phase diverse focusing method with an *SNR* threshold filter achieves a remarkably strong level of robustness. Aligning the phase diverse images before estimating is no longer required. We also derived a closed-form solution of the cost function using an approximation by Taylor expansion of the log-likelihood function around zero, which yielded a speedup of ~3×. We find that the recovery range in a defocus-only setup for the biological sample used in this study imaged at 10 nm pixel size, is around +-150 μm. With astigmatism, the recovery range is (+-120, +-50, +-50) μm. We also report that the usability of this method is not limited to biological samples. Non-organic samples, subject to good conductivity, imaged at vastly different magnifications also show good convergence.

Chapter 3

Data validation and visualization

For milling-based methods where the successive layers can only be exposed after the preceding layers are irrevocably lost, the scan must be validated for gaps before the next milling starts. Any gaps in the data can potentially lead to errors when the raw data is segmented and labeled to reconstruct neural pathways. To address this, we have created a set of functions that serve as checkpoints, detecting data gaps before the wafer exits the imaging chamber. Additionally, as our imaging setup includes two stages, one being a custom, in-house addition, it was essential to calibrate this stage to correct its spatial target positions, a part of which has been explained in this chapter.

Among the visualization tools, we have introduced functions for quick montage creation, allowing for the detailed examination of small data segments at high resolution or broader areas at lower resolution, which has been detailed here. Also, outlined in this chapter is a method we used to align small data stacks, facilitating the visualization of volumetric data.

3.1 Introduction

In the current development state, there are two stages in the imaging chamber of our GCIB-MultiSEM setup. The lower stage is the standard feature of a MultiSEM offered as an integrated microscope component from Zeiss, referred to as the mechanical stage as it uses mechanical motors to span a large space in the imaging chamber and has three degrees of freedom, x, y, and z. The second stage, fixed on top of the mechanical stage, was designed and developed by Winfried Denk and Jürgen Tritthardt in the Denk lab, and I assisted with its calibration. This stage has two degrees of freedom achieved by moving

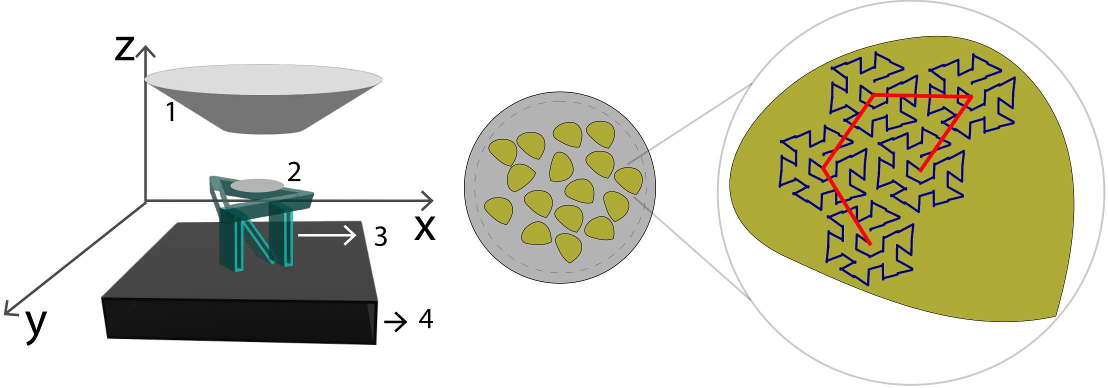


Figure 3.1: *Left:* The imaging chamber setup with two stages. 1: Objective lens, 2: sample wafer, 3: piezo stage, 4: mechanical stage. *Center:* brain samples on the silicon wafer. *Right:* movement of the piezo stage in blue and movement of the mechanical stage in red covering a part of the sample

two sapphire rods by applying voltages to two attached stacks of piezoelectric crystals. The mechanical stress changes by varying the two piezo voltages, producing a controlled two-dimensional movement. This stage is referred to as the piezo stage.

The range of motion of the piezo stage is designed to be within 1 mm, much lower than the mechanical stage, which can span a large volume of the imaging chamber. The utility of the piezo stage is in its ability to stabilize within ∼100 ms, which is far less than the ∼1 s ring downtime of the mechanical stage. This approach should effectively improve the acquisition speed by ∼ 2×. The entire area of the samples on the wafer is scanned by combining mechanical stage and piezo stage movements, shown in Fig. 3.1. The region covered through the movement of the piezo stage, while the mechanical stage remains stationary, is referred to as a *piezo scan*. In Fig. 3.1(rightmost), there are 5 piezo scans, each centered at a different position of the mechanical stage.

The piezo scan also differs from a regular scan in the way the xy deflection is achieved during the raster. In the regular $2D$ raster scan mode of a typical SEM, the beam is deflected in the xy direction by energizing two sets of electromagnet coils built into the microscope. In a piezo scan, however, the y deflection is achieved by moving the stage in the $-y$ direction instead of using the y beam deflection. This mode is called the linescan mode and is a part of the in-built microscope functionality. It only uses the x deflection, and the resulting scan is a line instead of a $2D$ image.

In our imaging setup, a single acquisition from the MultiSEM captures a hexagonal area with a diagonal measurement of ∼165 μm, referred to as the *hexagon*. The piezo movement captures 37 such hexagons covering a larger hexagon (a piezo scan) of a diagonal of around 1 *mm*. After the piezo stage has covered this path, the mechanical stage moves to a neighboring location, which causes the whole piezo stage, along with the wafer, to be centered at a new field of view. A new piezo scan is captured, and this process repeats until the whole sample is covered. The mechanical stage then moves to a new sample, usually a neighboring sample, and the cycle repeats.

3.2 Piezo Stage Calibration

The calibration of the piezo stage involved finding a set of parameters that could translate the piezo stage's xy movement vector into the voltages applied to the piezo stack. It was observed that the parameters had non-linear dependencies on the piezo stage location. Initially, the attempts to find the non-linear parameters were largely unsuccessful (I was only partially involved). We observed that as the errors were consistent, we could use them to adjust the 37 target piezo positions to compensate for the anticipated discrepancy. Only rotation and scaling parameters were kept, and the non-linearity was accounted for by the use of adjusted targets.

3.2.1 Method

The sapphire rods in the piezo stage are oriented at around 45° from the mechanical stage xy coordinate space. The space in which these rods move is referred to as the ab space, the two degrees of freedom corresponding to their movement. A coarse calibration orients and scales the ab space to the xy space as shown in Fig 3.2. The movement in the ab space is in green and the xy space is in red. The piezo scan corresponding to the 37 hexagons is montaged and displayed as an inset near the top center.

For further improvements after a coarse calibration, the error vectors are calculated between the input target xy positions and xy positions reconstructed from scan data. Error vectors are then used to correct the input target vectors to compensate for the error vector along the 37 positions.

Implementing a fast method that uses scan data to estimate the actual path that the piezo stage took actually served two purposes. We could use it to fine-calibrate the piezo target positions and validate the scan for sufficient overlaps along the border of the MultiSEM hexagons (hereafter referred to as hexagons). To compute this path estimate, referred to as the reconstructed piezo positions, the following step-by-step approach was found to be a working solution.

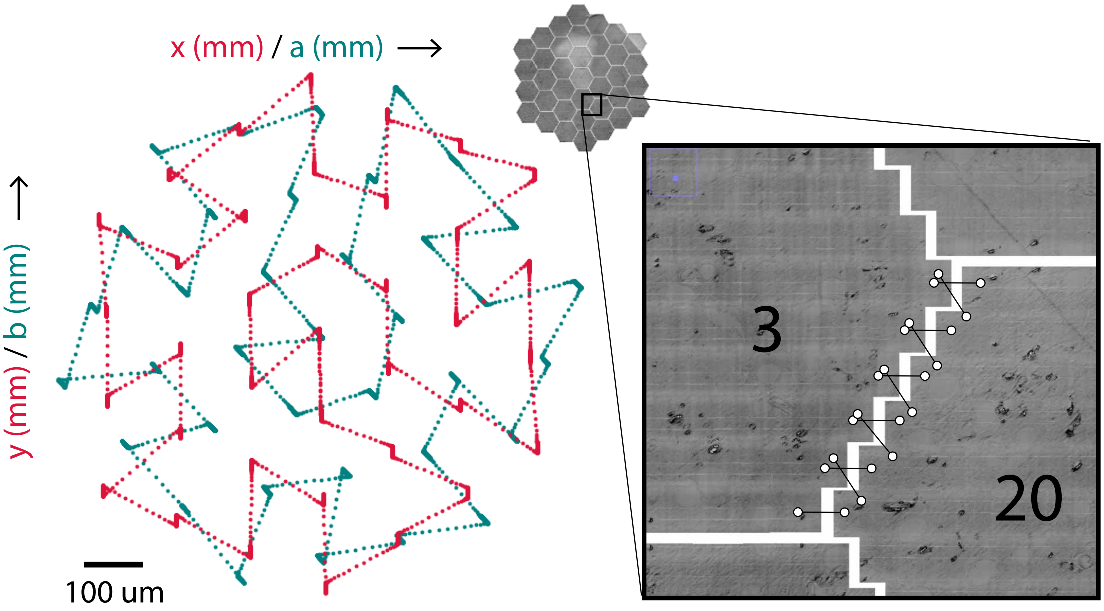


Figure 3.2: *Left:* The path taken by the piezo stage in the xy space (in red) and ab space (in green) sampled at 2 ms. *Right:* Magnified view between hexagons 3 and 20 of the piezo scan montage shown on the top left. The hexagons are positioned at deliberately stretched-out locations, which produces some space between them. The frames connected using circle-ended lines are examples of the frame neighbors between hexagons 3 and 20.

Listing neighboring frames

We first list the neighboring frames within the same piezo scan but belonging to different hexagons, an example of which is shown in 3.2 (right). A k-d tree of frames is used to find k nearest neighbors for each frame lying on the perimeter of each hexagon. Since k depends on the uncertainty of overlaps of neighbouring frames, a value of 2 was found to work for our setup.

Computing shift vectors

The next step is finding all the shift vectors between every neighbor. A standard way to determine the shift vector is by cross-correlating the frames and locating its highest peak (eq. 3.1). We also treat the cross-correlation to an operation that enhances the prominent features of the signal, which in our case is the peak, and also suppresses the noise before computing the shift vector. This operation is given by eq. 3.2. The vector from the center of the cross-correlation to the peak is the shift vector between the two frames. To be robust against large shifts that are greater than the frame's half-size, a well-known problem called phase-wrapping, we first pad the images with $0s$ along all 4 edges, increasing its size to twice the original, and then windowing it using a Tukey filter.

$$CC = F^{-1}(F(f).F^{*}(g)) \tag{3.1}$$

$$SCC = \Delta_x^2 CC \Delta_y^2 CC - G * \Delta_x^2 CC \Delta_y^2 CC \tag{3.2}$$

, where Δ_i^2 is the discrete second difference along axis i and the convolution with G is a Gaussian smoothing operation.

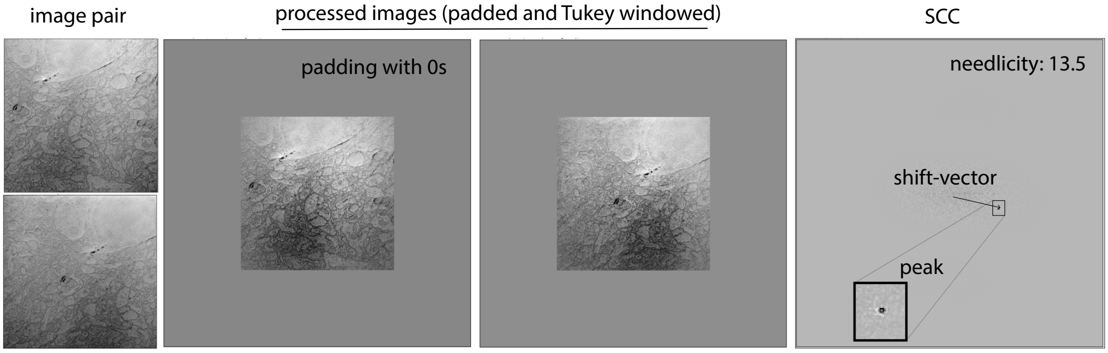


Figure 3.3: *From left to right:* The two EM scan pairs with some translation. After processing, the two images are padded with 0s from all sides, and a Tukey window is applied. The SCC (3.2) shows the shift vector from the center to the peak (also shown in the zoomed square).

A confidence metric, which we call *needlicity* is also calculated, given by eq. 3.3. It is an unpublished heuristic metric, developed by Winfried Denk, that gives a numeric value to the sharpness of the peak. This value is critical in filtering out the false shift vectors, often the result of cross-correlating frames with little to no overlap.

$$needlicity = 15 - 0.4 \times \ln\left(10^{-20} + \sum_{i=1}^{n} 1_{CC_i > \frac{\max(CC)}{20}}\right) \tag{3.3}$$

, where the value is $1_{CC_i > \frac{\max(CC)}{20}}$ stands for 1 when the cross-correlation value at index i is greater than 20 times the maximum value of the cross-correlation, otherwise it is 0. In most cases, the needlicity value is characteristically higher for true values but in case of little overlap between neighboring frames, the chances of the shift vector computation giving a false value substantially increase. To address this issue, we use the target location of the frames to estimate roughly the overlapping area of the neighboring frames. We then use this cropped intersecting area to compute the shift vector by cross-correlation as shown in Fig. 3.4.

Also, as the shift vector in the cropped case involves an extra translation of the second frame, absent in the full frame case, this new translation vector is added to the shift vector before returning. In our setup, the needlicity of 12.5 or greater is classified as qualifying. One challenge that requires further attention is that the needlicity threshold is susceptible to changing frame size. For validating shift vectors of full frames in our dataset, a needlicity threshold of 12.5 works well. However, if we compute the shift vectors of cropped frames, the effectiveness of the same needlicity threshold diminishes.

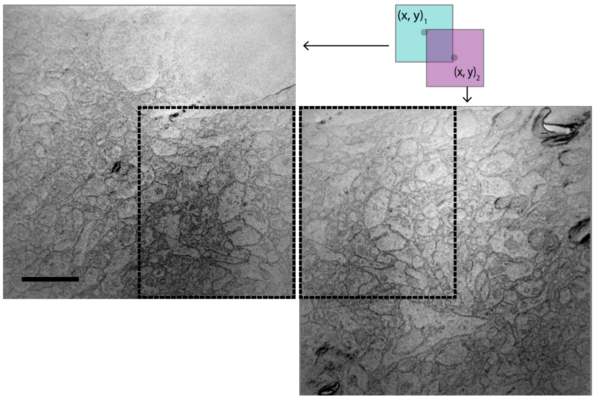


Figure 3.4: *Cropping the frames in case of little overlap.* $(x, y)_1$ and $(x, y)_2$ are the target positions of the two frames. The scale bar in black is 1 μm.

Reconstructing the piezo positions

After the computation of shift vectors, the reconstruction problem essentially boils down to just solving a system of linear equations, and constructing these linear equations from the shift vectors is the next step. Every shift vector estimation has a corresponding needlicity, and we use a needlicity threshold of 12.5 to reject the false values. Every neighboring frame pair is given an index of 5 dimensions: $(\alpha_a, \beta_a, \alpha_b, \beta_b)_\eta$. Here, α_a, α_b is the hexagon index, and β_a, β_b is the beam index of the first and second frame respectively. η is the index that runs from 1 to n_s, where n_s is the total number of qualifying shift vectors.

$$H(\alpha_a) - H(\alpha_b) = B(\beta_a) - B(\beta_b) + \delta_\eta \tag{3.4}$$

,where δ_η is the shift vector of neighboring frame pair index η , $H(i)$ is the reconstructed position of hexagon index i and $B(i)$ is the beam position of beam index i. There are n_s

equations like these, and the objective is to solve them for $H(i)$ where i runs from 1 to 37, for 37 hexagons in a piezo scan. Changing the left hand-side of the equation and moving to its matrix-form, we can write the set of equations as,

$$\begin{pmatrix} 1 & -1 & 0 & 0 & . & . & 0 \\ 0 & 1 & -1 & 0 & . & . & 0 \\ 0 & 0 & 0 & 1 & . & . & 0 \\ 0 & 1 & 0 & -1 & . & . & 0 \\ 1 & 0 & -1 & 0 & . & . & 0 \end{pmatrix} \times \begin{pmatrix} H(1) \\ H(2) \\ . \\ . \\ H(n_s) \end{pmatrix} = \begin{pmatrix} B(62) - B(83) + \delta_1 \\ B(72) - B(91) + \delta_2 \\ . \\ . \\ B(76) - B(88) + \delta_n \end{pmatrix} \tag{3.5}$$

Solving for $H(1), H(2), .., H(n_s)$

$$H(i) = S^+ \times (B(\beta_a) - B(\beta_b) + \delta_\eta) \tag{3.6}$$

where S^+ is the pseudoinverse of the sparse matrix in eq. 3.5 (with 1s, -1s and 0s). The beam numbers given in the matrix taken are just exemplary. The final solution is eq. 3.6.

Fine calibration of piezo stage target positions

We estimated the new targets by correcting the original targets using the reconstruction as given by eq. 3.7.

$$t_n(i+1) = t_n(i) + (t_o - r(i)) \tag{3.7}$$

$$t_n(0) = t_o$$

, where $t_n(i)$ are the new targets for iteration i, t_o are the original targets and $r(i)$ are the reconstructed positions for iteration i. We adjusted the target positions by incorporating the anticipated error, thereby offsetting the identified discrepancy. This cycle is repeated till the maximum discrepancy is in the window of tolerance.

3.2.2 Result

The results from a recent calibration procedure for the corrected targets of a piezo stage are presented here. The targets (t_0) are shown in Fig 3.5(left). The new targets, (t_n) are iteratively determined following the equation 3.7. The objective is to find a set of corrected targets t_n for which the reconstruction agrees well with t_0. After five iterations of calibration, the maximum discrepancy between the original target positions (t_0) and the reconstructed (or recalibrated) positions is approximately 1.3 μm, with the average value being around 0.2 μm. The reduction in error through subsequent iterations is visually represented in the Fig. 3.5(right). The analysis of errors involves comparing the target positions (t) and the reconstructed positions (r) using the equation $10 \times t - 9 \times r$. The red lines in Fig. 3.5(right) represent this comparison, highlighting the discrepancies between the t_0 and reconstructed positions and how they change iteratively over the course of calibrating.

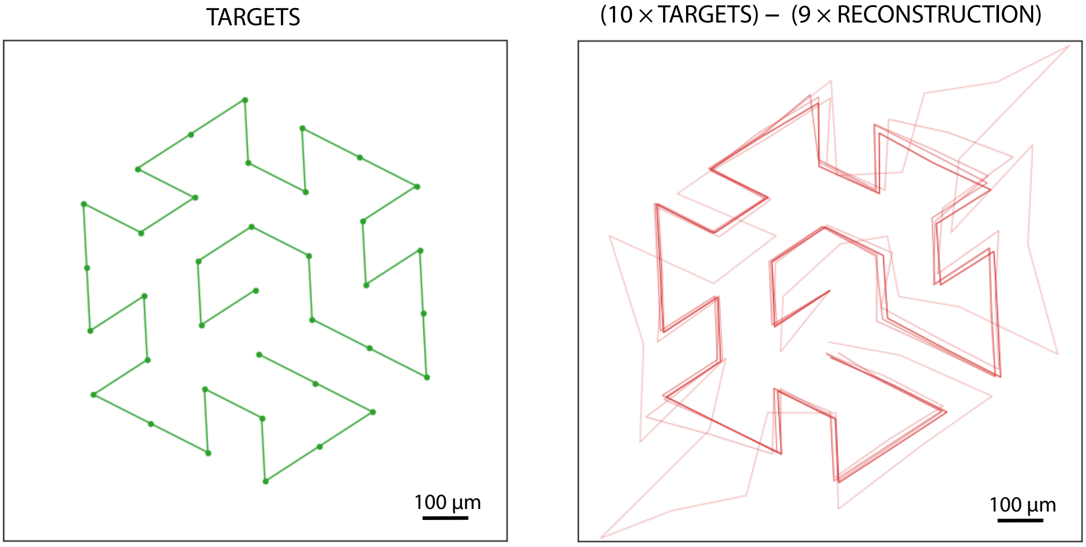


Figure 3.5: *Left:* The 37 xy target positions of the piezo stage are shown in green dots, and the target path is shown as the green line. *Right:* The error between the reconstructed and the target positions is visually represented in red through the equation 10×t - 9×r, where t are the target positions and r are the reconstructed positions. The variation in transparency among the lines corresponds to different iterations, ranging from 1 to 5, with the transparency decreasing progressively; the final iteration is depicted with the least transparency.

3.3 Detecting gaps within the piezo scan

Once the piezo is well calibrated, the likelihood of finding gaps is low. However, it is still important to continue validating the incoming piezo scans to keep the calibration in check. The presence of gaps in the data can be catastrophic at a later stage in the pipeline, rendering the full dataset or a part of it useless.

3.3.1 Method

The sufficient condition for a piezo scan to have no internal gaps is that the complete perimeter of each of its hexagons intersects with the neighboring hexagon's perimeter. There should not be any residual perimeter that has no intersection.

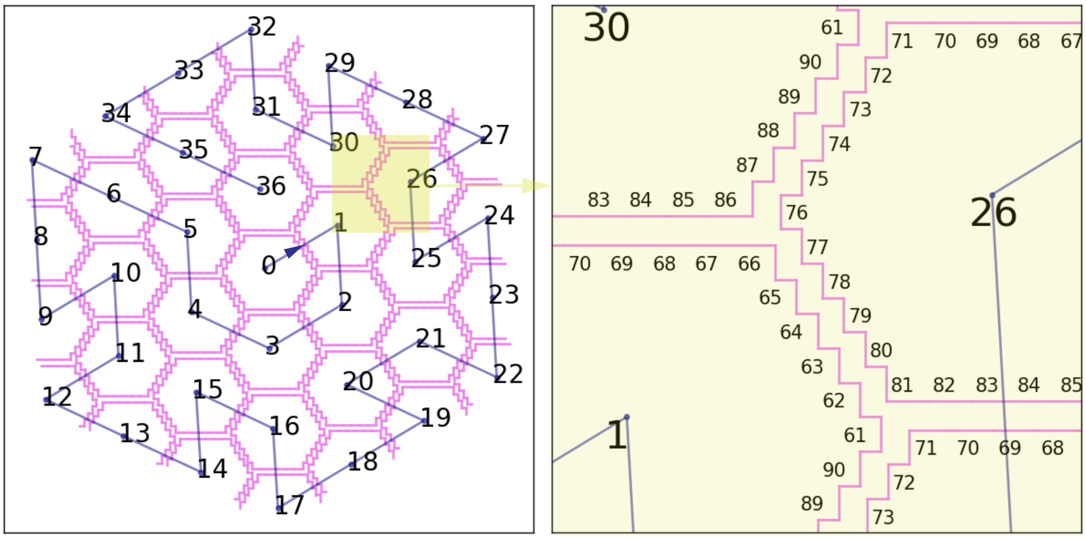

32
33
29
34
31
28
27
7
35
30
26
6
36
24
8
5
1
0
25
10
23
9
4
2
3
11
21
12
15
20
22
13
16
19
14
18
17
30
61
71 70 69 68 67
90
72
89
73
88
74
87
75
83 84 85 86
76
26
70 69 68 67 66
77
65
78
64
79
80
63
81 82 83 84 85
62
1
61
71 70 69 68
90
72
89
73

Figure 3.6: *Left:* The perimeter of 37 sequentially numbered hexagons of a piezo scan is shown in purple. The target sequence of the piezo stage is shown in blue. *Right:* A zoomed-in view of the yellow shaded area from the left image, showing the border beam frames' indices.

This perimeter is shown in Fig 3.6 as purple lines. The piezo positions are stretched out by 10 % to clearly show this perimeter. Every purple edge belongs to a beam frame which itself belongs to a hexagon. The amount of overlap is a simple subtraction problem involving the neighboring frames' width, height, and location. The complete routine has been presented in detail below.

Inputs

1. List of 37 piezo positions $(h_x, h_y)_i$ where i runs from 0 to 36.
2. List of frame positions for a zero-centered hexagon, $(b_x, b_y)_j$, where j runs from 0 to 90 for 91 beam frames
3. The width (w) and height (h) of the beam frames.

Output

A list of overlapping areas and a list of remaining exposed edges.

Steps

1. Create a lookup table in the form of a dictionary that maintains a record of the edges of each of the 37 hexagons in a piezo scan, including only the edges that border a frame belonging to the same hexagon, as shown by purple lines in Fig 3.6. This lookup table is referred to as exposed edges. Each frame should have four values corresponding to the top, bottom, left, and right edges.
2. Create a k-d tree using the 30 × 37 frame coordinates belonging to the 30 border frames of the 37 hexagons. These frames are indexed from 61 to 90 because only the border beams are taken.
3. Iterate over the frames in the exposed edges and query the k-d tree for 6 nearest neighbors within a radius of $\sqrt{w^2 + h^2}$.
4. Iterate over all the found neighbors. Skip and continue if a neighbor belongs to the same hexagon. If it belongs to a different hexagon, calculate the overlap using the

   $$overlap_x = w - |x_a - x_b| \; \& \; overlap_y = h - |y_a - y_b| \tag{3.8}$$

   , where (x_a, y_a) is the coordinate of the current frame and (x_b, y_b) is the coordinate of the neighboring frame.
5. If the overlap (eq.3.8) is positive, update the exposed edges of that frame by subtracting the intersecting perimeter from the corresponding direction - top, bottom, right, or left.

6. Append the positive overlaps to a result list.
7. Repeat the steps for every frame updating the overlaps result list and the exposed edges.

If the exposed edges have no positive values at the end of the iteration, then the piezo scan has no gaps and the acquisition proceeds without interruption. The result list can be used to keep track of the amount of overlap within the piezo scan.

3.3.2 Result

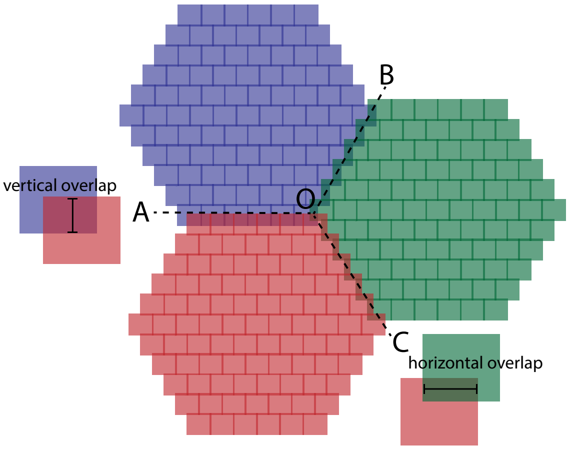


Figure 3.7: A triplet of neighboring hexagons with sufficient overlap.

Histograms are used to visualize the distribution of overlaps within a piezo scan. The vertical overlap within the piezo scan is only taken by frames belonging to either the top or the bottom edge of the neighboring hexagons, an example of which is shown by the red and blue hexagons in Fig. 3.7, which share the border OA. A single example of the vertical overlap has been shown by the red and blue frames. The horizontal overlap is only taken from pairs of frames belonging to the remaining 4 edges of neighboring hexagons, as shown by the red and green hexagons in the same figure, which share the border OC, and the blue and green hexagons, which share the border OB. The horizontal overlap is also shown using a single pair of red and green frames. The piezo calibration experiment presented in the previous section had accompanying histograms, which have been shown in Fig. 3.8. The first iteration revealed vertical gaps (in red) as well as horizontal gaps (in blue) as negative overlaps. The later histograms show a narrower distribution, indicating reduced errors and improved alignment between the target and the reconstruction.

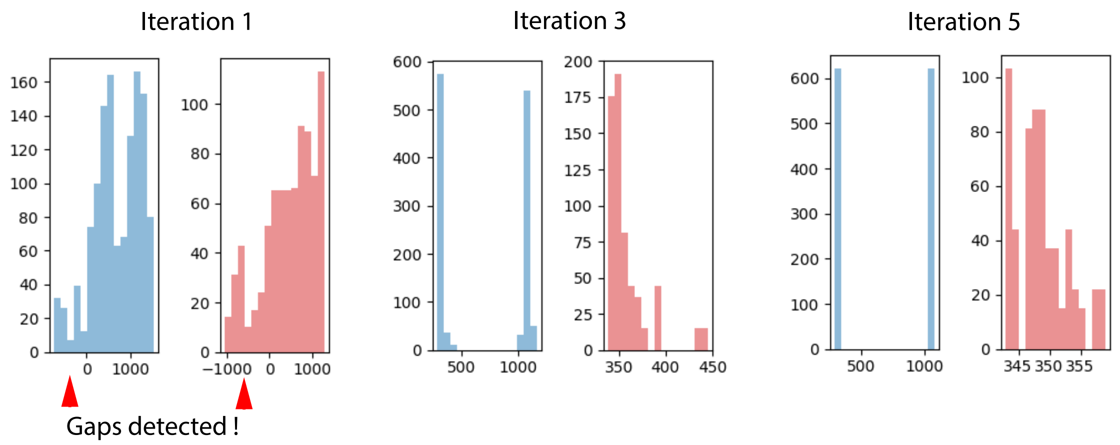


Figure 3.8: The evolution of the histogram, showing vertical overlaps in red and horizontal overlaps in blue (in pixel units), through a typical piezo target calibration process.

3.4 Montaging

Creating a montage from single or multiple piezo scans involves merging the encompassed beam frames into one comprehensive image. They provide a quick way to visually inspect the acquisition and validate the reconstruction's accuracy.

3.4.1 Method

A single piezo scan produces 91 image files, one for each beam. A single beam file contains accrued data from each of the 37 hexagons as shown in Fig 3.9b. This file also contains the transitory acquisition, the region above the top red arrow and below the bottom red arrow. The region within the arrows is the beam frame, which is cropped out when the montage is created.

Every piezo scan follows a trigger signal, which controls the time instances when the piezo will move to the next hexagon location. The time between two consecutive triggers is when the scan is taken. These signals are saved along with the rest of the metadata in the scan folder. The montaging function uses it to crop out 37 beam frames from the long accrued frames. The conversion from the trigger signals in time to 37 y-pixel locations from where the frame can be cropped is done using the set of eq. 3.9

$$y_{crop} = r_{ms} \ \times \ [j \mid \Delta_j > 0.1 \text{ for } j \in (0, 1, 2, 3, ...n)] \tag{3.9}$$

where Δ is the discrete difference along the time axis, j is the time index from 0 to n, and r_{ms} is rows traveled in 1 ms, which in our setup is ∼1.59 currently.

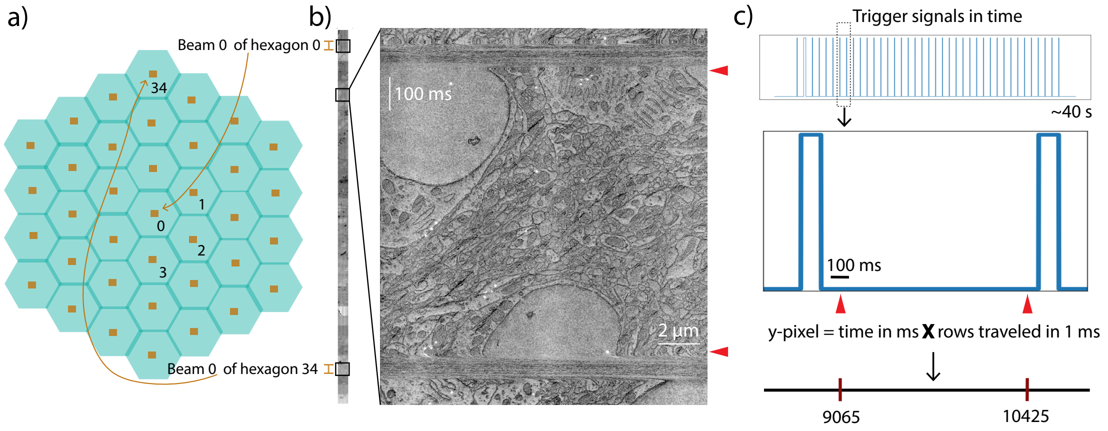


Figure 3.9: a) A piezo scan is illustrated in blue, with the center beam frames highlighted in orange. b) The long scan on the left is the center beams of the 37 hexagons accrued one after the other, with the expanded view of the center beam of hexagon 4 displayed on the side. c) The trigger signals are shown on the top with the fourth trigger expanded below, the time between the red arrows is when the frame was acquired. The corresponding y pixel location of the red arrows in (b) is plotted at the bottom.

We use the Python Imaging Library (PIL [57], PyPI repository: *pillow*) to create a blank canvas of maximum size 1 GB on which the individual frames are pasted.

In constructing a montage from a single piezo scan, the placement of each beam frame is determined by adding the piezo position of its hexagon to the beam location identified by its beam index. When the montage involves multiple piezo scans, this method is adapted to account for the varying center positions of the piezo scans. Specifically, the frames' paste locations are further adjusted by adding the piezo center positions, which translates the next piezo scan to a different location.

3.4.2 Result

A piezo montage is shown below (Fig 3.10). Some hexagon borders are depicted as black dashed lines to demonstrate effective stitching and successful piezo position calibration. The border that appears in white is concealed beneath the overlaying frames because the visibility of a border also depends on the sequence of the frame pasting on the montage.

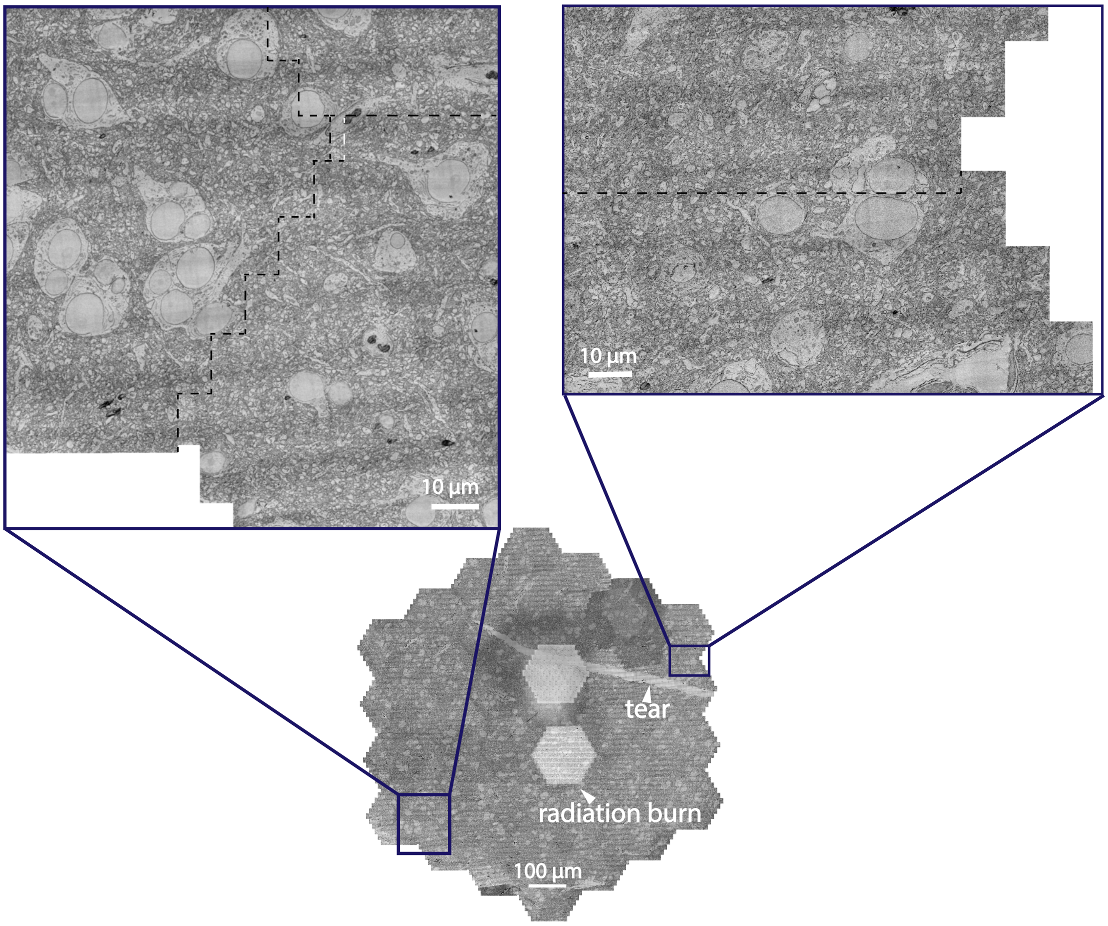


Figure 3.10: A piezo montage with two rectangular regions magnified. Black dashed lines are the hexagon borders. The white dashed line is the border concealed beneath the overlaying frames.

An overview montage containing multiple piezo scans covering almost the complete sample area of the two samples on the wafer (top) is shown in fig. 3.11.

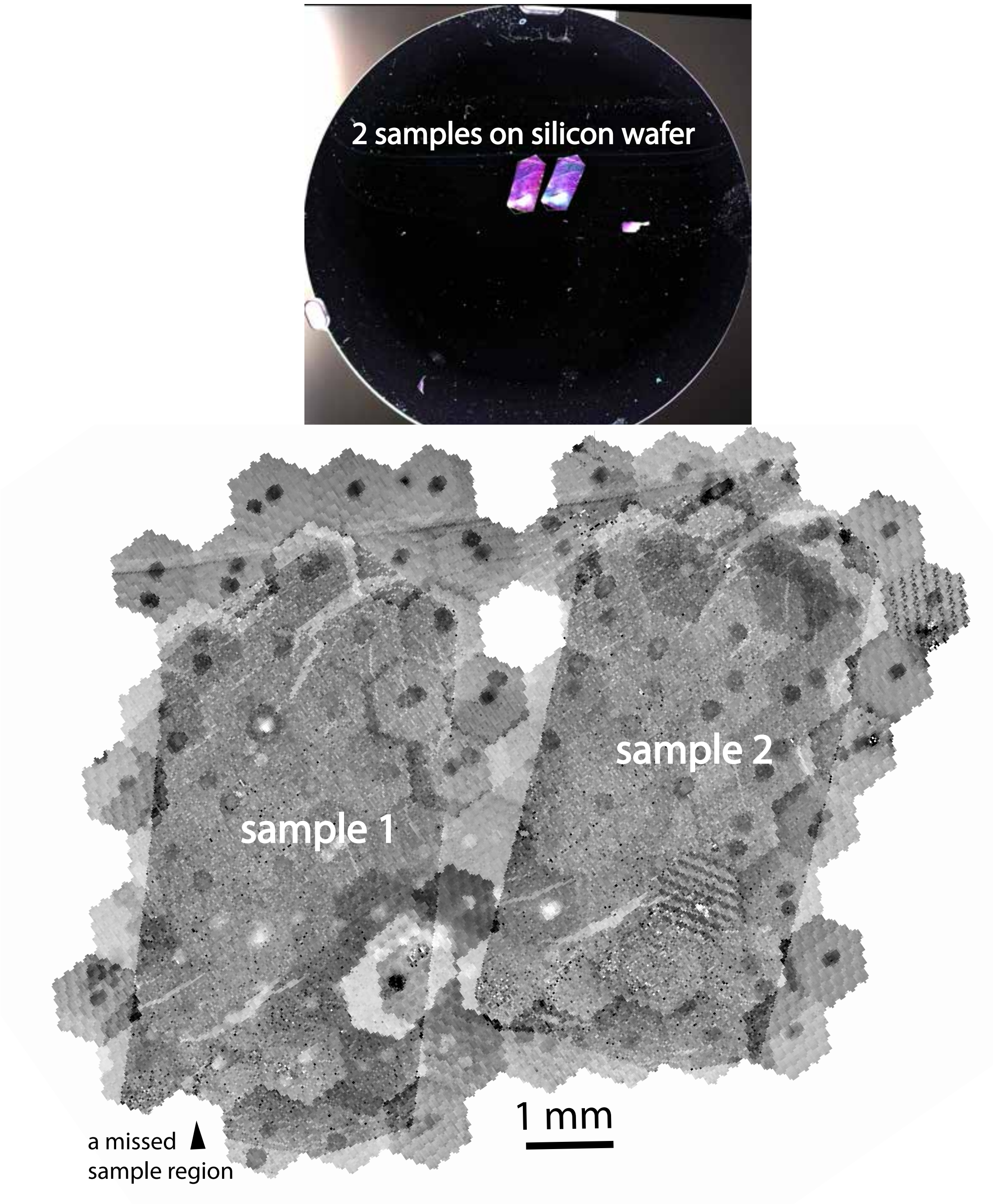

2 samples on silicon wafer
sample 1
sample 2
1 mm
a missed
sample region

Figure 3.11: *Top:* A camera image of a wafer with two slices of brain samples (purple) taken by Jonas Hemesath. The sample was also prepared by Jonas Hemesath (sample B 2.2.3). *Bottom:* An overview montage of acquisitions taken by Winfried Denk (setup B, 2.2.3)

3.5 Detecting gaps between piezo scans

As the mechanical stage transitions to another piezo center for subsequent piezo scanning, we have to ascertain that there are no gaps in the interstitial space between neighboring piezo scans. We use a similar method to the one that detects gaps within the piezo scan (3.3).

3.5.1 Method

This validation is done using only two piezo scans at a time. The first step is to find the beam frames along the border of the piezo scan which is neighboring the border of the adjoining piezo scan. The pairs of neighboring frames are cross-correlated to find the shift vectors between them. The piezo centers are reconstructed, two at a time, and the gaps are detected using the piezo exposed edges lookup table.

Finding neighboring frames

As the piezo scans are taken sequentially and the validation is done in parallel, the neighboring piezo scans are found using a k-d tree of the set of coordinates of the piezo scan centers that have already been acquired. The overlap validation is pair-wise.

To find the bordering frame indices between a pair of neighboring piezo scans, a k-d tree is created from the set of coordinates of all the frames belonging to them. The k-d tree is then queried for pairs of frames closer than $\sqrt{w^2 + h^2}$ and the ones that belong to different piezo scan indices are collected in a list. Every neighboring frame pair is labelled with 6 indices : $((\pi_a, \alpha_a, \beta_a), (\pi_b, \alpha_b, \beta_b))$, where (π_a, π_b) are the piezo indices, (α_a, α_b) are the hexagon indices and (β_a, β_b) are the beam indices of the two frames a and b.

Computing shift vectors

The shift vectors between the neighboring frames are computed between full frames following the same method explained in section 3.2.1. The list of shift vectors along with their needlicity and the six indices $((\pi_a, \alpha_a, \beta_a), (\pi_b, \alpha_b, \beta_b))$ is saved to a file.

Reconstructing piezo centers

The reconstructed piezo centers are computed following the same technique as in section 3.2.1.

The solution to the equation Eq. 3.10 gives the reconstructed vector. One qualifying shift vector is sufficient to solve this equation as every other quantity on the RHS is known.

$$\Pi_{r0}(\pi_a) - \Pi_{r1}(\pi_b) = (\Pi_0(\pi_a) - \Pi_1(\pi_b)) + (B(\pi_a, \alpha_a, \beta_a) - B(\pi_b, \alpha_b, \beta_b) - \delta_\eta) \tag{3.10}$$

, where $\Pi_{r0}(\pi_a) - \Pi_{r1}(\pi_b)$ is the reconstructed vector and $\Pi_0(\pi_a) - \Pi_1(\pi_b)$ is the target vector for the piezo indices (π_a, π_b), δ_η is the shift vector and $B(\pi_i, \alpha_i, \beta_i)$ is the frame beam frame location belonging to the π_i piezo index, α_i hexagon index, and β_i beam index for the two neighbouring frames i. where $i \in (a, b)$. As the set of equations is overdetermined, we use an average of the final result.

Finding gaps between the piezo scans

As in the case of finding gaps within a piezo scan where we use a lookup table of exposed edges along the borders of hexagons within a piezo, the gaps between two piezo scans use a lookup table of exposed edges along the border of the piezo scan (Fig. 3.12).

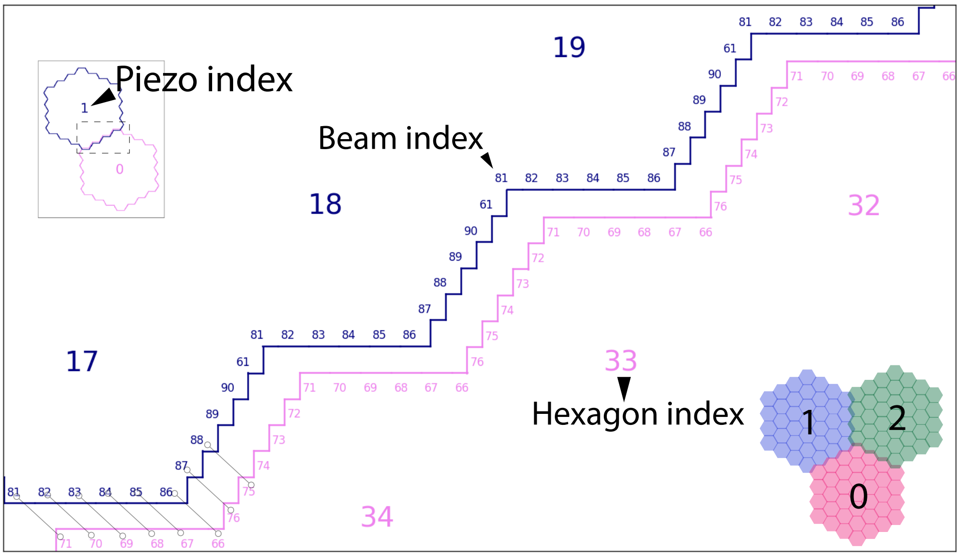


Figure 3.12: The border between two piezo scans (0 and 1) is shown in blue and pink. The hexagon and beam indices, which belong to the lookup table for this edge, are given. A few examples of the frames that are neighboring are indicated by connecting black lines.

A single piezo scan, being a hexagon in shape, can border a neighboring piezo in 6 possible directions. and can have a maximum of 6 neighbors. For every neighbor, the bordering exposed edge is selected. A single piezo scan's exposed edges lookup table is a dictionary with keys from 0 to 36 for 37 hexagon indices. The values are dictionaries from 0 to 90 for 91 beam indices. If a beam frame is contained within a piezo and is not at the border, then the corresponding exposed edge value is set to None. The steps of the routine are presented below.

Inputs

1. List of piezo centers $(p_x, p_y)_i$ where i runs from 0 to the total number of piezo centers on the wafer.
2. List of 37 piezo positions $(h_x, h_y)_i$ where i runs from 0 to 36.
3. List of frame positions for a zero-centered hexagon, $(b_x, b_y)_j$, where j runs from 0 to 90 for 91 beam frames
4. The width (w) and height (h) of the beam frames.

Output

A list of overlapping areas and a list of remaining exposed edges.

Steps

1. Create a new k-d tree with reconstructed piezo centers, using all the frame coordinates of the frames belonging to the two piezo scans for which the gap is being estimated. The coordinates of all the frames are calculated using their piezo index, hexagon index, and beam index.

   $$(f_x, f_y) = (p_x, p_y)_\pi + (h_x, h_y)_\alpha + (b_x, b_y)_\beta \tag{3.11}$$

   where π is the piezo index, α is the hexagon index, and β is the beam index of the frame.
2. Query the k-d tree to find pairs of neighboring frames belonging to different piezo scans. Iterate through them while estimating the overlap as in section 3.2.1.
3. If the overlap is positive, update the exposed edges of that frame by subtracting the intersecting perimeter from the corresponding direction - top, bottom, right, or left.

4. Append the positive overlaps to a result list.
5. Repeat the steps for every frame updating the overlaps result list and the exposed edges.

If no positive exposed edges are remaining for the border edge shared between the two piezo scans as shown in Fig. 3.12, then there is no gap. In the example, the piezo scan index 0 and 1 are immediate neighbors with a shared border in the north-west direction from 0. The exposed edge will only have frames from hexagon indices $17, 18, 19, 22$ from piezo 1 and $7, 32, 33, 34$ from piezo 0. Both the piezo scans are also bordering with another piezo: 2. This is repeated for all the immediate neighbors of the piezo scan.

3.5.2 Result

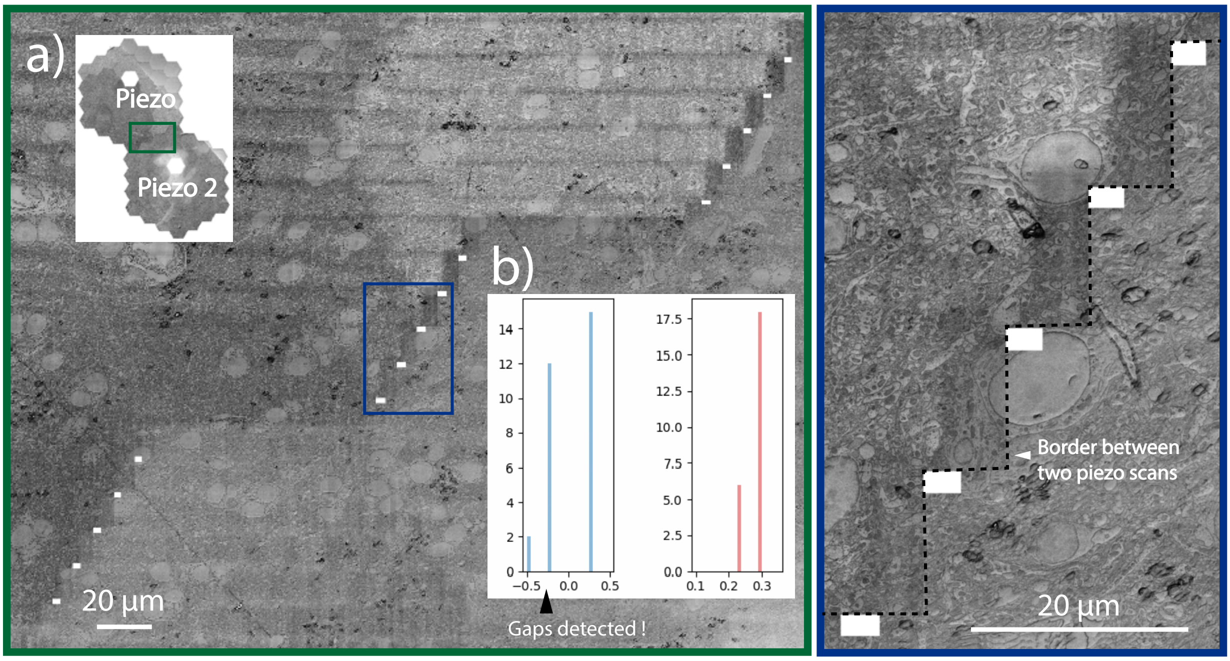


Figure 3.13: *A faulty scan:* The inset (a) shows two bordering piezo scans and the region inside the green window is magnified and shown in the left montage, with a green border. The inset (b) is the histogram of the overlaps (in μm) between the two piezo scans, with the horizontal overlaps in blue and vertical overlaps in red. The montage on the right with the blue border is the magnified region within the blue window on the left montage. The border between the two piezo scans is drawn as a shaded black line.

The vector connecting the center positions of the two neighboring piezo scans was reconstructed, and the updated frame coordinates were used to identify any gaps along their border as outlined in the previous section. This analysis revealed a negative horizontal overlap, indicating a gap between the two scans. Figure 3.13(left) shows a montage that magnifies the area between the two piezo scans (green window in the inset (a)), making the gaps clearly apparent. Further magnification of the area inside the blue window, as shown in the right montage, reveals the effective stitching between the scans. This visual inspection of the stitching quality serves as a reliable method for verifying the accuracy of the reconstruction during the project's development phase.

To streamline the process and maintain automation, without the need for time-consuming visual inspection of each neighboring pair montage, an accompanying histogram of the horizontal and vertical overlaps is generated as shown in the inset Fig. 3.13(b).

3.6 Stack alignment

Once multiple scan layers have been acquired, creating small volume stacks becomes beneficial for observing changes in neural matter along the z-direction, confirming sufficient milling.

The difficulty in frame alignment escalates with the increase in reinsertion errors within the system. In setups where the brain sample experiences minimal movement between milling sessions, aligning images is simpler since shifts typically do not go beyond single frames. In the latest developments of our setup, we observed lateral errors that span beyond a single frame and minor rotational differences among the different layers; because of this, the alignment is tackled in two steps - rough and fine alignment. The rough alignment determines the large movements and brings the shift down within a single frame. Currently, this adjustment has only been achieved visually for each layer. Automating the correction of insertion errors when the wafer comes back from the milling chamber is one of the objectives for future development. The fine alignment is done by cross-correlating overlapping frames from neighboring layers to find a set of frame coordinates that minimize the relative shifts.

In this study, we developed, tested, and employed the fine alignment tool to compile stacks from acquisitions made with the setup in the year 2020. Since then, the tool has seen limited use, and due to significant changes in the setup, the implementation has become outdated and is not included in the code repository. However, the underlying theoretical methodology remains unchanged and is explained below. The stack that was produced in 2020 has been presented in the results section.

3.6.1 Method

The rough alignment, which brings the shifts down within a single frame, was done visually by creating full-resolution hexagon montages that share the same piezo and hexagon index. The coordinates of any distinct common feature, like an organelle across the complete hexagon set, were recorded. The center coordinate of the hexagon belonging to the first layer is given $(0, 0)$. The center coordinates of all the layers on top are translated using the estimated shift as an offset so the common feature is aligned.

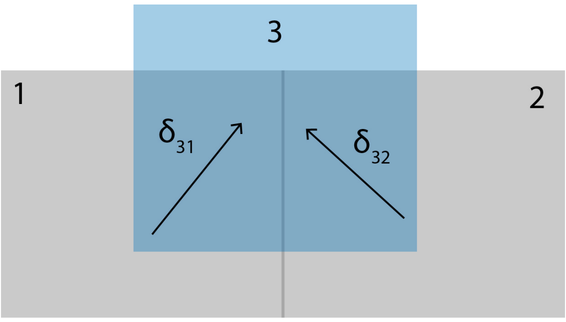


Figure 3.14: Two pairs of overlapping frames. The top layer is shown in blue, and the bottom layer is in gray.

The fine alignment step is essential for ensuring smooth transitions between frames in stack alignment, addressing any remaining relative shifts between frames. Starting with a simple case of two layers with two overlapping pairs of frames, (3,1) and (3,2), as shown in Fig. 3.14, we can formulate the following equations.

$$|(p_3 - p_1) - \delta_{31}| = \epsilon_{31}$$
$$|(p_3 - p_2) - \delta_{32}| = \epsilon_{32}$$
$$p_{ij} = p_i - p_j$$

, where p_i is the position for the frame i, δ_{ij} is the shift vector from frame j to frame i, and ϵ_{ij} is the discrepancy between the vectors p_{ij} and δ_{ij}, $i, j \in (1, 2, 3)$.

This can be extended to incorporate n pairs of overlapping frames, allowing for the formulation of a general equation given below.

$$|(p_i - p_j) - \delta_{ij}| = \epsilon_{ij} \tag{3.12}$$

, where i and j are all the overlapping pairs. A solution of this equation which gives the lowest errors of $\sum \epsilon_{ij} = 0$ is,

$$p_i - p_j = \delta_{ij} \tag{3.13}$$

Finding the positions for every frame in the stack with at least one overlapping neighbor is trivial with the final equation being,

$$P = S^{+}\Delta \tag{3.14}$$

, where P is the row matrix for the xy position coordinates of all the frames in the stack, S^{+} is the pseudoinverse of the sparse matrix, S, with each element being either 0, 1, or -1, that keeps track of the signs, and Δ is the matrix with the corresponding shift vectors.

All the shift calculations in this section are done using the method explained in section 3.2. The complete workflow involves multiple steps, outlined as follows.

1. All the frames in the dataset are labeled by the index $(\pi, \alpha, \beta, \gamma)$ where π is the piezo index, α is the hexagon index, β is the beam index, and γ is the layer index from 1 to n_l, the total number of layers being n_l.

2. $(n_l - 1)$ k-d trees are created for every layer except the first one.

3. A small number of neighboring seed frames are selected from the first layer, which will be the field of view of the final volume stack. Starting with the first seed frame, the overlapping frames from the next layer are identified by searching the k-d tree of that layer, using the current frame's location to find up to 5 nearest neighbors. The initial seed frame is also assigned a single index, 0. This index denoted as f, is then incremented by 1 each time a frame is labeled. A dictionary that maintains the index $(\pi, \alpha, \beta, \gamma)$ of the frame belonging to each index f is also recorded.

4. During the iteration over the resulting frames, each one is checked for overlap by verifying if the distance between any two frames is less than the frame's width and height. The f tuple for each overlapping frame pair is added to a list, which maintains a record of all such pair indices, referred to as F.

5. Steps 3 and 4 are iteratively applied to each seed frame and its selected overlapping frames, continuing this process until all layers have been covered.

6. The set F is iterated over, and shift estimations are performed on the pairs of frames corresponding to each f tuple while keeping track of the signs in the matrix S and the corresponding shift vectors in the matrix Δ.

7. Equation 3.14 is used to solve for P.

While creating montages for every layer, the dictionary that kept track of the index $(\pi, \alpha, \beta, \gamma)$ for every f is used to retrieve the appropriate frame from the dataset and identify the layer to which it belongs.

3.6.2 Result

The presented stack was acquired by Winfried Denk and the brain sample was prepared by Maria Kormacheva and Alexandra Rother. Fig. 3.15 shows three slices from the 39-slice stack to show the common field of view. The changing neural matter in the subsequent slices indicates successful milling, and the re-sliced view in the xz plane shows the alignment as seen from the side. The link to the stack has also been given in the Fig. 3.15.

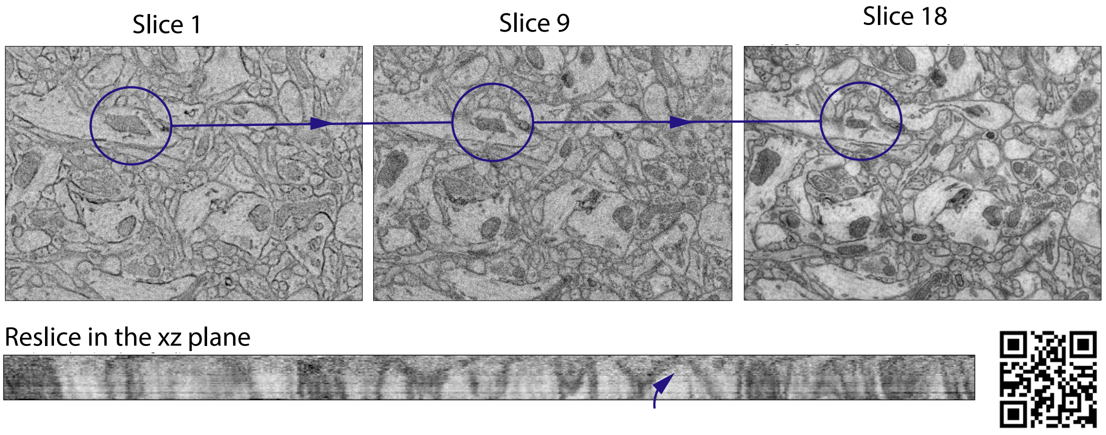


Figure 3.15: Three aligned scan slices illustrate the changes in neural matter due to milling, highlighted within blue circles. The bottom image is a re-sliced view of the stack in the xz plane, with the blue arrow indicating the alignment in the z-direction. The QR code can be scanned to directly access the link where the stack is uploaded. The direct link is : `https://jmp.sh/nyrN0Sm2`

Chapter 4

Conclusion and Outlook

At the outset of the project, our EMs lacked reliable focusing routines. We have since devised and implemented a novel focusing technique that significantly extends the recovery range by $\sim 4\times$ and have made it available on our machines. The newly developed focusing routine has been operational for two years, with the latest update using the closed-form solution implemented over the past year. With the power-only approach and effective noise filtering, the recovery range now falls within the machine movement's uncertainties.

While local focusing works well at specific locations, ensuring that the acquisition stays in focus throughout requires further investigation to predict the wafer surface focus model within tolerance limits. Moreover, we lack a solid theoretical foundation to adaptively adjust the phase diversity (presently at $\pm 4\mu m$) as the system zeroes in on the precise focus. Furthermore, the method is susceptible to sample charging, and our current implementation involves a heuristic approach to detecting and excluding the charging patches from the input data prior to aberration calculation. However, this heuristic method has been omitted from the current report due to its lack of general applicability.

We developed toolkits for the acquisition pipeline that enable the detection of data gaps and facilitate rapid montaging. The reconstruction method initially devised for detecting gaps in piezo positions was effectively adapted to pre-correct target positions of the piezo, proving to be valuable in the calibration of the stage. We successfully aligned a subset of the acquired data stack as proof of concept.

While the toolkits are available for independent use, the development of a controlling program that would automatically trigger and synchronize these functions during data acquisition is still pending. Furthermore, a significant challenge anticipated in stack alignment involves aligning data from multiple slices on the wafer, as these slices may be oriented in various directions on the sample. This process requires sequencing the slices in the order they were cut before standardizing their rotation.

Appendix A

Tools Used

ChatGPT4 [12] was used to find synonyms for words that were closest to my intended thought. It was also used to improve the clarity of specific single sentences, which sounded confusing after the first draft of the thesis. This was done with around 11 sentences in the Introduction, around 3 in Chapter 2, and around 13 in Chapter 3. The reply was never copied directly but a modified version was used in combination with my own changes to write a sentence closest to my intended thought. It was also used heavily to find information related to biology in the Introduction, which was then used as a filter to further research on Google/Google Scholar. Grammarly was used almost constantly to check for any grammatical mistakes and also for better sentence formation. Equations 2.8, 2.9, 2.12-2.15 were derived using Mathematica(13.2). The 4 black and white graphics used in Fig 1.1 were made using ChatGPT4 and have been cited.

Appendix B

Acronyms

PSF: Point Spread Function
MTF: Modulation Transfer Function
PD: Phase diversity
EM: Electron Microscope/Microscopy
VEM: Volume Electron Microscopy
SEM: Scanning Electron Microscope
FEG: Field Emission Gun
GCIB: Gas Cluster Ion Beam
FOV: Field Of View
SNR: Signal to Noise ratio
HHMI: Howard Hughes Medical Institute

Appendix C

Supplementary

C.1 MTF Derviation

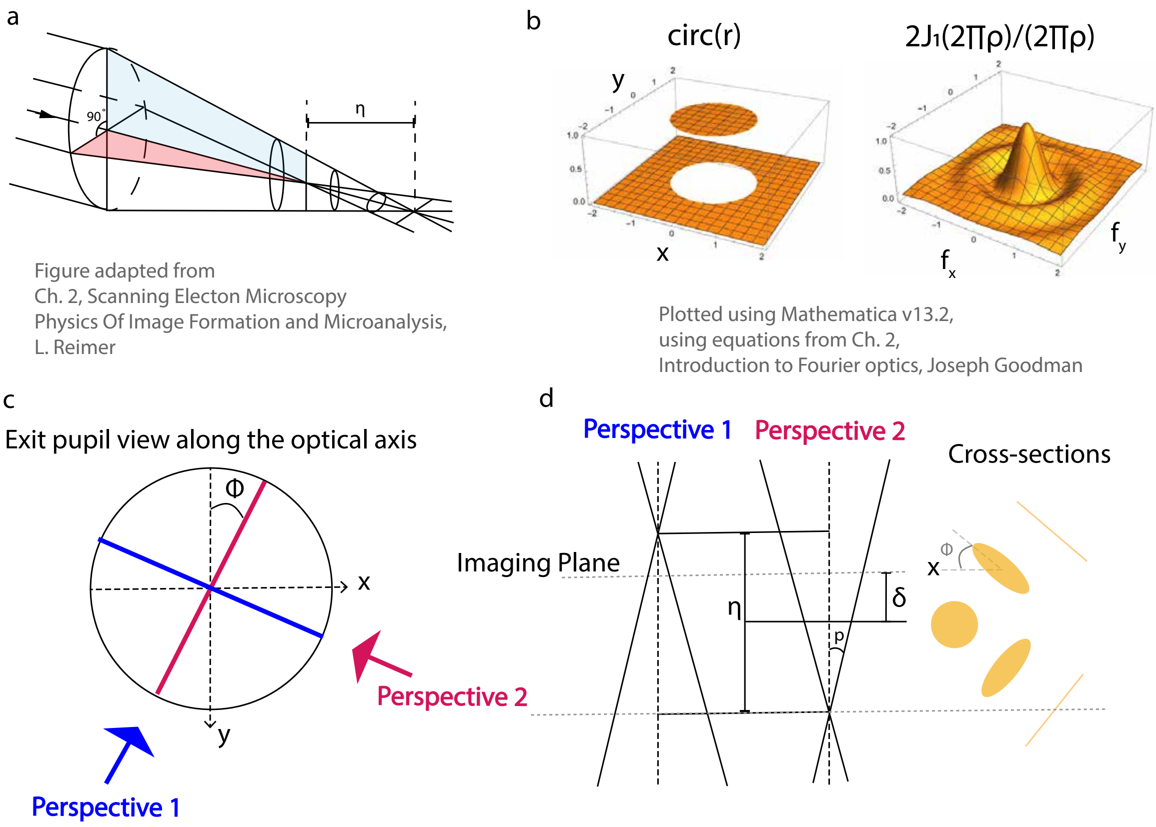

a
90°
η
b
circ(r)
2J₁(2∏ρ)/(2∏ρ)
y
x
f_x
f_y
1.0
0.5
0.0
-2
-1
0
1
2
c
Exit pupil view along the optical axis
Φ
x
y
Perspective 2
Perspective 1
d
Perspective 1
Perspective 2
Cross-sections
Imaging Plane
η
δ
p
x
Φ

Figure adapted from
Ch. 2, Scanning Electon Microscopy
Physics Of Image Formation and Microanalysis,
L. Reimer

Plotted using Mathematica v13.2,
using equations from Ch. 2,
Introduction to Fourier optics, Joseph Goodman

Figure C.1

Assuming geometrical optics, in an optical system with defocus as the only aberration present, and a flat illumination function the cross-section of the beam is a circular disk of radius $r_c = \rho/f * \sqrt{\delta^2}$, where ρ is the pupil radius, f is the focal length and δ is the defocus. The beam probe cross-section becomes elliptical when astigmatism is present in the system. The two semi-axes of the ellipse, A and B depend on the focal length of the sagittal and meridional ray fans which are shifted from the ideal focal point by $\pm\eta/2$.

$$A = na * (\delta + \eta/2)^2, B = na * (\delta - \eta/2)^2 \tag{C.1}$$

, where the numerical aperture $na = sin(p) \approx tan(p)$, the p being the probe convergence angle (see Fig. C.1d). The eccentricity of the ellipse depends on the orientation of the two ray fans from the scan direction xy as shown in Figure C.1a, c, and d. The locus of the boundary of the elliptical disk, defined by the propagation of the marginal rays at the exit pupil, follows the equation

$$\frac{(x\cos(\phi) - y\sin\phi)^2}{A^2} + \frac{(x\sin\phi + y\cos\phi)^2}{B^2} = 1 \tag{C.2}$$

and the resulting, PSF can be written as a piece-wise function,

$$PSF(x, y, \delta, \eta, \phi) = \left\{ \begin{array}{ll} 1, & \text{if } \sqrt{(x\cos(\phi) - y\sin\phi))^2 B^2 + (x\sin\phi + y\cos\phi)^2)A^2} < AB \\ 0, & \text{if } \sqrt{(x\cos(\phi) - y\sin\phi))^2 B^2 + (x\sin\phi + y\cos\phi)^2)A^2} > AB \end{array} \right\} \tag{C.3}$$

Computing, the Fourier transform of an elliptical disk becomes significantly simpler when starting with the Fourier transform of a circular disk, the result of which is widely published. Goodman (p. 16, [52]) shows that the Fourier transform of a circular disk is a Bessel function of the first kind (Fig C.1b). A normalized version of this function with unity at the origin is eq. C.6

$$x' -> (x\cos(\phi) - y\sin\phi)^2 B^2 \tag{C.4}$$

$$y' -> (x\cos(\phi) - y\sin\phi)^2 A^2 \tag{C.5}$$

$$F(circ(\sqrt{x'^2 + y'^2})) = 2\frac{J_1(2\pi\rho)}{2\pi\rho} \tag{C.6}$$

, where $\rho = \sqrt{f_x^2 + f_y^2}$

A rotation and scaling transformation of the axes can then be used to derive the Fourier transform of an elliptical disk,

$$f_x/(2\pi) = \sqrt{(k_x \cos(\phi) - k_y \sin\phi)^2 B^2} \tag{C.7}$$

$$f_y/(2\pi) = \sqrt{(k_y \cos(\phi) - k_y \sin\phi)^2 A^2} \tag{C.8}$$

As the input stigmation vector in the microscope is in cartesian coordinate, we also do the following transformation

$$\eta \rightarrow \sqrt{\alpha^2 + \beta^2} \tag{C.9}$$

$$\phi \rightarrow \arctan(\frac{\alpha}{\beta})/2 \tag{C.10}$$

Plugging, the above equations in C.6, we arrive at the formula for the *MTF*,

$$MTF(k_x, k_y, \delta, \alpha, \beta, na) = 2\frac{J_1(\sqrt{na^2((k_x^2 + k_y^2)(\delta^2 + \alpha^2 + \beta^2) + 2\delta\alpha(k_x^2 - k_y^2) - 4\delta\beta k_x k_y))}}{\sqrt{na^2((k_x^2 + k_y^2)(\delta_2 + \alpha^2 + \beta^2) + 2\delta\alpha(k_x^2 - k_y^2) - 4\delta\beta k_x k_y}}) \tag{C.11}$$

A close Gaussian approximation of this formula for our imaging parameters is

$$MTF(k_x, k_y, \delta, \alpha, \beta, na) = \exp(\frac{-na^2}{8}((k_x^2 + k_y^2)(\delta^2 + \alpha^2 + \beta^2) + 2\delta\alpha(k_x^2 - k_y^2) - 4\delta\beta k_x k_y)) \tag{C.12}$$

This was first published in Binding 2013 [23]. The rationale for presenting this proof lies in the observation that I did not find the explanation provided in the supplementary part of the thesis of Jonas Binding sufficiently detailed and that future research involving the *MTF* must have a thorough understanding of the foundational derivations.

Bibliography

1. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode Caenorhabditis elegans: the mind of a worm. *Phil. Trans. R. Soc. Lond* **314,** 340 (1986).
2. Bargmann, C. I., Hartwieg, E. & Horvitz, H. R. Odorant-selective genes and neurons mediate olfaction in C. elegans. *Cell* **74,** 515–527 (1993).
3. Briggman, K. L., Helmstaedter, M. & Denk, W. Wiring specificity in the direction-selectivity circuit of the retina. *Nature* **471,** 183–188 (2011).
4. Kornfeld, J. *et al.* EM connectomics reveals axonal target variation in a sequence-generating network. *Elife* **6,** e24364 (2017).
5. Zheng, Z. *et al.* A complete electron microscopy volume of the brain of adult Drosophila melanogaster. *Cell* **174,** 730–743 (2018).
6. Bock, D. D. *et al.* Network anatomy and in vivo physiology of visual cortical neurons. *Nature* **471,** 177–182 (2011).
7. Takemura, S.-y. *et al.* A visual motion detection circuit suggested by Drosophila connectomics. *Nature* **500,** 175–181 (2013).
8. Kornfeld, J. & Denk, W. Progress and remaining challenges in high-throughput volume electron microscopy. *Current opinion in neurobiology* **50,** 261–267 (2018).
9. Lichtman, J. W., Pfister, H. & Shavit, N. The big data challenges of connectomics. *Nature neuroscience* **17,** 1448–1454 (2014).
10. Scheffer, L. K. *et al.* A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9** (eds Marder, E., Eisen, M. B., Pipkin, J. & Doe, C. Q.) e57443. ISSN: 2050-084X. `https://doi.org/10.7554/eLife.57443` (Sept. 2020).
11. Dorkenwald, S. *et al.* Neuronal wiring diagram of an adult brain. *bioRxiv* (2023).

12. OpenAI. *ChatGPT-4: Technical Report* `https://openai.com/`. Accessed: 2023-02-21. 2023.
13. Lena Eberle, A. *et al.* Multiple-beam scanning electron microscopy. *Microscopy Today* **23,** 12–19 (2015).
14. Nickell, S. & Zeidler, D. A 331-Beam Scanning Electron Microscope. *Microscopy and Microanalysis* **25,** 568–569 (2019).
15. Hayworth, K. J. *et al.* Gas cluster ion beam SEM for imaging of large tissue samples with 10 nm isotropic resolution. *Nature methods* **17,** 68–71 (2020).
16. Denk, W. & Horstmann, H. Serial Block-Face Scanning Electron Microscopy to Reconstruct Three-Dimensional Tissue Nanostructure. *PLoS Biology* **2,** e329 (2004).
17. Thermo Fisher Scientific. *DualBeam FIB-SEM Microscopes* `https://www.thermofisher.com/de/de/home/electron-microscopy/products/dualbeam-fib-sem-microscopes.html`. Accessed: 2023-02-19. 2023.
18. Paxman, R. G., Schulz, T. J. & Fienup, J. R. Joint estimation of object and aberrations by using phase diversity. *JOSA A* **9,** 1072–1085 (1992).
19. Gonsalves, R. A. Phase Retrieval And Diversity In Adaptive Optics. *Optical Engineering* **21,** 215829. `https://doi.org/10.1117/12.7972989` (1982).
20. Gerchberg, R. W. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35,** 237–246 (1972).
21. Dean, B. H. & Bowers, C. W. Diversity selection for phase-diverse phase retrieval. *JOSA A* **20,** 1490–1504 (2003).
22. Gardner, J. P. *et al.* The james webb space telescope. *Space Science Reviews* **123,** 485–606 (2006).
23. Binding, J., Mikula, S. & Denk, W. Low-dosage maximum-a-posteriori focusing and stigmation. *Microscopy and Microanalysis* **19,** 38–55 (2013).
24. *Google Connectomics Team* `https://research.google/teams/connectomics/`. 2023.
25. Jain, V. Adversarial Image Alignment and Interpolation. *arXiv* **abs/1707.00067.** `https://arxiv.org/abs/1707.00067` (2017).
26. Januszewski, M. *et al.* High-precision automated reconstruction of neurons with flood-filling networks. *Nature methods* **15,** 605–610 (2018).

27. *KNOSSOS - 3D Neuron Reconstruction and Annotation* `https://knossos.app/`. 2023.

28. *Hemibrain Neuroglancer Demo* `https://hemibrain-dot-neuroglancer-demo.appspot.com/`. 2023.

29. Reimer, L. *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis* 2nd. ISBN: 978-3-540-63938-1 (Springer, Berlin, Heidelberg, 1998).

30. Wikipedia contributors. *Zernike polynomials — Wikipedia, The Free Encyclopedia* `https://en.wikipedia.org/wiki/Zernike_polynomials`. Accessed: 2023-02-21. 2023.

31. Hawkes, P. Aberration correction past and present. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367,** 3637–3664 (2009).

32. ZEISS. *GeminiSEM Family - ZEISS* `https://www.zeiss.com/microscopy/en/products/sem-fib-sem/sem/geminisem-family.html`. Accessed: 2023-02-21. 2023.

33. Keller, A. L., Zeidler, D. & Kemen, T. *High throughput data acquisition with a multibeam SEM* in *Scanning Microscopies 2014* **9236** (2014), 69–74.

34. Goldstein, J. I. *et al. Scanning Electron Microscopy and X-Ray Microanalysis* 4th. ISBN: 978-1-4939-6676-9 (Springer, New York, NY, 2017).

35. Rudnaya, M., ter Morsche, H. G., Maubach, J. & Mattheij, R. M. M. A derivative-based fast autofocus method in electron microscopy. *Journal of Mathematical Imaging and Vision* **44,** 38–51 (2012).

36. Rudnaya, M., Mattheij, R. & Maubach, J. Evaluating sharpness functions for automated scanning electron microscopy. *Journal of microscopy* **240,** 38–49 (2010).

37. Kudryavtsev, A. V., Dembélé, S. & Piat, N. Autofocus on moving object in scanning electron microscope. *Ultramicroscopy* **182,** 216–225 (2017).

38. Cui, Le, Marturi, Naresh, Marchand, Eric, Dembélé, Sounkalo & Piat, Nadine. Closed-Loop Autofocus Scheme for Scanning Electron Microscope. *MATEC Web of Conferences* **32,** 05003. `https://doi.org/10.1051/matecconf/20153205003` (2015).

39. Dembélé, S., Lehmann, O., Medjaher, K., Marturi, N. & Piat, N. Combining gradient ascent search and support vector machines for effective autofocus of a field emission–scanning electron microscope. *Journal of microscopy* **264,** 79–87 (2016).

40. Harada, M., Obara, K. & Nakamae, K. A robust SEM auto-focus algorithm using multiple band-pass filters. *Measurement Science and Technology* **28,** 015403. `https://dx.doi.org/10.1088/1361-6501/28/1/015403` (Dec. 2016).

41. Rudnaya, M. Automated focusing and astigmatism correction in electron microscopy (2011).

42. Batten, C. F. Autofocusing and astigmatism correction in the scanning electron microscope. *Mphill thesis, University of Cambridge* (2000).

43. Lee, W. *et al.* Robust autofocusing for scanning electron microscopy based on a dual deep learning network. *Scientific reports* **11,** 20933 (2021).

44. Schubert, P. J., Saxena, R. & Kornfeld, J. DeepFocus: Fast focus and astigmatism correction for electron microscopy. *Nature Communications* **15,** 1–10 (2024).

45. Feinberg, L. D. *et al. TRL-6 for JWST wavefront sensing and control* in *UV/Optical/IR Space Telescopes: Innovative Technologies and Concepts III* **6687** (2007), 67–90.

46. Greenbaum, A. Z. & Sivaramakrishnan, A. In-focus wavefront sensing using non-redundant mask-induced pupil diversity. *Optics Express* **24,** 15506–15521 (2016).

47. McElwain, M. W. *et al.* The James Webb Space Telescope Mission: Optical Telescope Element Design, Development, and Performance. *Publications of the Astronomical Society of the Pacific* **135,** 058001 (2023).

48. Löfdahl, M. G. & Scharmer, G. Wavefront sensing and image restoration from focused and defocused solar images. *Astronomy and Astrophysics Suppl., Vol. 107, p. 243-264 (1994)* **107,** 243–264 (1994).

49. Paxman, R. G., Seldin, J. H., Lofdahl, M. G., Scharmer, G. B. & Keller, C. U. *Evaluation of phase-diversity techniques for solar-image restoration* tech. rep. (1995).

50. Hirzberger, J., Feller, A., Riethmüller, T., Gandorfer, A. & Solanki, S. Performance validation of phase diversity image reconstruction techniques. *Astronomy & Astrophysics* **529,** A132 (2011).

51. Hayes, T. L. Biological Low-Voltage Scanning Electron Microscopy. Edited by Heide Schatten and James B. Pawley. Springer Science Business Media, LLC, New York; 2008, 317 pages. ISBN 978-0-387-72970-1; e-ISBN 978-0-387-72972-5. *Microscopy and Microanalysis* **15,** 564–566 (2009).

52. Goodman, J. W. *Introduction to Fourier optics* (Roberts and Company publishers, 2005).

53. Bonnans, J.-F., Gilbert, J. C., Lemaréchal, C. & Sagastizábal, C. A. *Numerical optimization: theoretical and practical aspects* (Springer Science & Business Media, 2006).

54. Brady, G. R. & Fienup, J. R. Nonlinear optimization algorithm for retrieving the full complex pupil function. *Opt. Express* **14,** 474–486. `https://opg.optica.org/oe/abstract.cfm?URI=oe-14-2-474` (Jan. 2006).

55. Kormacheva, M. *Cluster-based milling method for large-field-of-view volume electron microscopy* PhD thesis (lmu, 2023).

56. Song, K., Feng, Z. & Helmstaedter, M. High-contrast en bloc staining of mouse whole-brain and human brain samples for EM-based connectomics. *Nature Methods,* 1–5 (2023).

57. Clark, A. *et al.* Pillow (pil fork) documentation. *readthedocs* (2015).

Acknowledgement

Winfried has this quality that seeps into your head and sits there, telling you almost constantly how loose a grasp you have on your subject. The learning curve is shallow, with a very small gradient day-to-day, easily overlooked. But the trend is always upward, and you stop feeling so stupid one day. I am so happy I could keep up with the embarrassing years because the reward in my head seems invaluable. Thank you Winfried. You taught me science.

My esteemed colleagues, who became my close friends by the end - Maria Kormacheva and Alexandra Rother. It seemed the frustration truly brought us together which quickly branched out to the topics of the world. I will miss our Denk meditation nights. Thank you, Maria, for your help with the gun, and Alexandra, for your help on the microtome. Thank you, Jonas Hemesath, for helping me with the integrated chip samples.

Franz Rieger, Laura Bauer, Hashir Ahmad, Csilla Pataki, and Carlos Aguilar (and also Jonas, Maria, and Alex) - thank you for all the laughs.

Thank you, Joergen Kornfeld, for writing such comprehensive papers on connectomics. I used them a lot to understand the challenges faced by the field.

Thank you, Barbara Strasser, for being such a warm presence in the office, for helping me with the scheduling so many times, and for the lively chats. Thank you, Juergen Tritthardt, for dropping me home so many times and for the help with the SEMs. Thank you, Petra Winterholler, for introducing me to the microscopes.

Thank you mummy and papa, for giving me a joyous upbringing and all the love and care. And also the money, haha. I am not thanking my husband here because he told me not to and that it wouldn’t be enough anyway.