Walsh Mode Based Neural Network for Adaptive Optics in Two-Photon Microscopy

Yuyao Xiao

Hertford College

University of Oxford

Supervisor: Prof. Martin Booth

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2025

This thesis is dedicated to Yingshi Ji and Kai Xiao for their love and support.

Acknowledgements

The work presented in this thesis would not have been possible without the support of many people.

I am deeply grateful to Prof. Martin Booth for his intelligence, guidance, supervision, and encouragement throughout my doctoral studies.

I also wish to thank Dr. Qi Hu for her dedicated support, without whom I could not have accomplished this work.

I am also thankful to Dr. Jingyu Wang for his technical assistance in microscopy.

Finally, I am grateful to Dr. Raphael Turcotte and Dr. Jiahe Cui for their support at the beginning of my doctoral journey.

Statement of Originality

I hereby declare that except where specific references to the work of others are given, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except in the following:

1. The two-photon microscope in Chapter 6 was originally designed and built by Dr. Raphael Turcotte. The system was later modified and aligned by me.

2. The two-photon microscope in Chapter 9 was originally designed and built by Dr. Jingyu Wang, and later aligned by Dr. Jingyu Wang and Dr. Qi Hu. The system aberration correction software was designed by Biwei Zhang.

Contents

Abstract

Microscopy is a vital tool in biomedical research. In deep tissue two-photon imaging, scattering induces severe wavefront distortions, which degrade image quality. Adaptive optics is used to restore image quality; however, conventional continuous modes are insufficient for compensating complex wavefronts.

This thesis introduces neural-network–based sensorless adaptive optics methods using pixelated Walsh modes for complex wavefront compensation. These methods are designed to be robust and more efficient than conventional algorithms in improving image quality. The neural networks exploit the unique geometrical properties of Walsh modes and, in the machine-learning-assisted wavefront-sensorless adaptive optics control method (MLAO), also utilize aberration structure information extracted from extended images. Both approaches can outperform the conventional $2N + 1$ algorithm and the sequential $3N$ algorithm in efficiency for image quality improvement while maintaining robust performance under various conditions.

Neural-network-based sensorless adaptive optics methods using Walsh modes show significant potential for deep biomedical imaging. Their high efficiency reduces aberration correction time and specimen exposure while maintaining high image quality, and their robust performance ensures versatility across diverse imaging conditions. These results indicate that combining Walsh modes with neural networks provides a powerful framework

for advancing sensorless adaptive optics in complex wavefront compensation.

Chapter 1

Introduction

Microscopy is an important tool in many research areas for revealing microscale and even nanoscale structural and functional properties, particularly in biomedical imaging. Two-photon microscopy, owing to its deep imaging depth, three-dimensional specimen reconstruction ability, high signal-to-noise ratio, and low background emission, is widely used in biomedical applications. However, imaging biological samples with microscopes often suffers from aberrations and scattering, which degrade image quality, reduce signal levels, and limit imaging depth. Adaptive optics (AO) has been introduced as an aberration correction method to restore microscopy image quality, and it has been extensively developed over the past few decades [1–10]. Many optical components, algorithms, and strategies have been developed in the context of AO. Nevertheless, as the light beam propagates deeper into inhomogeneous tissue, wavefront distortion becomes more severe due to scattering, which limits the effectiveness of conventional continuous Zernike polynomials for aberration compensation. In this context, pixelated Walsh

modes are introduced for complex wavefront compensation. Walsh modes and closely related Hadamard modes have been widely studied, particularly for scattering compensation [11–23].

The unique mathematical properties of Walsh modes present both opportunities and challenges in their application. Multiple Walsh modes can generate an image-quality metric space with a lattice-like structure, where the fundamental unit is known as a Voronoi cell. These mathematical properties have remained largely unexplored. To exploit them, neural networks are adopted to handle this high-dimensional, lattice-like metric space—an approach that has not been designed before.

1.1 Research Contributions

The aim of this thesis is to develop a neural-network–based sensorless AO method capable of compensating discontinuous wavefronts, while being robust and more efficient than conventionally used algorithms.

Neural networks solely based on the lattice-like Strehl ratio metric space were constructed, trained, simplified, and optimized. Such a neural network is capable of surpassing both the conventional $2N + 1$ algorithm [24] and the sequential $3N$ algorithm in efficiency, while maintaining high robustness.

To both exploit the mathematical properties of Walsh modes and fully utilize the aberration structure information contained in images, a machine-learning-assisted wavefront-sensorless adaptive optics method (MLAO) was constructed and trained. A scheme for analyzing MLAO was also established, aiming to improve its performance

and provide guidance for future research. The MLAO can outperform the conventional algorithms in efficiency while maintaining high robustness.

Both simulations and experiments were conducted to validate the performance of the proposed neural networks.

The proposed neural-network–based sensorless AO methods utilizing Walsh modes show promise for deep-imaging scattering compensation. This approach may contribute to the future design of AO systems that integrate Walsh-related modes with neural networks.

Chapter 2

Background

2.1 A Brief History of Optical Microscopes

The optical microscope has been an essential instrument across biology, medicine, materials science, and related fields, due to its ability to reveal structural and functional properties at the microscale and even nanoscale.

This section provides a concise overview of the development of optical microscopes. It discusses the motivation for designing such systems, their key components, operating principles, and the advantages and limitations associated with their use.

2.1.1 Optical Compound Microscope

The optical compound microscope is generally credited to the Dutch lens maker Hans Janssen and his son Zacharias, who developed it around 1595 [25, 26]. A significant discovery enabled by the optical compound microscope was the existence of microor-

ganisms. It was first made by Robert Hooke and Antoni van Leeuwenhoek between 1665 and approximately 1678. In 1665, Hooke published *Micrographia* [27], the first illustrated book on microscopy, which highlighted the importance of optical compound microscopes. He was the first to publicly describe a microorganism, the microscopic fungus *Mucor*, and he coined the word 'cell' while examining the tiny rooms in cork. He also refined and described the microscope designs, including improvements in mounting and illumination. Antoni van Leeuwenhoek was the first to observe and describe bacteria and protozoa. He was able to grind lenses to achieve about $300\times$ magnification and 1.4 µm resolution. He is often regarded as the 'first of the microbe hunters' [25, 26, 28].

Figure 2.1 illustrates the principle of the optical compound microscope. It typically consists of two lenses: an objective lens and an eyepiece lens. The objective lens forms an enlarged, inverted, real image of a small object. The eyepiece lens then produces a larger, inverted, virtual image, which enters the observer's eye. The total magnification of the optical compound microscope is the product of the magnifications of the individual lenses. In practice, objects can typically be magnified by $1000 - 1500\times$ [29–32].

Optical compound microscopes offer several advantages, including simplicity, compactness, and ease of use [29]. However, they also present some disadvantages. In this wide-field microscopic technique, control over the image field of view (FOV), brightness (total image intensity), magnification, and contrast depends strongly on the optical components and illumination strategy. Additionally, most observations are based on qualitative microscopic images rather than quantitative measurements [33, 34]. To address these limitations, J.Z. Young and F. Roberts presented the idea of a 'flying-spot

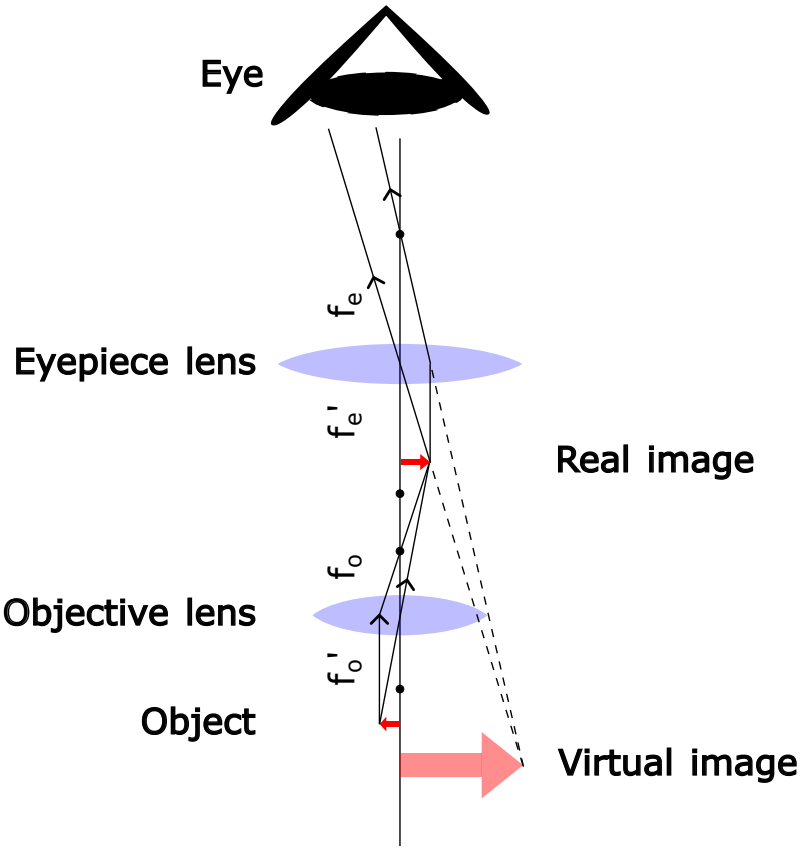


Figure 2.1: A schematic diagram of an optical compound microscope. It contains an objective lens (focal length $f = f_o = f'_o$) and an eyepiece lens (focal length $f = f_e = f'_e$). The two focal points do not intersect. An object (the small red arrow) forms an inverted, enlarged real image (the larger red arrow) by the objective lens. Eventually, the light enters the eye, which forms an inverted, enlarged virtual image (the large pink arrow).

microscope' in 1951 [33]. This was the first scanning optical microscope, which will be introduced in Section 2.1.2.

Microscope Objective

The microscope objective is probably the most important component of an optical microscope, as it largely determines the image quality, resolution, and magnification.

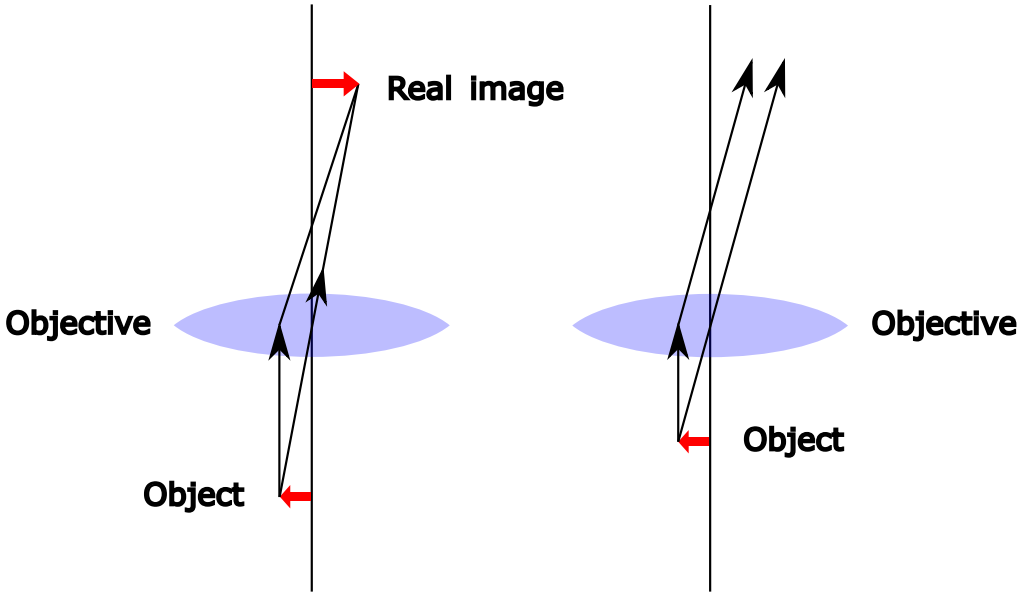


Figure 2.2: Two types of microscope objective designs: (a) finite conjugate and (b) infinite conjugate. For clarity, the diagram represents the compound objective lens as a single lens. In the finite conjugate design, the objective forms a real image at a finite distance. In the infinite conjugate design, the objective forms a real image at infinity.

Modern objectives consist of a series of internal lenses, rather than a single lens as in early designs, to achieve high imaging quality and flatness of the field [35]. Objective designs are generally classified into two main types: finite conjugate and infinite conjugate [36].

In the case of a finite conjugate objective, the real image of the object is formed at a finite distance. The most significant benefits of this design are cost efficiency and the ease of adjusting magnification by swapping objective lenses. This design is employed by most of the optical compound microscopes.

The alternative design is the infinite conjugate objective, in which the real image of

the object is formed at infinity. Although the tube lens needs to be matched when using this type of objective, it allows the integration of some special optical components, such as filters, polarizers, beamsplitters, and adaptive optics (AO) [35]. AO is often placed conjugate to the infinite conjugate objective's back aperture. This is to ensure a uniform aberration correction across the whole FOV. By allowing such flexibility, the infinite conjugate objective greatly extends the functionality and applications of optical microscopes and is therefore widely adopted in scanning optical systems. Illustrations of both designs are shown in Figure 2.2.

2.1.2 Scanning Optical Microscope

A scanning optical microscope illuminates a small spot on the sample, while a photoreceptor collects the emitted signals. By scanning the spot across the sample and recording the corresponding positions, the detected signals are digitized and reconstructed as a two-dimensional (2D) image for display. In their original work, J.Z. Young and F. Roberts outlined several advantages of using such a microscope design: it allows the control of imaging FOV, brightness, magnification, and contrast, which facilitates the visualization of fine details; it improves resolution by eliminating interference of scattered light from adjacent structures; it enables more straightforward quantitative analysis; and it suggests the possibility to probe the samples with invisible light [33, 34].

2.1.3 Fluorescence Microscope

The fluorescence phenomenon has been known for a long time [37]. In 1852, George Gabriel Stokes observed that the emitted fluorescence wavelengths were always longer than those of the excitation light [38]. He also introduced the term 'fluorescence'. In 1873, Ernst Abbe proposed the limitations of using transmitted light in microscopes, and in 1874, Herman von Helmholtz suggested that fine structures could be more easily distinguished if they emitted light. These insights made fluorescence a favorable method for microscopy [39].

In 1911, Oskar Heimstädt and Hans Moritz Lehmann developed the first fluorescence microscope as the outgrowth of the ultraviolet microscope. The early stage fluorescence microscopes, however, were limited by reliance on object autofluorescence, use of dark-field condensers, and transmitted illumination [40]. These challenges were addressed by Max Haitinger, who introduced the use of exogenous fluorescent dyes. The development of this 'secondary fluorescence' method played a vital role in the design of the epi-fluorescence microscope, which enabled light-efficient biomedical imaging, even in living opaque samples.

The epi-fluorescence microscope was designed by Philipp Ellinger and August Hirt in 1929 [40]. In this configuration, the light source and objective are positioned on the same side of the sample. Excitation light passes through the objective to illuminate the sample, and the emitted fluorescence is collected through the same objective [40]. An important innovation in this design is the dichroic beamsplitter (DBS), which allows efficient separation of excitation and emission light. An illustration of this epi-fluorescence

microscope excitation and emission pathway is shown in Figure 2.3.

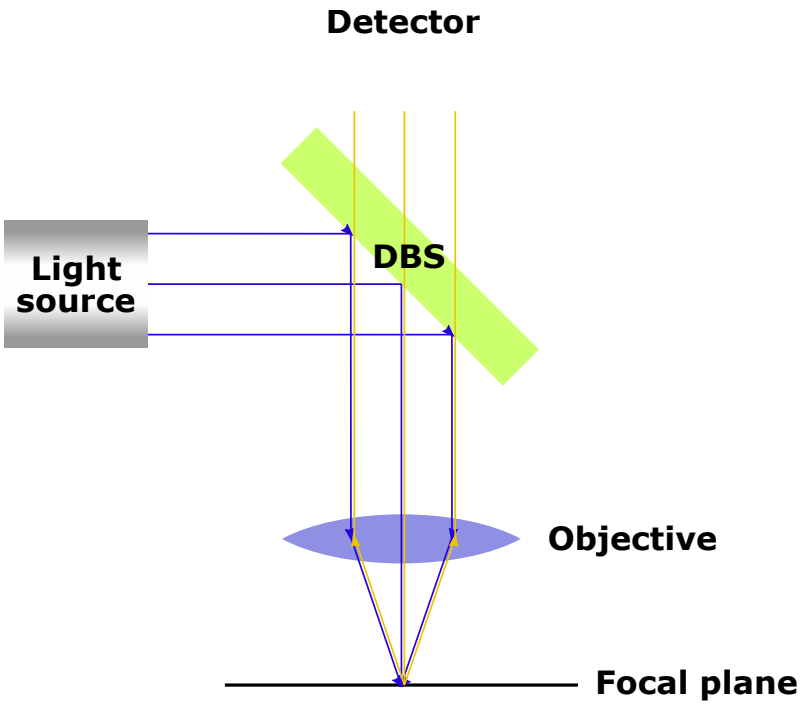


Figure 2.3: Epi-fluorescence microscope excitation and emission pathway. Excitation light is reflected by the DBS, focused through the objective, and excites fluorescence at the focal plane. Emitted fluorescence is collected by the same objective, passes through the DBS, and is eventually detected.

The evolution of fluorescent labelling methods has greatly accelerated the development of fluorescence microscopy as well. Organic chemists have designed a wide range of fluorescent probes for diverse biological applications [41]. Since the 1960s, the green fluorescent protein (GFP) was discovered and subsequently developed [42]. The cloning of the GFP gene [43, 44] and the development of a family of related proteins with distinct spectra [45] made them indispensable tools for genetically tagging proteins in cells of all living organisms *in vivo* [42, 46]. The fluorescent proteins, organic dyes, and secondary

antibodies are now widely employed for protein and structure localization in cells and tissues, as well as for monitoring rapid dynamics in living cells [47].

Consequently, owing to their high signal-to-noise ratio (SNR), high sensitivity, and high specificity, fluorescence microscopes have become reliable and popular tools in biological research.

2.1.4 Confocal Microscope

The confocal microscope was invented by Marvin Minsky in 1955 [48, 49]. Its basic principle is to position the detection pinhole conjugate to the focal spot in the sample, thereby rejecting out-of-focus light before it reaches the photoreceptor. The first confocal microscope utilized the brightest light source available in the 1950s—a carbon arc [49]. Then in 1960, Theodore Maiman invented the first optical laser, marking the beginning of a new era of microscopy [50]. In 1969, Paul Davidovits and M. David Egger designed and built the confocal laser scanning microscope (CLSM) [51]. By incorporating a laser light source, which is highly monochromatic, collimated, and bright, the CLSM enables the formation of extremely small, high-intensity focal spots in tissue [34, 52, 53].

Figure 2.4 presents a schematic diagram of a confocal microscope. Conjugation of the pinhole with the focal point effectively eliminates out-of-focus light.

The CLSM offers several advantages, including increased imaging resolution in deep tissue and the ability to perform three-dimensional (3D) reconstruction by optical sectioning [47]. Although the CLSM technique has improved the image quality and pen-

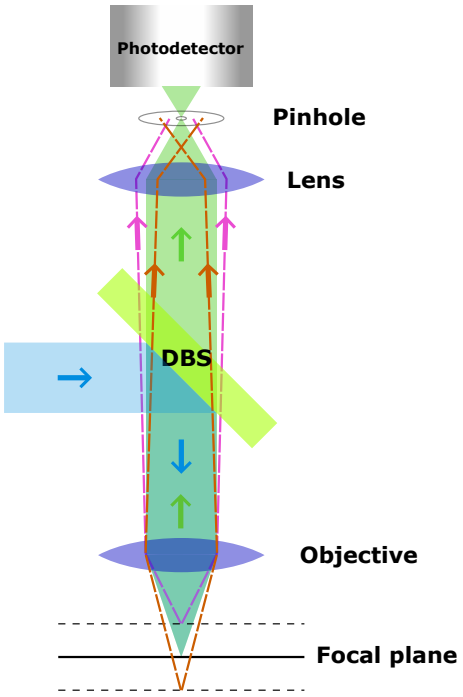


Figure 2.4: Schematic of a confocal microscope. All light-propagating directions are annotated with corresponding coloured arrows. The excitation beam is reflected by the DBS and focused into the sample, illustrated as a blue transparent light path. Emitted fluorescence, with a longer wavelength than the excitation laser, passes through the DBS and is depicted as a green transparent light path. The pinhole in front of the photodetector is conjugate to the focal spot, allowing in-focus fluorescence to reach the detector. Out-of-focus emission, shown as pink and red dotted lines, is blocked by the pinhole and not detected.

etration depth, when imaging at greater depths (beyond 100 µm), the imaging process becomes increasingly challenging. Tissue diffraction, refraction, and especially scattering can all reduce those qualities [54]. In this case, the idea of combining nonlinear optical phenomena, such as high-order harmonic generation and multiphoton fluorescence excitation, with scanning optical microscopy was proposed and demonstrated by Sheppard et al. [55–58]. Building on this foundation, Winfried Denk and colleagues

made two-photon laser scanning fluorescence microscopy (2PLSM) technically feasible in 1990 [59].

2.1.5 Two-Photon Laser Scanning Fluorescence Microscope

In general, the working principle of multiphoton microscopy is based on the 'localized nonlinear light-matter interaction' that generates signals [54]. 2PLSM relies on the process of two-photon absorption (2PA), which was theoretically analyzed by Göppert-Mayer in 1931 [60] and experimentally demonstrated by W. Kaiser and C. G. B. Garrett in 1961 [61]. In the 2PA process, two photons of identical energy are absorbed nearly simultaneously (less than 0.5 fs) by one molecule in its ground state, exciting it to a higher energy state. The excited molecule emits the energy as fluorescence and returns to the ground state [54, 60]. Figure 2.5 compares 2PA and one-photon absorption (1PA), in which a single, high-energy photon excites the molecule and produces the same emission. A key distinction between 2PA and 1PA is that 2PA depends quadratically on the incident light intensity, whereas 1PA depends linearly. This quadratic dependence leads to several advantages in utilizing this 2PA process.

The main advantages of 2PLSM are as follows. First, it typically uses femtosecond near-infrared lasers with wavelengths from 700 nm to 1000 nm. Longer wavelengths are less susceptible to tissue scattering and hence allow deeper penetration [54, 60, 62].

Second, because 2PA depends quadratically on the incident laser intensity, regions outside the focal volume receive insufficient intensity to induce significant fluorescence. The spatial confinement to the perifocal region of the 2PA, which is referred to as lo-

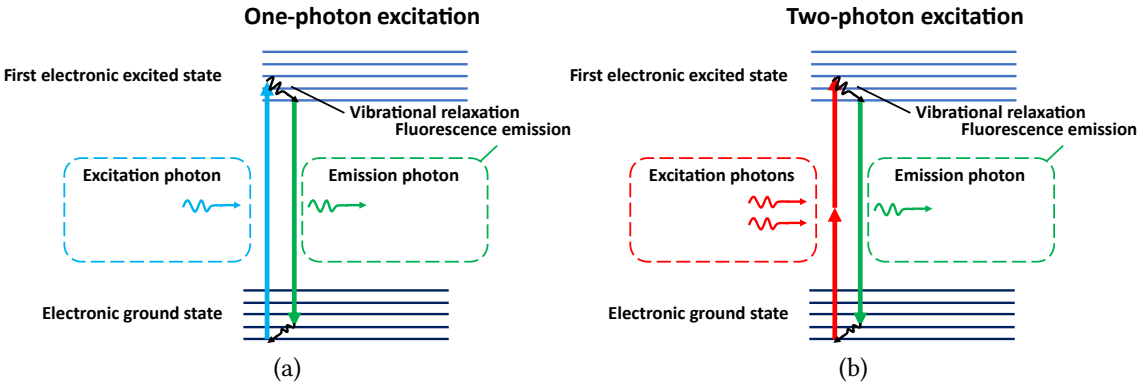


Figure 2.5: Jablonski diagrams of (a) 1PA and emission processes, and (b) 2PA and emission processes. (a) In 1PA, a single photon is absorbed, exciting the molecule to the first electronic excited state. Through a vibrational process, the molecule relaxes to the lowest energy level of that state. Fluorescence emission then returns the molecule to the ground state, releasing a single photon. (b) In 2PA, two photons of the same energy are absorbed nearly simultaneously via a temporary intermediate state, exciting the molecule to the first electronic excited state. After vibrational relaxation, fluorescence emission returns the molecule to the ground state, releasing a single photon.

calized nonlinear light-matter interaction, enables 2PLSM to achieve optical sectioning without a pinhole and greatly reduces out-of-focus photobleaching. In contrast, 1PA produces emission throughout the entire excitation light cone. A comparison is shown in Figure 2.6. The absence of 2PA in the out-of-focus region reduces the background emission and improves the SNR.

A schematic diagram of the 2PLSM is shown in Figure 2.7. The excitation laser is typically chosen in the near-infrared range to match the absorption properties of commonly used fluorescent labels. To enhance the efficiency of 2PA, which requires a high photon density in the temporal domain, a pulsed laser with a pulse width of approximately 100 fs to 200 fs and high peak intensity is used, usually provided by a mode-locked laser

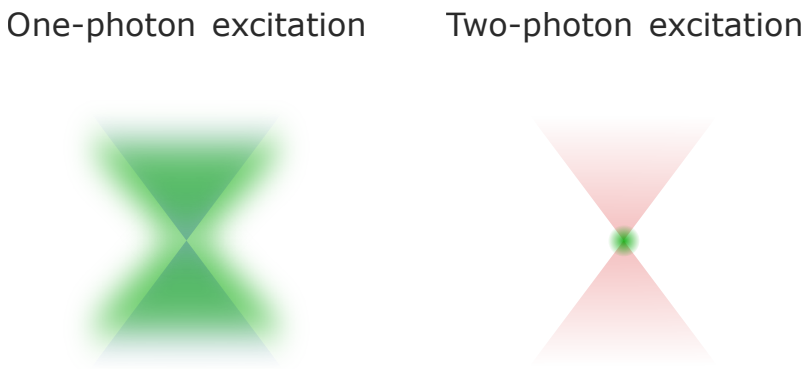


Figure 2.6: Comparison of linear and nonlinear emission volumes. The left cone shows one-photon excitation, where emission occurs throughout the entire light cone (the green cone). The right cone shows two-photon excitation, where near-infrared light excites fluorescence only within the focal region (the green spot).

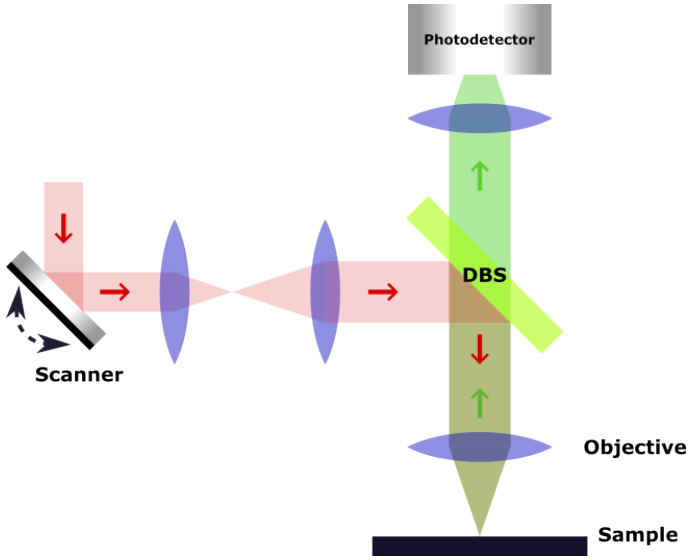


Figure 2.7: A schematic diagram of a 2PLSM. All light-propagating directions are annotated with corresponding coloured arrows. The laser beam is modulated by the scanner, reflected by the DBS, and focused into the tissue, illustrated by the red transparent light path. Two-photon emission occurs only in the focal region. The emitted fluorescence is then collected by the objective, passes through the DBS, and is detected by the photodetector. The emission path is illustrated by the green transparent light path.

[54]. A high numerical aperture (NA) objective is also employed to increase the spatial photon density (see Section 2.2.2 for the NA definition). The laser beam is focused into the tissue, where 2PA occurs exclusively in the focal region. The resulting fluorescence, emitted at a shorter wavelength than the excitation light, passes through the DBS and is collected by the photodetector.

2.2 Numerical Modeling of Optical Systems

In this section, we derive basic numerical models. These models are essential for microscopy method design, as they provide fundamental theoretical understanding prior to hardware implementation.

2.2.1 Models of Light

Light is electromagnetic radiation, which can be described using three models: ray, wave, or particle. They are modeled with different limitations and applied in different situations.

Light can be modelled as rays. Rays represent straight-line paths along which energy is transmitted from the light source to the energy receptor. When interrupted by an opaque obstacle, no energy reaches the receptor [63]. This model forms the basis of geometrical optics [64, 65].

The wave model, derived from Maxwell's equations, describes light as a transverse electromagnetic wave. It is essential for explaining phenomena such as interference,

diffraction, and polarization [63, 66–68]. This model serves as the primary framework in this thesis.

However, classical wave theory alone could not explain some phenomena, including nonlinear optical effects such as 2PA. To address this, the particle model of light was introduced, treating light as discrete energy packets—photons. This model highlights the particle-like nature of light and underpins the principle of wave–particle duality. Within this framework, light–matter interactions involving quantized energy transfer can be successfully explained [63, 69].

Wave Model of Light

In a linear, isotropic, homogeneous, and nondispersive dielectric medium, assuming no free charges or currents, Maxwell's equations reduce to the following form (expressed in SI units) [66]:

$$\nabla \cdot \epsilon\vec{E} = 0 \quad \nabla \cdot \mu\vec{H} = 0$$
$$\nabla \times \vec{E} = -\mu\frac{\partial\vec{H}}{\partial t} \quad \nabla \times \vec{H} = \epsilon\frac{\partial\vec{E}}{\partial t}$$

In the above equations, $\vec{E}$ and $\vec{H}$ denote the electric and magnetic field vectors, respectively. Both are functions of time t and position $\vec{r}$. The parameters μ and ϵ represent the permeability and permittivity of the propagation medium. The operators $\times$ and $\cdot$ denote the vector cross and dot products, respectively. The differential operator is given by:

$$\nabla = \frac{\partial}{\partial x}\vec{x} + \frac{\partial}{\partial y}\vec{y} + \frac{\partial}{\partial z}\vec{z}$$

Here, $\vec{x}$, $\vec{y}$, and $\vec{z}$ are the unit vectors along the x, y, and z directions. The position vector

is defined as:

$$\vec{r} = x \cdot \vec{x} + y \cdot \vec{y} + z \cdot \vec{z}$$

Then we do:

$$\nabla \times (\nabla \times \vec{E}) = \nabla \times (-\mu \frac{\partial \vec{H}}{\partial t})$$

And consider [66]:

$$\nabla \times (\nabla \times \vec{E}) = \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E}$$

We can derive the wave equation for the electric field in a medium:

$$\begin{aligned} \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} &= -\mu \nabla \times \frac{\partial \vec{H}}{\partial t} \\ -\nabla^2 \vec{E} &= -\mu \frac{\partial (\nabla \times \vec{H})}{\partial t} \\ -\nabla^2 \vec{E} &= -\mu\epsilon \frac{\partial}{\partial t}(\frac{\partial \vec{E}}{\partial t}) \\ -\nabla^2 \vec{E} &= -\mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} \end{aligned} \tag{2.1}$$

We define a wave propagating along the z-axis. The equation is:

$$-\frac{\partial^2 \vec{E}}{\partial z^2} = -\mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} \tag{2.2}$$

A harmonic (monochromatic) wave solution to the above equation can be expressed as:

$$\vec{E}(z,t) = \vec{E_0} e^{j(kz - \omega t)} \tag{2.3}$$

In Equation 2.3, ω is the angular frequency, and $\vec{E_0}$ is the original electric field. The wave number k is defined as:

$$k = \frac{2\pi}{\lambda} \tag{2.4}$$

By combining Equation 2.2 and Equation 2.3, we obtain:

$$k^2 = \omega^2 \mu \epsilon \tag{2.5}$$

Additionally, the refractive index of the medium can be calculated as:

$$n = \sqrt{\frac{\mu\epsilon}{\mu_0 \epsilon_0}} \tag{2.6}$$

Here, μ_0 and ϵ_0 denote the permeability and permittivity of free space, respectively. The wave number k can then be calculated as:

$$k = \omega n \sqrt{\mu_0 \epsilon_0} \tag{2.7}$$

2.2.2 Diffraction and Fourier Optics

Diffraction and Approximations

Diffraction is the deviation of light from rectilinear propagation when it encounters an obstruction [70]. The first theory explaining diffraction was proposed by Christian Huygens in 1678. He stated that if each point on a wavefront (see Section 2.2.3 for definition) is considered a source of a secondary spherical wave, then the wavefront at a later instant is constructed as the 'envelope' of all these secondary spherical waves [66]. The ideas of Huygens and Thomas Young (who brought the concept of interference) were mathematically merged by Augustin Jean Fresnel for near-field diffraction in 1818. Then Joseph von Fraunhofer simplified these formulas for far-field diffraction. In 1882, Gustav Kirchhoff derived a rigorous diffraction formula directly from the wave equation, thereby justifying the Huygens-Fresnel principle [66].

We assume that the primary wavefront lies in the plane $z = 0$, with its normal parallel to the z-axis [71]:

$$E(x_2, y_2) = \frac{-j}{\lambda} \int\int \left(\frac{1 + cos\theta}{2}\right) \frac{e^{jkr}}{r} E(x_1, y_1) dx_1 dy_1 \tag{2.8}$$
$$r = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + z^2}$$

λ is the wavelength, k is the wave number, r is the distance between a point (x_2, y_2, z) on the secondary wavefront and a point $(x_1, y_1, 0)$ on the primary wavefront, and θ is the angle between the normal to the primary wavefront and the vector $\vec{r}$. $\left(\frac{1 + cos\theta}{2}\right)$ is the obliquity factor given by Kirchhoff.

If $|x_2 - x_1|$ and $|y_2 - y_1|$ are much smaller than z, a Taylor expansion gives:

$$r \approx z + \frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2z}$$

In this case, the distance r in the denominator of $\frac{e^{jkr}}{r}$ can be approximated as z, and the angle θ is sufficiently small so that the obliquity factor can be approximated as 1. Then the Equation 2.8 can be written as:

$$E(x_2, y_2) \approx \frac{-je^{jkz}}{\lambda z} \int\int e^{jk[\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{2z}]} E(x_1, y_1) dx_1 dy_1 \tag{2.9}$$

This is the Fresnel approximation.

If we further approximate that:

$$z \gg \frac{k}{2} \max(x_1^2 + y_1^2) \tag{2.10}$$

$\max(x_1^2 + y_1^2)$ represents the square of the incident aperture radius (assuming a circular aperture). Equation 2.10 describes the far-field condition, which holds when the propagation distance z is much greater than the square of the incident aperture radius. Under

this approximation:

$$E(x_2, y_2) \approx \frac{-je^{jkz}}{\lambda z} e^{jk(\frac{x_2^2+y_2^2}{2z})} \int\int e^{-jk(\frac{x_1x_2+y_1y_2}{z})} E(x_1, y_1) dx_1 dy_1 \tag{2.11}$$

We can have the Fraunhofer approximation.

If we define $u = \frac{kx_2}{z}$ and $v = \frac{ky_2}{z}$, we obtain:

$$E(u, v) \approx \frac{-je^{jkz}}{\lambda z} e^{\frac{jz(u^2+v^2)}{2k}} \int\int e^{-j(x_1u+y_1v)} E(x_1, y_1) dx_1 dy_1 \tag{2.12}$$

The above Fraunhofer approximation takes the form of a Fourier transform. It can be written as:

$$E(u, v) \propto \int\int e^{-j(x_1u+y_1v)} E(x_1, y_1) dx_1 dy_1 \propto \mathcal{FT}\left[E(x_1, y_1)\right] \tag{2.13}$$

where $\mathcal{FT}$ denotes the Fourier transform operation. This approximation is valid only when the far-field condition is satisfied. The Fresnel number F is defined based on Equation 2.10:

$$F = \frac{k}{2z} \max(x_1^2 + y_1^2) \tag{2.14}$$

The approximation of Equation 2.11 is valid when $F \gg 1$, corresponding to the far-field condition, whereas Equation 2.9 applies when $F \approx 1$ [71].

Thin Lens Model

We assume a thin, double-convex (converging) lens with spherical surfaces at $z = 0$. The coordinates on the lens pupil plane are (x_1, y_1). The phase transmission is given by [71]:

$$t_{\text{phase}}(x_1, y_1) = e^{-j\frac{k}{2f}(x_1^2+y_1^2)} \tag{2.15}$$

f is the focal length of the thin lens. And the amplitude transmission is given by:

$$t_{\text{amp}}(x_1, y_1) = \begin{cases} 1 & \sqrt{x_1^2 + y_1^2} \leq 1 \\ 0 & \sqrt{x_1^2 + y_1^2} > 1 \end{cases} \tag{2.16}$$

By combining Equation 2.9, Equation 2.15, and Equation 2.16, and applying the paraxial approximation (assuming waves travel close to the optical axis at small angles) with $z = f$, we obtain:

$$\begin{aligned} E(x_2, y_2) &\approx \frac{-je^{jkf}}{\lambda f} e^{jk(\frac{x_2^2+y_2^2}{2f})} \int\int e^{-jk(\frac{x_1x_2+y_1y_2}{f})} E(x_1, y_1) t_{\text{amp}}(x_1, y_1) dx_1 dy_1 \\ &\propto \int\int e^{-jk(\frac{x_1x_2+y_1y_2}{f})} E(x_1, y_1) t_{\text{amp}}(x_1, y_1) dx_1 dy_1 \end{aligned} \tag{2.17}$$

The above equation also takes the form of a Fourier transform. By defining $u = \dfrac{kx_2}{f}$ and $v = \dfrac{ky_2}{f}$, we obtain:

$$E(u, v) \propto \mathcal{FT}\left[E(x_1, y_1) t_{\text{amp}}(x_1, y_1)\right] \tag{2.18}$$

This thin-lens model is valid only under the paraxial approximation and for small numerical aperture (NA) conditions (NA < 0.5) [71]. The NA is defined as:

$$\text{NA} = n \frac{\sqrt{\max(x_1^2 + y_1^2)}}{f} \tag{2.19}$$

n is the medium refractive index.

Although a vectorial lens model exists for diffraction analysis in high NA situations, which provides a full 3D electric field solution at the focal plane, the thin-lens scalar model offers a good approximation when the polarization state is not explicitly considered. This approximation does not significantly affect the results and is therefore adopted throughout this thesis.

2.2.3 Phase Property

By extracting the temporally invariant part of Equation 2.3, we obtain [71]:

$$e^{jkz} = e^{j(\Phi+2\pi m)} = e^{j\Phi}e^{j2\pi m} = e^{j\Phi} \quad m \in \mathbb{Z} \tag{2.20}$$

Φ denotes the spatial phase. In a propagating beam, a set of points where the waves have the same spatial phase Φ of the sinusoid is defined as the wavefront (by Euler's formula, $e^{j\Phi} = \cos(\Phi) + j\sin(\Phi)$) [72]. For a collimated beam, a flat wavefront is usually considered aberration-free.

Given that $k = \omega n\sqrt{\mu_0\epsilon_0}$, when waves in a collimated beam with a flat wavefront propagate through a medium with inhomogeneous refractive index n, the phase becomes:

$$\Phi = kz = \omega n\sqrt{\mu_0\epsilon_0}z \tag{2.21}$$

The angular frequency ω remains constant. Consequently, different portions of the beam acquire different spatial phase shifts, distorting the initially flat wavefront and introducing aberrations. These aberrations degrade image quality, reduce signal levels, and limit penetration depth. Such distortions can be corrected using well-established AO, which is the central focus of this thesis. Further discussion on how distorted wavefronts impair the image quality is provided in Section 2.2.5.

2.2.4 Interference and Coherence

When two waves are combined, their complex amplitudes are superimposed, taking into account the phase difference between them. This phenomenon is known as

interference. If the waves are in phase, the resulting amplitude increases, producing constructive interference; if the waves are out of phase, the resulting amplitude decreases, producing destructive interference [68]. Coherence expresses the potential for two waves to interfere, which can be categorized into two types: spatial coherence and temporal coherence.

Spatial coherence describes the degree of correlation between waves at different points in space, measured at the same time. Similarly, temporal coherence describes the degree of correlation between waves at the same spatial point but at different times. Considering the temporal part of the light wave model (see Equation 2.3), we have: $e^{j\omega t}$, where ω is the angular frequency. From Equation 2.4 and Equation 2.5, it follows that:

$$\omega = \frac{k}{\sqrt{\mu\epsilon}} = \frac{2\pi}{\lambda\sqrt{\mu\epsilon}} \tag{2.22}$$

High temporal coherence corresponds to a narrow frequency spectrum, which also implies a narrow wavelength spectrum, as in the case of monochromatic light.

Although coherence is fundamental in many optical applications, this thesis focuses on incoherent imaging systems. Incoherent imaging systems allow the imaging process to be formulated using linear system theory. This approach includes the two-photon microscope used in this thesis. [73].

2.2.5 Point Spread Function

The point spread function (PSF) is the impulse response of a focused, isoplanatic (the PSFs formed by different points at the focal plane are the same), linear optical system. It corresponds to the diffraction pattern formed in the image plane by a point source at

the focal plane and characterizes its normalized intensity distribution [73]. In an ideal scenario with no aberration, the PSF is diffraction-limited. The PSF is calculated using the diffraction integral described in Section 2.2.2, which is given by the squared modulus of the Fourier transform of the thin lens pupil function [74].

The functional form of the diffraction-limited PSF is [73]:

$$\mathrm{PSF}(r) = \left[2 \frac{J_1\left(\frac{2\pi \mathrm{NA}}{\lambda} r\right)}{\left(\frac{2\pi \mathrm{NA}}{\lambda} r\right)} \right]^2 \tag{2.23}$$

J_1 is the first-order Bessel function of the first kind, λ is the wavelength, NA is the numerical aperture, and r is the radial distance from the center of the image plane. This function is valid under the far-field condition, where Fraunhofer diffraction applies. The resulting intensity distribution is referred to as the Airy pattern. The Airy pattern, which plays a central role in defining the diffraction-limited resolution. Resolution is commonly defined as the minimum separation distance between two point sources at which they can still be distinguished as distinct. It characterizes the ability of the imaging system to resolve structural details. Two major criteria are employed to quantify this distance: the Abbe distance and the Rayleigh criterion. The Abbe distance is:

$$r_{\mathrm{Abbe}} = \frac{\lambda}{2\mathrm{NA}} \tag{2.24}$$

which can be approximated by the full width at half maximum (FWHM) value of the central bright spot of the Airy pattern — the Airy disk.

According to the Rayleigh criterion, the resolution is the distance from the centre of the Airy pattern to the first minimum value:

$$r_{\mathrm{Rayleigh}} \approx \frac{0.61\lambda}{\mathrm{NA}} \tag{2.25}$$

The PSF can be used not only to evaluate the imaging quality of a system but also to model the image formation process. Based on the discussion in Section 2.2.2, in a one-photon system, the PSF is calculated as the squared modulus of the Fourier transform of the objective pupil function. Within the linear system framework, the image (Img) can be simulated as the convolution of the PSF (PSF) with the object function (Obj):

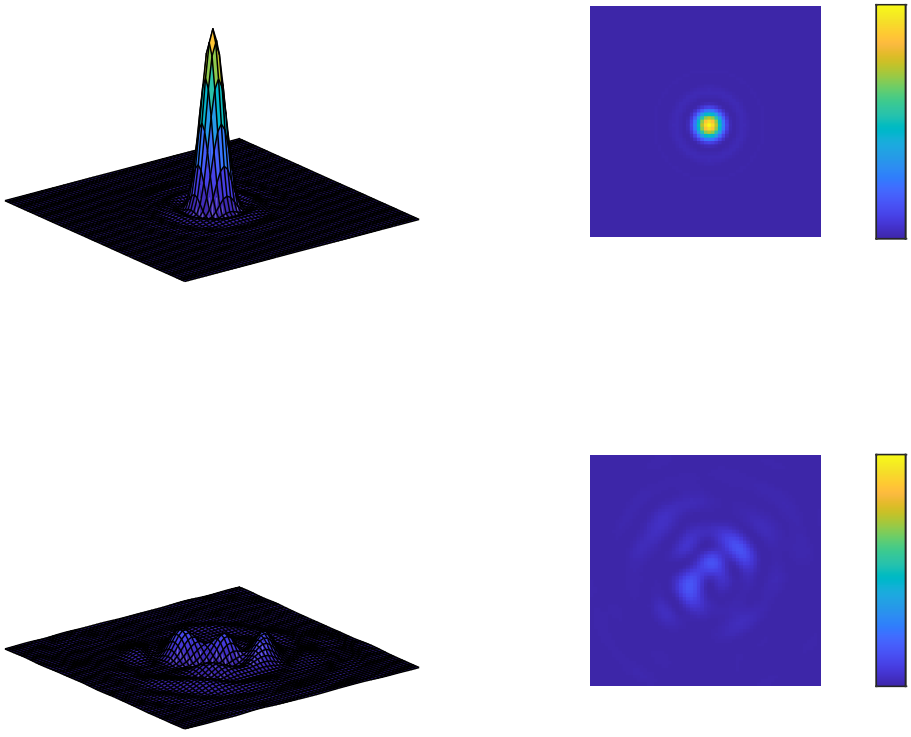
$$PSF \propto |\mathcal{FT}\left[\mathcal{P}(x_1, y_1)\right]|^2 \tag{2.26}$$

$$Img \propto PSF * Obj \tag{2.27}$$

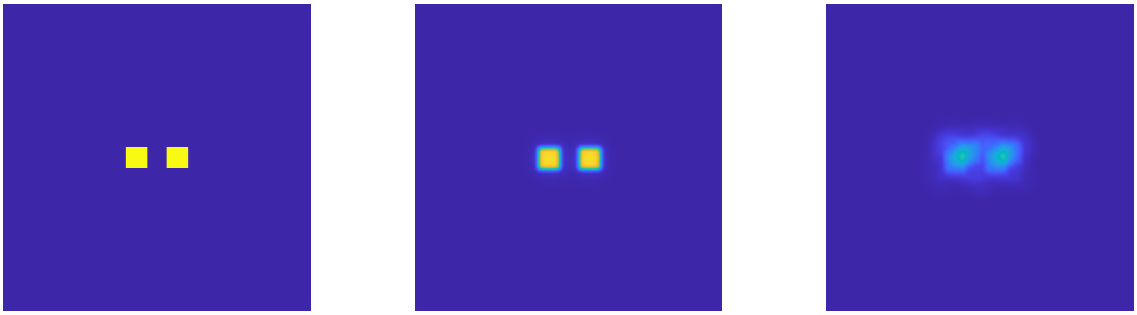
Here, $\mathcal{P}(x_1, y_1)$ denotes the objective pupil function. $*$ represents the convolution operator. Equations 2.26 and 2.27 illustrate how aberrations affect image quality. If the wavefront at the objective pupil is distorted, or if distortion occurs within the imaging sample, the resulting PSF deviates from the diffraction-limited case and becomes aberrated, leading to degraded image quality. For further details, see Section 2.3.1.

In 2PLSM, the two-photon fluorescence intensity depends quadratically on the incident light intensity (see Section 2.1.5). Consequently, when simulating the 2PLSM image formation process, the PSF is computed as the fourth power of the modulus of the Fourier transform of the objective pupil function.

The simulated images of the diffraction-limited and aberrated PSFs are shown in Figure 2.8(a), for a one-photon system. In the diffraction-limited PSF image, the central bright spot corresponds to the Airy disk, and the surrounding dim concentric rings together form the Airy pattern. In the aberrated PSF image, the central spot is dim, and the outer rings are distorted. Figure 2.8(b) illustrates image formation by convolving the



(a)



(b)

Figure 2.8: (a) Top: Diffraction-limited PSF. Bottom: Aberrated PSF. In both surface plots, the height represents intensity, and the color scale also encodes intensity, with blue indicating zero and yellow indicating maximum. (b) Object (left), convolution of the object with the diffraction-limited PSF (middle), and convolution of the object with the aberrated PSF (right).

PSF with the object: when using the diffraction-limited PSF, two point sources remain distinguishable, whereas convolution with the aberrated PSF produces a blurred image in which the points cannot be resolved.

2.3 Adaptive Optics

Optical microscopes are indispensable in biomedical imaging; however, achieving optimal image quality is often hindered by aberrations originating from the sample, the optical system, or other sources. Adaptive optics (AO), originally developed in astronomy to correct atmospheric distortions [75], has been applied to microscopy to compensate for these aberrations.

In this section, the nature of aberrations is introduced and the fundamental concepts of AO are discussed.

2.3.1 Sources and Effects of Phase Aberrations in Microscope Imaging

Aberration can be broadly defined as the deviation of light from its ideal form. In the context of AO, it is typically considered a deviation of the wavefront phase from its ideal form. For a collimated beam, the ideal wavefront is planar. When such a beam is focused by an ideal lens, the rays converge to a single point, and the planar wavefront becomes a spherical cap.

Aberrations can arise from various sources within the optical system, such as sys-

tem misalignment [76], low-quality optical components, poor system design, operation outside design specifications (for example, wavelength or temperature) [77], or sample mounting. Even in the absence of these factors, optical components are never perfect and inherently introduce aberrations.

Specimens can also introduce aberrations. A common example occurs when light is focused through two media with different refractive indices separated by a planar interface, leading to refraction at the interface. The resulting focal shifts and spherical aberrations are the primary consequences, and are common in applications such as the use of immersion objectives with a mismatched immersion medium or cover glasses of incorrect thickness [3]. Numerous research studies have investigated these effects [76, 78–84].

Inhomogeneous refractive indices within the specimen can also induce aberrations. As light propagates through such regions, rays undergo unpredictable diffraction and refraction, broadening the focal spot and distorting the wavefront. This degrades image resolution and distorts structural visualization. It also reduces excitation intensity at the focus, thereby compromising signal generation, lowering optical efficiency, reducing SNR, and limiting penetration depth.

The wavefront distortions caused by refractive index inhomogeneities are complex, with both their magnitude and complexity increasing with imaging depth. These inhomogeneities vary not only among specimens but also across different regions of the same specimen, rendering generalized solutions such as correction collars ineffective. This limitation underscores the necessity of AO in biomedical microscopy for aberration correction.

Physical illustrations of these processes are presented in Figure 2.9.

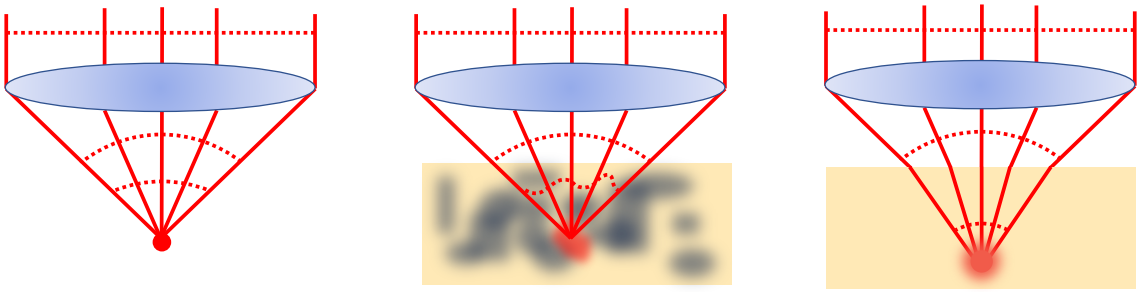


Figure 2.9: Focus and wavefront shapes for different aberration sources. Red lines denote light rays, and red dashed lines represent the wavefront. Left: In the ideal case, all light rays converge to a single point, and the planar wavefront forms a spherical cap. Middle: Imaging through an inhomogeneous specimen (shown in black) broadens the focal spot and distorts the wavefront. Right: A refractive index mismatch between media refracts the light, shifting the focal point and introducing spherical aberrations.

2.3.2 Adaptive Optics Elements

The aim of AO is to restore the wavefront phase to its ideal form. The correction is achieved by introducing equal but opposite phase profiles that cancel the aberrations described in the previous section. This is accomplished by using adaptive elements. Adaptive elements are controlled via programmable instructions to modify their optical properties and generate the desired compensation phase profiles [3, 4, 10, 85–87]. There are two major types of widely used adaptive elements: the deformable mirror (DM) and the spatial light modulator (SLM). In addition to these, a promising new type of adaptive elements, the deformable phase plate (DPP), has been introduced. Each type operates on distinct principles and is suited to specific applications, with respective advantages and limitations.

Deformable Mirror

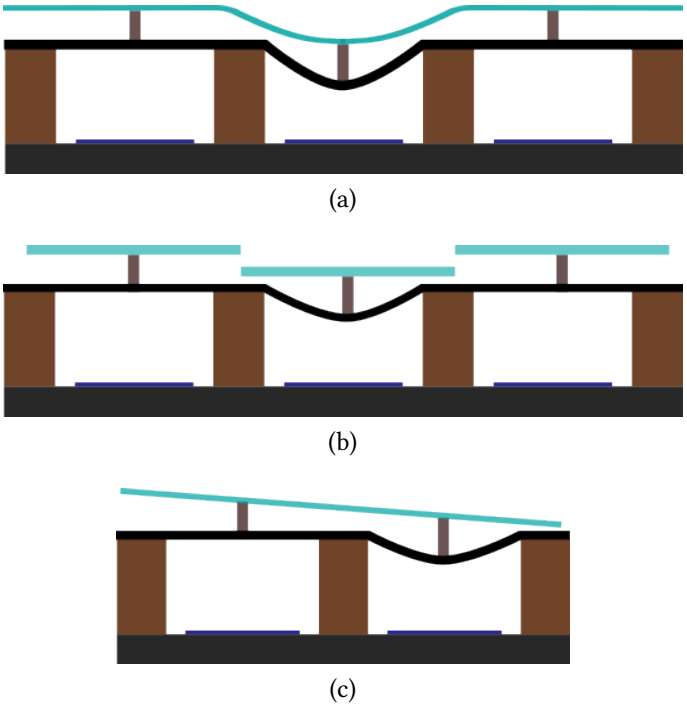


Figure 2.10: Continuous DM (a) and segmented DMs (b) and (c). DMs (a) and (b) modulate phase, whereas DM (c) modulates both phase and direction. Light blue regions indicate reflective mirror surfaces. Dark blue regions represent electrodes that drive the actuator arrays, deforming the surface to generate the desired compensation phase profile.

The DM consists of micromechanical arrays of electronically addressable actuators and reflective surfaces [88, 89]. The overall surface profile can be modified via programmable configurations of the actuators. The actuators are driven by electrostatic, electromagnetic, or piezoelectric forces. Reflected light from the surface is modulated in phase and/or direction [90]. Most DMs have a continuous surface, though segmented ones are also available.

Each reflective surface, with suitable coatings, can reflect over 95% of the incident light. The reflective nature of DMs allows low energy loss, polarization independence, and broadband operation. DMs with tens to hundreds of actuators, coupled with appropriate electronic drivers, can operate at frequencies up to several kilohertz, generally faster than other adaptive elements [91]. Illustrative figures of the DMs are shown in Figure 2.10.

Spatial Light Modulator

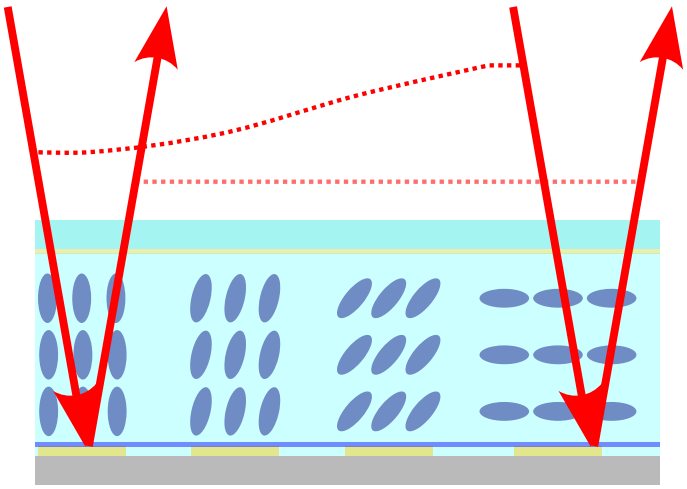


Figure 2.11: Schematic of the SLM. Red arrows represent incident and reflected light rays (shown at an angle since the SLM operates in reflection mode), and red dashed lines indicate the wavefronts before and after modulation. Ellipses denote LC molecules. Incident light passes through the transparent cover glass and electrode layer (light yellow beneath the blue cover glass), then enters the LC layer, where modulation occurs. The refractive index of each pixel is controlled by its corresponding electrode (light yellow rectangles above the gray silicon-based control electronics). The applied voltage decreases from left to right. In the rest state (no applied voltage), LC molecules align parallel to the surfaces; with increasing voltage, they incline progressively toward the surface normal. Reflected light from the mirror coating (dark blue layer above the control electrode array) is further modulated by the LC layer before exiting the device, restoring the distorted wavefront to a planar form.

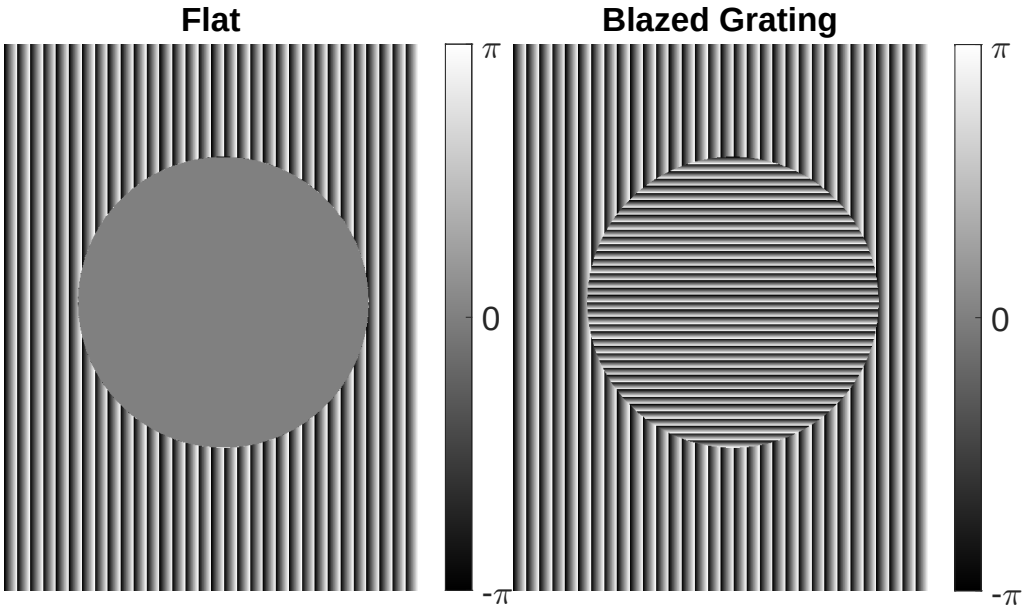


Figure 2.12: Hologram images of the SLM. (a) Blazed gratings (stripes) appear only in the inactive background region, while the active circular region remains flat. (b) Blazed gratings are present in both the active and background regions.

The term SLM most commonly refers to a liquid-crystal-on-silicon (LCOS) device, described comprehensively in [92].

Liquid crystals (LCs) are fluid-like materials with molecules arranged in regular order [92]. The molecules can be represented as slender rods, whose long axes are mostly aligned. When an electric field is applied, the molecules incline toward the field. Because of the anisotropic nature of LCs, the refractive index varies with the molecular alignment. This phenomenon is known as the electro-optic birefringence effect [93]. SLM operation relies on this principle [94].

A schematic of the SLM is shown in Figure 2.11. The refractive index is controlled

pixel by pixel using an electrode array. The incident wavefront is modulated by the LC layer according to a prescribed refractive index profile.

SLMs usually operate in reflection mode, providing phase-only spatial modulation. For pure phase modulation, the incident light should be linearly polarized and aligned with the fast axis of the SLM device. Other types of SLMs are reviewed in detail in [93].

The pixelated structure of the SLM also functions as a 2D grating, producing both unmodulated zeroth-order and higher-order diffraction patterns. To surpass the zeroth-order diffraction light and improve modulation efficiency, blazed gratings are often incorporated into the SLM surface pattern. In this configuration, a spatial filter, such as a pinhole, selectively transmits the high-intensity, stable, modulated first-order (positive) diffraction light while blocking other orders. Figure 2.12 shows hologram images of the SLM with and without the blazed grating.

SLMs enable high-resolution wavefront compensation, with up to about ten million pixels, making them suitable for correcting complex aberrations [93]. However, their optical efficiency is reduced by diffraction, and their polarization sensitivity limits certain applications.

Deformable Phase Plate

The DMs and SLMs described above usually operate in reflection modes. In this configuration, the optical paths should always be folded. To simplify system design, researchers have been working on designing and constructing transparent, plug-in devices [95–97]. One such device is the deformable phase plate (DPP). It consists of a high-refractive-index liquid chamber sealed between a soft, continuous transparent membrane and a solid glass substrate. The glass substrate contains multiple transparent pix-

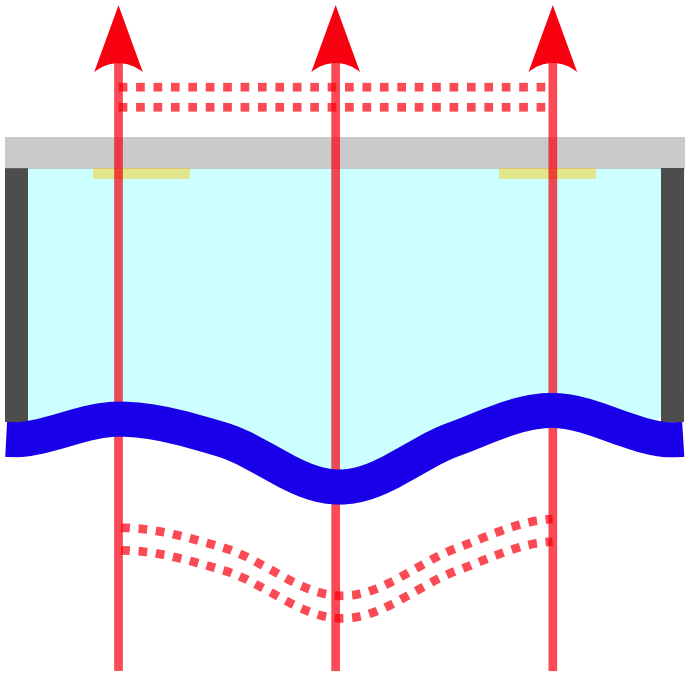


Figure 2.13: Schematic of the DPP. A high-refractive-index fluid (light blue) is enclosed between a conductive transparent membrane (dark blue) and a glass substrate (gray). Voltage applied to the transparent electrodes (yellow blocks) displaces the corresponding membrane regions, modulating the wavefront (red dashed lines) of the transmitted light (red arrows).

elated electrodes. When voltage is applied, the corresponding region of the conductive membrane is displaced via electrostatic actuation [98], altering the optical path length and modulating the wavefront phase according to Equation 2.21.

This device offers several benefits: it is easy to integrate, polarization-independent, compact, high-resolution, and scalable [98]. However, as a newly designed device, the resolution and speed are not comparable to conventional SLMs and DMs. Further details are provided in [99–102]. A schematic of the DPP is shown in Figure 2.13.

2.3.3 Adaptive Optics Implementation

Wavefront Sensors

Wavefront sensors are broadly categorized into two types: those based on geometrical optics [103–105] and those based on diffractive optics [106–108].

A typical geometrical-optics-based wavefront sensor is the Shack-Hartmann sensor (SHS). The typical configuration of SHS consists of a microlens array and a charge-coupled device (CCD) camera [109]. The microlenses sample the incident wavefront and focus it onto the focal plane, where the CCD is placed. Wavefront gradients are determined by measuring displacements of the focal spots relative to reference positions formed by a plane wavefront [110–112]. This method requires a point-like light source and is therefore not universally applicable to all specimens [113–116]. A schematic of the SHS is shown in Figure 2.14.

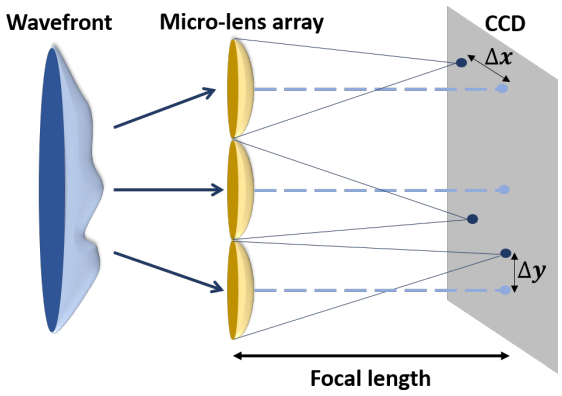


Figure 2.14: Fundamental configuration and operating principle of a SHS. An aberrated wavefront is sampled by a microlens array onto a CCD. Light blue spots indicate reference positions for a planar wavefront, while dark blue spots show positions after sampling the aberrated wavefront. The displacements, Δx and Δy, are measured to determine local wavefront gradients.

Diffractive-optics-based interferometers operate by interfering a distorted wavefront with a reference planar wavefront to measure the phase difference. Because this method relies on interference, it requires spatially coherent light and is normally used for DM calibration or illumination beam measurement. Several techniques can extract the distorted wavefront from the interference pattern. A common approach is off-axis digital hologram reconstruction: Fourier-transform the interference pattern and identify the first-order side lobes that encode frequency information. One side lobe is then shifted to the origin, followed by an inverse Fourier transform. This procedure yields the phase profile of the distorted wavefront [117]. Moreover, phase-shifting interferometry, which can eliminate the intrinsic phase sign ambiguity in a single interference measurement, has become a standard wavefront sensing technique [118–120]. By recording multiple interferograms with various phase shifts of the reference wavefront, the distorted wavefront can be calculated without the requirement of a known reference wavefront. The concept of this technique is also used in Section 3.4.1.

Sensorless Adaptive Optics

Wavefront-sensorless adaptive optics (sensorless AO) determines correction aberrations indirectly through a series of image measurements, each acquired with a different aberration—the 'bias aberration'—introduced by the adaptive element. Image quality optimization serves as the criterion guiding adjustments of the adaptive element [121].

Although speed is its main limitation, sensorless AO offers significant advantages. It requires minimal modifications to microscope hardware, making it cost-effective and avoiding cumbersome optical designs. It also eliminates non-common path errors (Figure 2.15) [122]. Moreover, optical efficiency is improved because all light is used for

imaging rather than diverted to a wavefront sensor. Finally, sensorless AO is applicable in scenarios where wavefront sensors cannot be used, such as in scattering tissues or under extremely low-light conditions. This thesis investigates and advances sensorless AO.

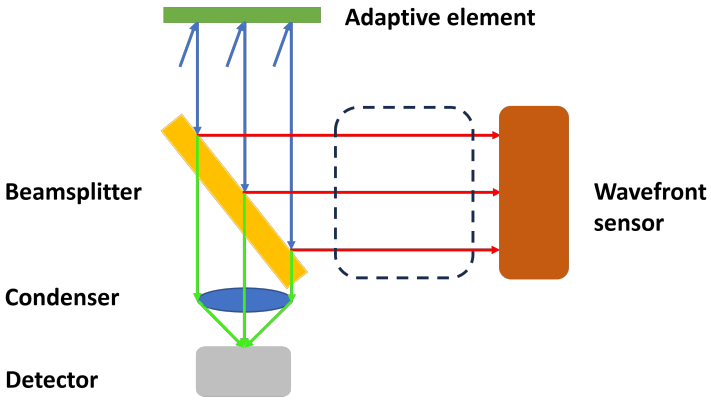


Figure 2.15: Non-common path error in wavefront sensing (detection path). Each optical path contains distinct elements that introduce different aberrations, illustrated by the dashed block—aberrations exist only in the wavefront sensing path. Consequently, the aberrations detected by the wavefront sensor differ from those at the imaging detector.

2.4 Sensorless Adaptive Optics

The following section discusses the three main components of the sensorless AO scheme: the pupil function decomposition method for applying bias aberrations, the image-quality metric, and the estimation algorithm for obtaining the desired compensation phase shape.

2.4.1 Pupil Function Decomposition

The pupil function decomposition methods can be classified into three common approaches: the pixelated piston-based method, the pupil-segmentation zonal method, and the modal method [123].

Pixelated piston-based method

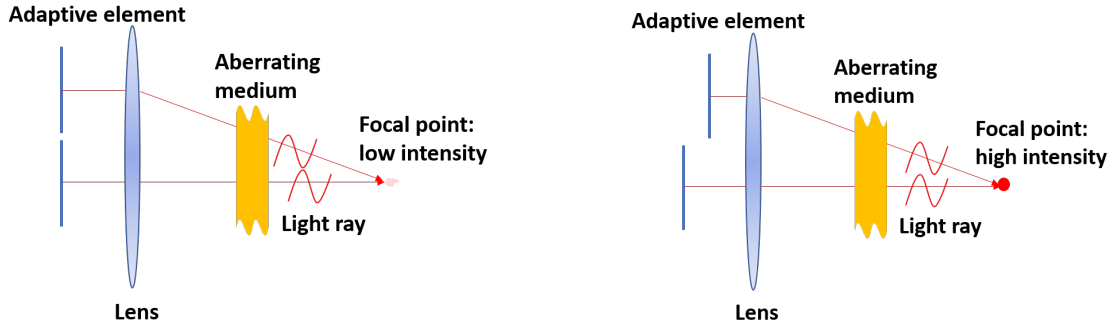
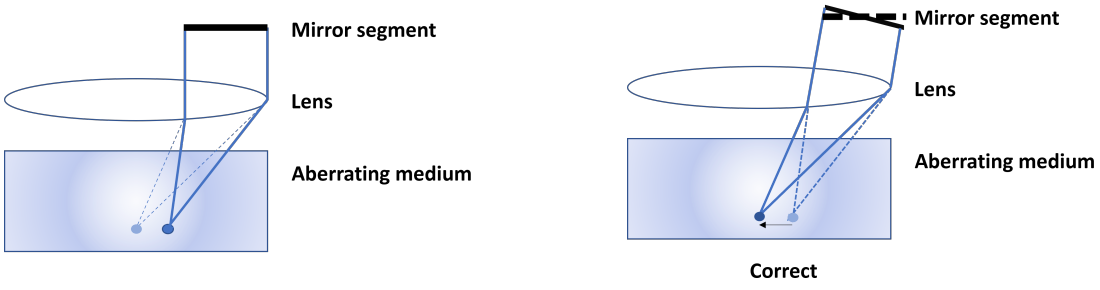


Figure 2.16: Pixelated piston-based method. (a) When imaging through an aberrating medium, the waves become out of phase, causing destructive interference and reduced image intensity. (b) By adjusting the piston of each mirror element on a segmented DM or the refractive index of each SLM pixel, the phase differences among the waves are corrected, ensuring constructive interference at the focal point and increasing image intensity. This procedure is repeated for every pixel on the adaptive element.

In the pixelated piston-based method, the pupil phase function is discretized into pixels, each assigned a constant piston value that can be individually controlled. In piston-only correctors, as shown in Figure 2.16, each mirror element of a segmented DM or each individual SLM pixel's refractive index is adjusted to restore constructive interference at the focal point, with maximized image intensity serving as the optimization criterion. In some applications, the phases of selected pixels are modulated at distinct frequencies.

The resulting intensity is recorded as a video, and a Fourier transform is then applied to extract the phase information of each pixel [124–126].

Zonal method



(a) (b)

Figure 2.17: Zonal method. A single pupil zone is illuminated and modulated here, with a point object for clarity. (a) In the presence of aberrations, the image shifts, shown as the dark blue spot, while the light blue spot represents the desired image position. (b) The image displacement determines the tilt of each mirror segment, ensuring that all rays converge at the same focal point. The shifted image (light blue spot) is corrected to the desired position (dark blue spot) via mirror segment tilt. This procedure is repeated for each mirror segment. To restore constructive interference, the piston of each mirror segment is adjusted following the same steps as in the pixelated piston-based method.

In the zonal method, the pupil is divided into several zones [87, 127, 128], and each zone is illuminated sequentially. When imaging through an aberrating medium, the local tip/tilt of each wavefront segment, induced by the specimen's refractive index structure, causes the image to shift, resulting in measurable displacements. These displacements determine the slope of the wavefront in each zone. The required tilt of each mirror segment is then calculated to ensure all corrected rays converge at the focal point. The process is illustrated in Figure 2.17. Additionally, the piston of each mirror segment can be adjusted to maximize focal intensity, following the principle used in the piston-based

method. Alternatively, full-pupil illumination with intensity-maximisation optimisation is also a valid approach [129]. This zonal method can be implemented using the DM shown in Figure 2.10(c).

Modal method

In contrast to the previous decomposition methods, which divide the pupil function into discrete segments, the modal method represents the pupil phase as a sum of orthogonal basis modes, each weighted by a specific mode coefficient.

In sensorless AO, the image quality serves as the criterion for aberration correction. When using the modal method (Figure 2.18), the presence of a specific aberration mode degrades image quality, reducing image intensity. As a result, the peak of the image quality optimization curve, the metric curve (here using intensity), is shifted from zero. By measuring intensity variations with the mode coefficient applied to the adaptive element, the optimal coefficient for correcting the aberration can be determined. After applying the mode with this optimal coefficient, the image becomes brighter. This procedure can be repeated for all other aberration modes present [130].

Several orthogonal basis mode sets have been developed. A widely used set for circular pupil systems is the Zernike polynomials, $Z_n^m(r, \theta)$. These are a sequence of orthogonal and continuous polynomials defined over a unit disk [131]. After first being employed by F. Zernike in his phase contrast method [132], they have several advantages that make them widely used. Zernike polynomials provide an appropriate mathematical description of the wavefront and a balanced representation of classical aberrations, yielding minimum variance over a circular pupil [133]. Low-order Zernike polynomials correspond closely to classical aberrations. When using appropriately normalized

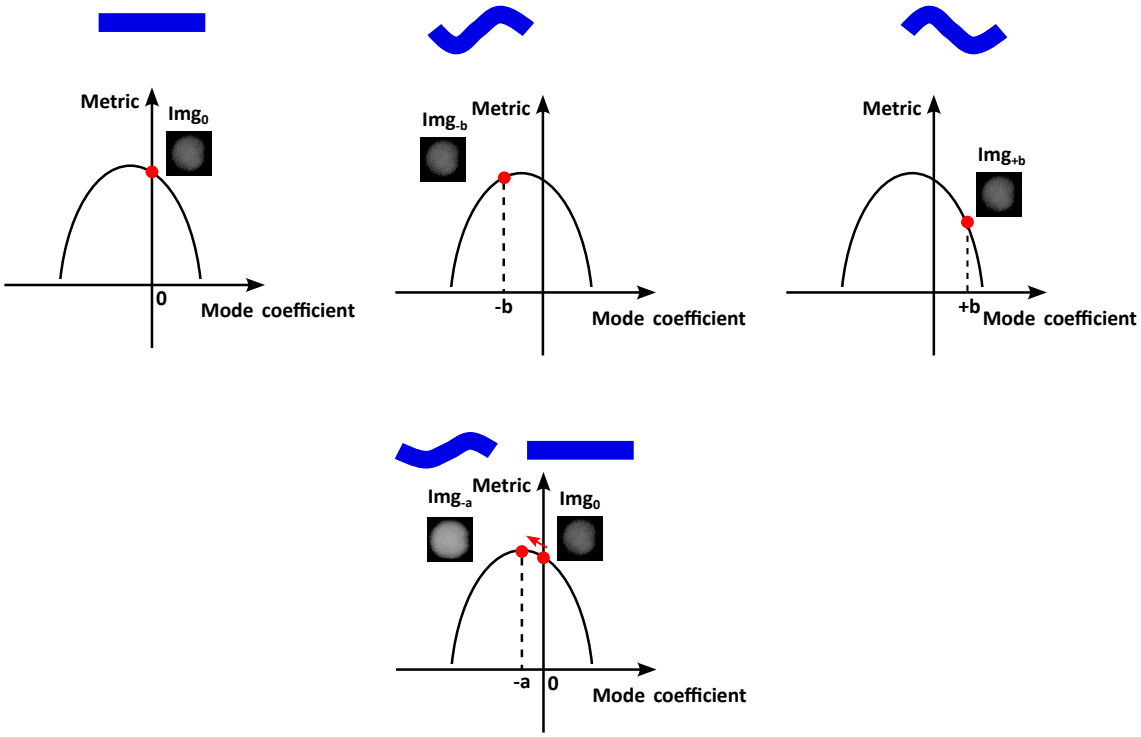


Figure 2.18: Modal method. Top left: With a specific aberration mode, the peak of the metric curve shifts from 0 to $-a$. The image obtained with a planar adaptive element (blue straight block), Img_0, is dim. Top middle and top right: Two biased images with mode coefficient $\pm b$ (bias aberrations illustrated as blue curved blocks) are collected as $Img_{\pm}b$. These metric values are used to determine the optimal mode coefficient $-a$. Bottom: After applying the mode coefficient $-a$ to the adaptive element, the corrected image Img_{-a} becomes brighter.

Zernike equations, the coefficient of each mode represents the root mean square (RMS) wavefront error attributable to the total RMS wavefront error. Furthermore, the orthogonality ensures that the coefficients are independent; that is, the coefficient of one mode does not contribute to any other mode.

For the original mathematical definitions of Zernike polynomials, see [134]. Here, we present the recursive method developed by Shakibaei and Paramesran [135]. The recursive process is illustrated by the following equations:

$$Z_n^m(r,\theta) = N_n^m R_n^m(r)\Theta(m\theta) \tag{2.28}$$

$$N_n^m = \sqrt{\frac{2(n+1)}{1+\delta_{m0}}}, \delta_{m0} = \begin{cases} 1, & \text{if } m = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\Theta(m\theta) = \begin{cases} cos(m\theta), & \text{m}\geq 0 \\ -sin(m\theta), & \text{m}<0 \end{cases}$$

$$\begin{cases} R_0^0(r) = 1 \\ R_n^m(r) = r[R_{n-1}^{|m-1|}(r) + R_{n-1}^{m+1}(r)] - R_{n-2}^m(r) \\ R_n^m(r) \equiv 0, \text{when } n < m \end{cases}$$

In Equation 2.28, $Z_n^m(r,\theta)$ are the Zernike polynomials, N_n^m are the normalization factors, which ensure that the effect of all modes on the signal is similar, $R_n^m(r)$ are the radial polynomials, and $\Theta(m\theta)$ are the azimuthal polynomials. r is the radial distance within range $r \in [0,1]$, and θ is the azimuthal angle within range $\theta \in [0, 2\pi)$. m and n are integer Zernike indices, with $n = 0, 1, 2, \ldots$, and $(n-m)$ even and non-negative.

The first 15 Zernike modes, labeled and listed in Noll convention, are shown in Figure 2.19. The first four radial orders are shown.

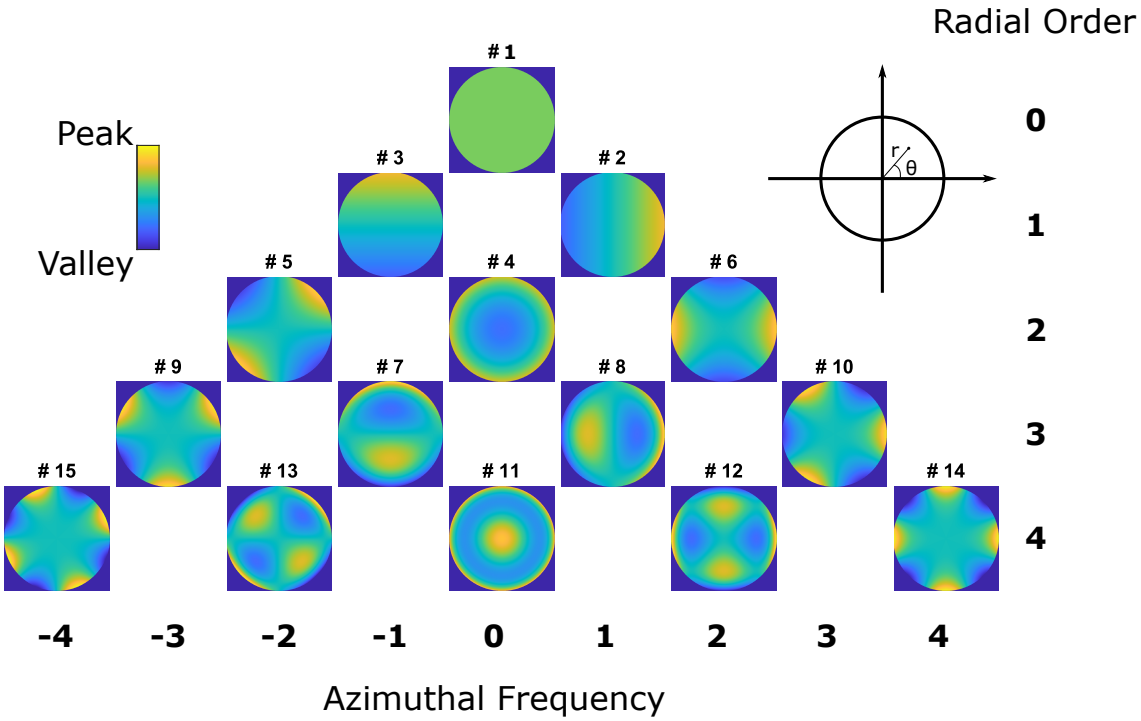


Figure 2.19: Surface plots of the first four radial orders of the Zernike polynomials, defined over the unit disk. The mode patterns are labeled according to the Noll convention, with both the azimuthal frequency and radial order indicated. Color represents the scaled mode values. The representation of r and θ within the unit disk is also shown.

In addition to the advantages noted above, Zernike expansions also have certain limitations. When aberrations are highly complex or discontinuous, using continuous Zernike polynomials can be restrictive. In this case, a pixelated and binary basis mode set, the Walsh-function-generated modes (Walsh modes), may be used. The Walsh mode set, used in this thesis, will be introduced in Chapter 3.

2.4.2 Metric

The metric is chosen as a proxy for image quality, which the AO system aims to optimize. Ideally, the metric should be a smooth, unimodal function, with a global extremum occurring when all aberrations are corrected, and it should also be sensitive to aberrations [130]. One commonly used metric is total intensity, which is particularly applicable to 2PLSM, as the 2PA is quadratic with respect to the illumination intensity and therefore highly sensitive to aberrations.

In this thesis, the Strehl ratio (SR) is generally used as the metric during simulations for normalization purposes. SR provides a simple and widely used measure of the effect of aberrations on image quality, originally introduced by Karl Strehl [136]. It is defined as the ratio of the maximum intensity of the diffraction pattern in the presence of aberrations in the objective pupil function to the maximum intensity in their absence [137].

Other metrics, such as image sharpness, are also employed in different microscope configurations [138].

2.4.3 Estimation Algorithm

The estimation algorithm is a vital component of sensorless AO for obtaining the desired compensation phase shape. These algorithms are divided into two main categories: model-based and model-free. The model here refers to the mathematical relationship between the metric function and its variation with the aberration amplitudes [139].

In the model-based algorithm, the problem is converted into fitting a parabola with

several data points and finding its maximum, since the metric function can usually be approximated by a parabola near the maximum [140, 141]. The data points are collected by measuring the metric values of images with bias aberrations applied to the adaptive element at various amplitudes. Eventually, the aberration amplitude that maximizes the metric function is obtained.

In the model-free algorithm, however, if there is no prior knowledge of the metric function, more measurements should be taken to locate the extremum of the curve, for example, using the hill-climbing algorithm [142]. Due to the increased number of measurements, the total exposure time also increases, which may lead to issues such as photobleaching.

Because of their high speed and reduced specimen exposure, most applications use model-based algorithms.

Established Estimation Algorithms

The previously introduced estimation algorithms are designed to correct a single aberration. When multiple aberrations are present, two common approaches are employed: the conventional $2N + 1$ algorithm (CA) [24] and the sequential $3N$ algorithm (SA) [130].

In the CA, metric values for all aberrations are measured first, after which the compensation phase shape is applied to the adaptive element. The workflow of this algorithm is illustrated in Figure 2.20. For clarity, the case with two bias aberration amplitudes $\pm b$ is shown here. At the end of the correction, CA collects $2N + 1$ metric values, where N is the number of aberrations corrected.

The other approach is the SA. In this method, two biased metric measurements and

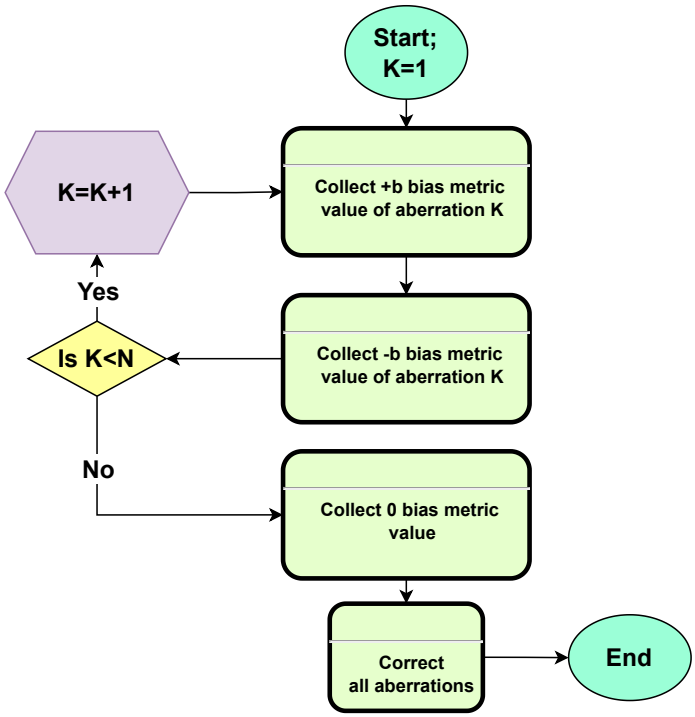


Figure 2.20: Workflow of the CA: Starting with aberration $K = 1$, two biased metric values are collected, bias amplitudes are $\pm b$, then the process moves to the next aberration, $K = 2$. This pattern continues until all N aberrations are addressed, without updating the 'zero-bias' pupil. After collecting the final unbiased metric value, all aberration amplitudes are estimated via parabola fitting, and the resulting compensation phase pattern, obtained by summing all aberrations with their estimated amplitudes, is applied.

one unbiased metric measurement are taken for a single aberration to determine its optimal amplitude. At this stage, the corresponding aberration compensation is applied to the adaptive element, and the 'zero-bias' pupil function is updated. The workflow of this algorithm is illustrated in Figure 2.21. At the end of the correction, SA collects $3N$ metric values.

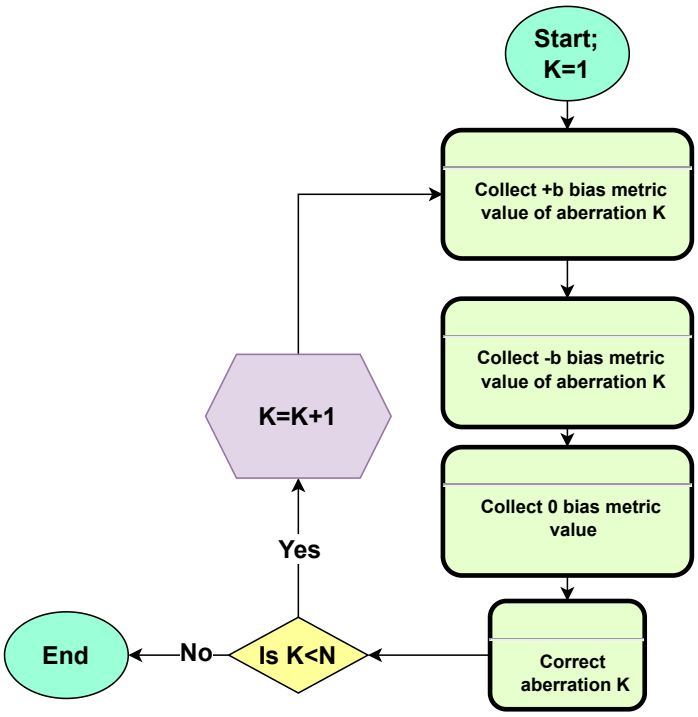


Figure 2.21: Workflow of the SA: Starting with aberration $K = 1$, two biased metric values (with bias amplitudes $\pm b$) and one unbiased metric value are collected. The amplitude of aberration K is estimated using parabola fitting, and the corresponding phase pattern is applied to the adaptive element to update the 'zero-bias' pupil function. The process then moves to the next aberration, $K = 2$, and continues in the same manner until all N aberrations are addressed and applied.

Although those algorithms are well-designed, they have to condense the properties contained in an image, such as the aberration structure information, into a single metric value. This approach does not fully exploit the useful information in the images. They are also limited when applied to a multidimensional metric space, which involves collecting a large number of images and can be time-consuming. Recently, a new class

of estimation algorithms—neural networks—has emerged, which will be introduced in Section 2.5. This approach can effectively address the limitations mentioned above.

2.5 Neural Network

A neural network (hereafter referred to as NN) is a computational problem-solving paradigm inspired by the human brain, mimicking its basic components and processes [143]. It was originally developed to better process and exploit the increasingly large amounts of data available today. In conventional approaches based on pre-programmed instructions, a computer follows explicit algorithms provided by humans, which can only perform simple tasks and extract limited information from the data. In contrast, a NN learns solutions directly from datasets [144]. Users only need to define the task, rather than provide the algorithm. By analyzing datasets, the NN extracts relevant information and constructs a model that maps inputs to outputs. Because it is more directly tied to the data than conventional algorithms, a NN can make more effective use of the information and perform more complex tasks.

2.5.1 Neural Network for Adaptive Optics

Currently, many studies focus on using NNs for AO [145–156]. In the context of estimation algorithms for sensorless AO, selecting a suitable learning method requires consideration of the mathematical model of the microscope imaging process. This model follows Equations 2.26 and 2.27, allowing the resulting PSF and image to be calculated. While the forward optical model is well defined, constructing a backward model is chal-

lenging because the phase information is not directly accessible from the PSF. Therefore, supervised learning is ideal for training NNs as sensorless AO estimation algorithms, in which each data point in the dataset comes with a corresponding label, enabling the model to learn the relationship between inputs and outputs. [157–160].

Although supervised learning is ideal for training NNs as sensorless AO estimation algorithms, it has certain limitations. One major limitation is that NN learning relies entirely on finite training datasets. When presented with data outside this domain, there is no guarantee that the NN will produce a correct or desirable output. Consequently, most supervised learning NNs use images of point-like objects, such as PSFs. In this thesis, however, we employ a specially designed NN, called the machine learning–assisted wavefront-sensorless AO control method (MLAO) [146]. This approach allows the NN to handle extended images instead of being limited to point-like objects (see Chapter 7).

Supervised learning still faces challenges, including the need for large training datasets, which are often difficult to obtain. Consequently, self-supervised learning has emerged as an alternative, offering improved data efficiency [161–163]. This approach embeds the physical relationships of the imaging process, enabling the extraction of desired information from a relatively small or even single dataset. Rather than manual labels as in supervised learning, it generates pseudo-labels from the inherent structure of the data. An example of its application in AO is CoCoA [164].

In summary, supervised learning is well-suited for online or real-time correction, as the network can operate immediately once trained; however, it requires large training datasets. In contrast, self-supervised learning is generally used in offline settings, since it must be retrained for each new imaging scenario. Its main advantage is that it does

not require pre-training or reliance on pre-existing datasets, as it can learn directly from the data of the current experiment.

In this thesis, we choose the supervised learning approach for its real-time correction capability, which is relevant to live imaging applications.

Chapter 3

Introduction to the Walsh Modes

Although smooth polynomials such as Zernike modes are commonly used for aberration correction, their smooth nature limits their effectiveness in compensating relatively gradual wavefront distortions. In contrast, many biological structures—such as the skull, dura, and muscle tissues—are highly scattering, producing discrete and sharp wavefront distortions. Walsh modes, with their binary and discrete nature, have been proposed as complementary to Zernike modes and may be better suited for correcting wavefront distortions induced by scattering tissues.

3.1 Inhomogeneous Refractive Indices of Biological Specimens

Deep tissue imaging is of significant interest in many biomedical fields. As a light beam propagates through inhomogeneous tissue, it undergoes absorption, refraction,

diffraction, and scattering. These effects become more pronounced with increasing imaging depth. As imaging depth increases, scattering that originates from inhomogeneities at spatial scales comparable to or smaller than the wavelength becomes an increasingly important contributor to optical opacity [165]. Consequently, optical wavefronts are increasingly distorted with imaging depth.

Conventional continuous modes, such as Zernike polynomials, are limited to smooth and low-order shapes, which are unsuitable to compensate for more complex sample-induced aberrations. As demonstrated in [165] and [166], discrete wavefront compensation approaches are therefore preferred. To address this problem, pixelated and binary Walsh mode sets provide an appropriate choice.

3.2 Introduction of Walsh Modes

Walsh functions were first applied by electrical engineers, such as Frank Fowle [167], in studies of overhead conductor transmission around the early 1900s [168]. They were later introduced into mathematics by J. L. Walsh [169]. In the 1960s, interest in Walsh functions grew rapidly due to their potential applications in communications, signal and system analysis, and related areas [170–173].

Today, Walsh modes are widely used in wavefront sensing [11–15], beam shaping [174], adaptive element control [175], and particularly in scattering compensation, as their pixelated structure makes them well suited for complex wavefront compensation [176–181]. In scattering compensation, the closely related Hadamard basis set is also widely employed [16–23]. Walsh modes also continue to show promise in quantum

applications [182] and in astronomy [15, 183].

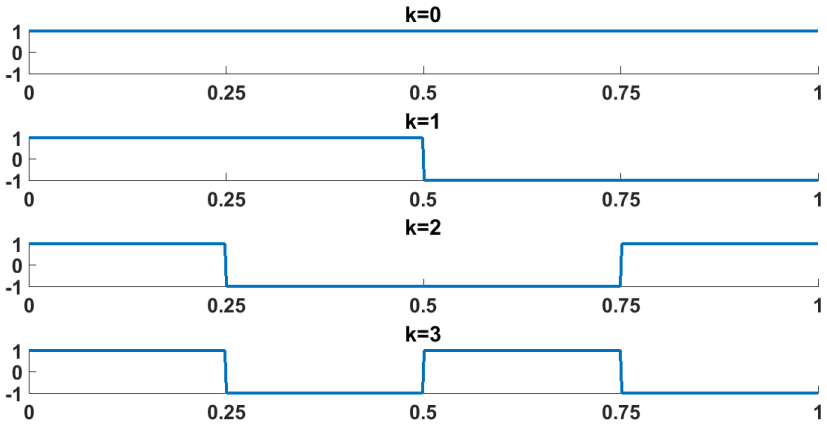


Figure 3.1: The first four Walsh functions with index k.

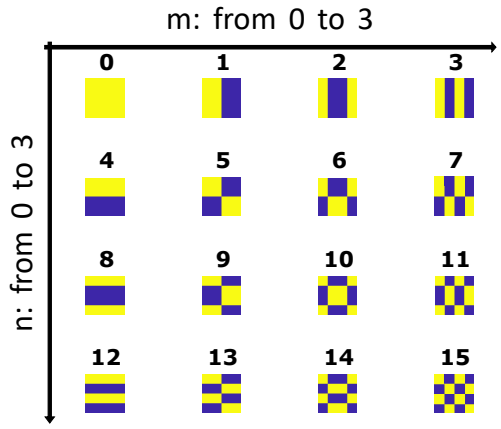
The mathematical definition of the Walsh function is [184]:

$$WAL_0(x) = \begin{cases} 1, & \text{when } x \in [0,1] \\ 0, & \text{when } x \in (-\infty, 0) \cup (1, \infty) \end{cases}, \tag{3.1}$$

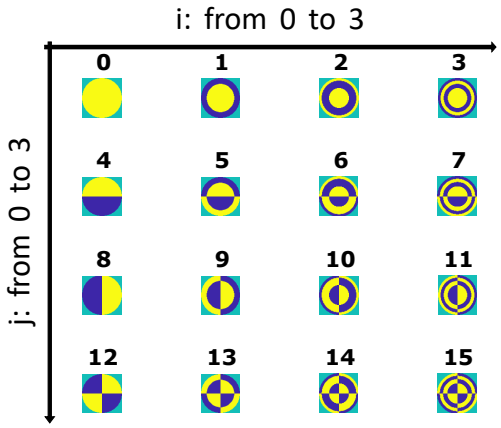
$$WAL_{2j+p}(x) = WAL_j(2x) + (-1)^{j+p} WAL_j(2x-1), \tag{3.2}$$

in which $p = 0, 1$; $j = 0, 1, \ldots$ In Equation 3.2, j and $2j + p$ are both indices of Walsh functions. The Walsh function index k in WAL_k indicates the number of signal changes within the interval $x \in [0, 1]$. This indexing method is referred to as the 'sequency' order [185]. The first four Walsh functions are shown in Figure 3.1, where only the interval $x \in [0, 1]$ with binary values ± 1 is displayed.

For practical use in optical systems, 2D Walsh modes are required. The Cartesian Walsh modes [186], defined as the pointwise products of Walsh functions in the Carte-



(a)



(b)

Figure 3.2: (a) The first sixteen Cartesian Walsh modes; the m- and n-axes represent the indices of the Walsh functions. Yellow indicates a value of $+1$, and blue indicates -1. Each Cartesian Walsh mode is also labeled in a 1D sequence for clarity. (b) The first sixteen polar Walsh modes, labeled in a 1D sequence; the i-axis represents the indices of radial Walsh functions, and the j-axis represents the indices of azimuthal Walsh functions. Color represents mode values, with yellow indicating $+1$ and blue indicating -1. The domain is a unit disk.

sian coordinates, are introduced, as illustrated in Equation 3.3:

$$WD_{m,n}(x,y) = WAL_m(x) \times WAL_n(y), \tag{3.3}$$

in which $m, n = 0, 1, 2, \ldots$ In Equation 3.3, $WD_{m,n}(x,y)$ denotes the resulting Cartesian Walsh mode, where $WAL_m(x)$ and $WAL_n(y)$ are Walsh functions with indices m and n, respectively. An example of the first 16 Cartesian Walsh modes is shown in Figure 3.2(a). For simplicity of representation, each mode is labeled in a one-dimensional (1D) sequence, with each row and column arranged in sequency order.

The use of Cartesian Walsh modes on circular pupils has limitations: they are only orthogonal over a square aperture rather than a circular aperture, hence they are not applicable.

To better accommodate circular domains, polar Walsh modes were developed in polar coordinates. Defined over a circular aperture, these modes are orthogonal when used to correct aberrations of a circular pupil. The derivation process is described in [186]. In general, the polar Walsh modes can be illustrated as:

$$WP_{i,j}(r,\theta) = WAL_i(r^2) \times WAL_j(\frac{\theta}{2\pi}), \tag{3.4}$$

in which $i, j = 0, 1, 2, \ldots$ In Equation 3.4, $WP_{i,j}(r,\theta)$ is the resulting polar Walsh mode, where $WAL_i(r^2)$ and $WAL_j(\frac{\theta}{2\pi})$ are radial and azimuthal Walsh functions, i and j are their respective indices. Here, r and θ denote the radius and angle in polar coordinates, respectively. Figure 3.2(b) shows the first sixteen polar Walsh modes labeled in a 1D sequence, with each row and column arranged in sequency order [185].

3.3 Walsh Mode Metric Function

In this section, we derive Walsh-mode–based metric functions, including those for single-Walsh-mode and multiple-Walsh-mode cases.

3.3.1 Single Walsh Mode Metric Function

The single-Walsh-mode metric function is derived as follows. The one-photon simplified system model we use is shown in Figure 3.3.

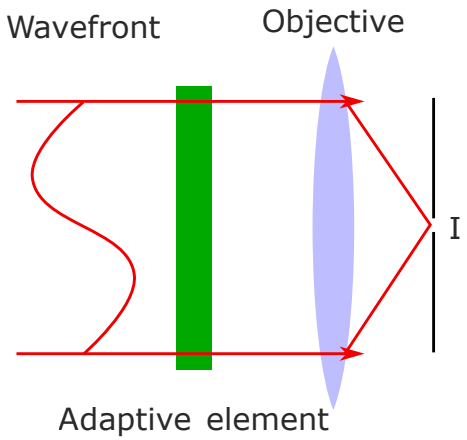


Figure 3.3: The simplified optical system model. The wavefront (red curve) passes through an adaptive element (green block) for compensation and is then focused by the objective. A vanishingly small pinhole detector on the optical axis measures the intensity I.

First, following the Equation 2.26, we calculate the on-axis intensity of a point object:

$$I = I_0 \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} e^{j\Phi(r,\theta)} r \, dr d\theta \right|^2 \tag{3.5}$$

Here, I_0 denotes the illumination intensity, $\Phi(r, \theta)$ is the phase function, r and θ are the radial and azimuthal coordinates. The pupil radius is defined as 1.

The pupil phase function is represented by a single polar Walsh mode WP_k. k is the 1D index labeled in Figure 3.2(b), and is non-zero here. Owing to the pixelated nature of Walsh modes, the intensity I can be calculated as:

$$I = I_0 \left| \frac{\sum_{m,n \in WP_k} e^{j\phi_{m,n}}}{N} \right|^2 \tag{3.6}$$

m, n are the coordinates of pixels on the polar Walsh mode, $\phi_{m,n}$ represents the phase of each polar Walsh mode pixel, and N is the total number of polar Walsh mode pixels.

For a single polar Walsh mode WP_k, half of the pixel phases are positive and the other half are equal in magnitude but negative. Using this property, the Equation 3.6 can be simplified as:

$$I = I_0 \left| \frac{e^{j\varphi}}{2} + \frac{e^{-j\varphi}}{2} \right|^2 = I_0 \left| \cos \varphi \right|^2 \tag{3.7}$$

φ is the mode coefficient. In this equation, the intensity varies periodically with a coefficient period of π. For simplicity, the metric is normalized to the Strehl ratio (SR). The resulting metric function is:

$$\mathrm{SR} = \left| \cos \varphi \right|^2 \tag{3.8}$$

The metric function generated by a single Walsh mode is shown in Figure 3.4.

The periodic, 'phase-wrapping' property distinguishes Walsh modes from Zernike modes, which only exhibit a single metric maximum at a coefficient of zero. This periodicity introduces both unique challenges and potential advantages for using Walsh modes in sensorless AO aberration correction. When using this Walsh-mode-based metric, the bias aberration coefficient and the optimal Walsh mode coefficient lie within a single period, simplifying Walsh-mode-based correction and ensuring that the correction capability is not limited by the modulation range of the adaptive elements.

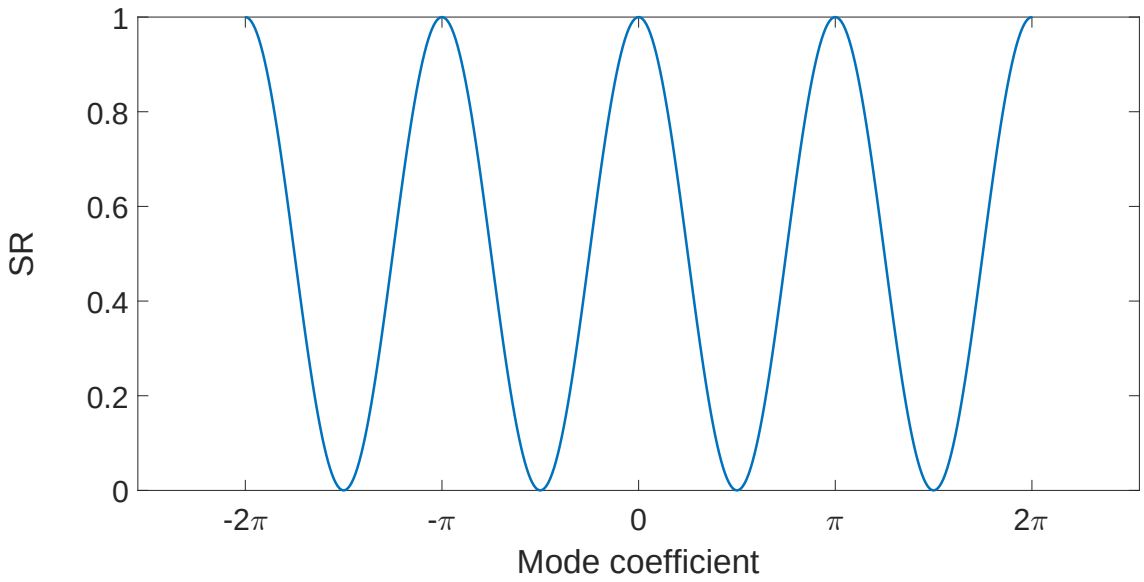


Figure 3.4: Single Walsh mode metric function. The function is periodic with a period of π.

3.3.2 Walsh Mode Metric Space

We further investigate the metric function structure generated by multiple Walsh modes.

The function is defined over N dimensions, spanned by N Walsh mode coefficients. Due to the discrete and binary nature of Walsh modes, the Walsh mode metric space (defined as space S) has multiple maxima. As shown in [187], space S has a repetitive lattice-like structure. The lattice structure has a maximum at the origin, and if there are maxima at points u and v in space S, then there are also maxima at points $u + v$ and $u - v$ [188].

To illustrate the lattice-like structure of the space S, we generate an N-dimensional Walsh mode set W_N in matrix form, for example, with $N = 4$. In this matrix, each row

represents a single Walsh mode, and each column corresponds to a pixel on the pupil. The first row, representing the piston mode, is removed because it has no effect on the signal and is ignored during training. The resulting reduced matrix is denoted as W'_N. At this stage, the matrix is still redundant, as the number of columns exceeds its rank. To resolve this, any one column can be removed. Here, the first column is selected. The entire procedure is illustrated in Figure 3.5. The final matrix is defined as W''_N.

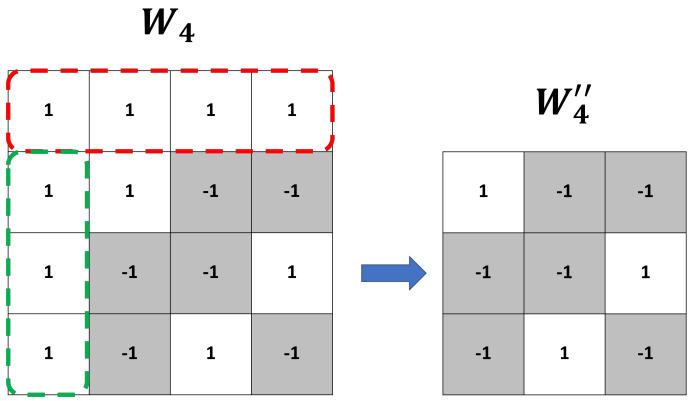


Figure 3.5: The process of removing the redundant mode and pixel from a Walsh mode set. The 4×4 matrix represents a 4-pixel Walsh mode set, with each row corresponding to a Walsh function. The red dashed area indicates the first row, representing the piston mode, which is removed. Next, the first column (green dashed area), representing a pixel, is removed. The final output is the matrix on the right, with dimensions 3×3.

According to [187], the multi-dimensional metric function can be expressed as $I(a')$. I represents the intensity at the focusing point, the one-photon simplified system model is shown in Figure 3.3. a' is a vector containing the phase value of each pixel with the first pixel removed, whose length is $(N-1)$. a' can be expanded in terms of Walsh modes: $a' = W''^{T}_{N} b'$, b' is a vector, each element of which represents the corresponding Walsh mode coefficient; T denotes transposition. The metric function can then be rep-

resented as: $I(W_N''^T b')$. Due to the phase-wrapping effects inherent in the Walsh mode, the maxima of $I(a')$ form a lattice [187]. Likewise, after the transformation by $W_N''^T$, the maxima of $I(W_N''^T b')$ also retain a lattice structure. The procedure of W_N'' generating the lattice-like metric function is shown in Figure 3.6. Note that only three dimensions are shown in the figure, despite there being four Walsh modes, because the piston mode has been removed as it has no effect.

If a metric space S is generated using the matrix W_N'', the resulting metric function exhibits a lattice-like structure, with each row of the matrix corresponding to a lattice vector [188]. The properties of this lattice can be characterized by its Gram matrix, defined as the matrix of inner products of the lattice vectors. For an N-pixel Walsh mode set, the Gram matrix is given by:

$$G_N = W_N''^T W_N'' = \begin{pmatrix} N-1 & -1 & \cdots \\ -1 & N-1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{3.9}$$

In this example the Gram matrix G_4 is:

$$G_4 = W_4''^T W_4'' = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \tag{3.10}$$

The Gram matrix is identical to that of the lattice known as A_{n-1}^* [188]. Figure 3.7 illustrates an example of the lattice described by A_3^*, which is the same as the lattice structure shown in Figure 3.6 [187].

In this space S, each axis corresponds to the coefficient of a specific Walsh mode, and any point in S represents the metric value generated by the Walsh modes with the

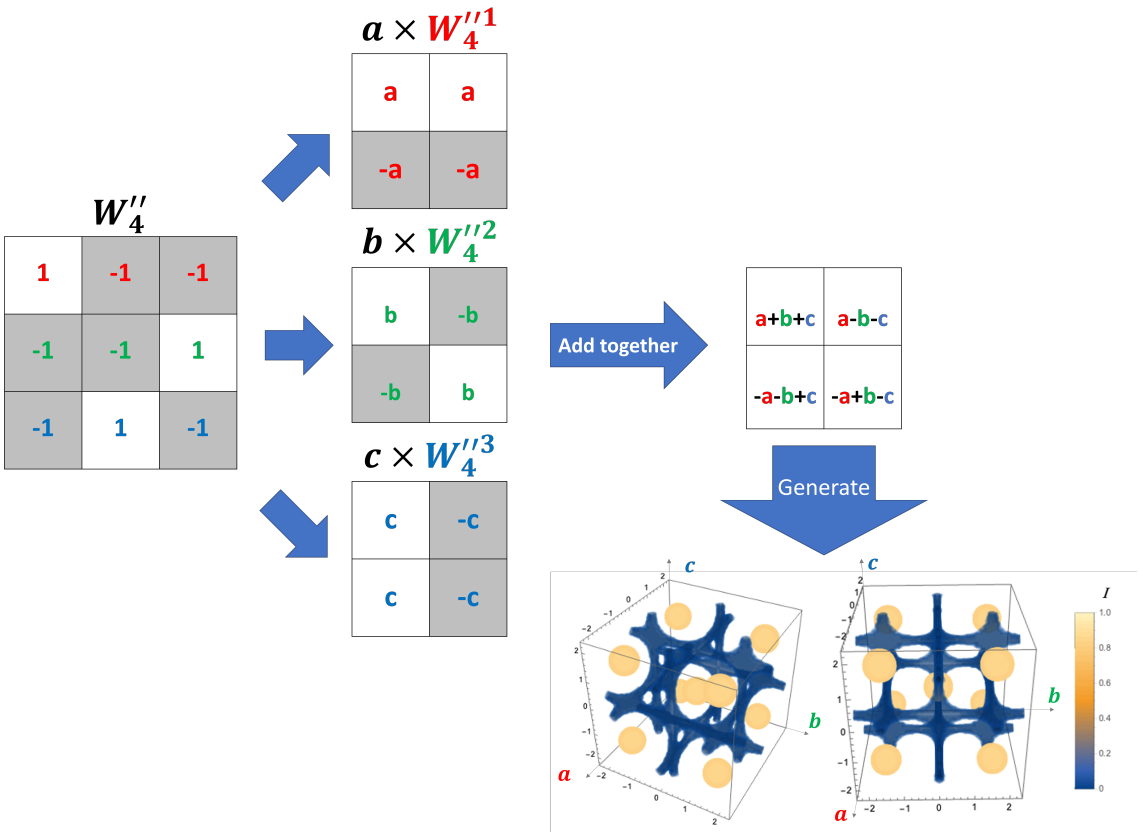


Figure 3.6: Lattice structure of the metric function generated by a 4-pixel Walsh mode set. a, b, and c are the coefficients of the three Walsh modes except the piston mode. Colors of red, green and blue represent specific Walsh modes. The left 3×3 matrix illustrates the reduced 4-pixel Walsh mode set, the W_4'', with each row representing a certain Walsh mode. The pupil phase function (top-right) is constructed by multiplying the Walsh modes (represented by three 2×2 matrices formed by reshaping each row in W_4'', as well as restoring the top-left pixel which has been removed in W_4'') with their corresponding coefficients and then adding them together, thereby generating a lattice-like metric function (bottom). Only three dimensions of the metric function are shown, as the piston mode is neglected; the axes represent a, b, and c in radians. The structures are visualized from different angles. Blue contours are set at $I = 0.01$, indicating the positions of zeros, and orange contours are set at $I = 0.8$, indicating the maxima [187]. Values between the 0.01 and 0.8 contours are rendered transparent for clarity.

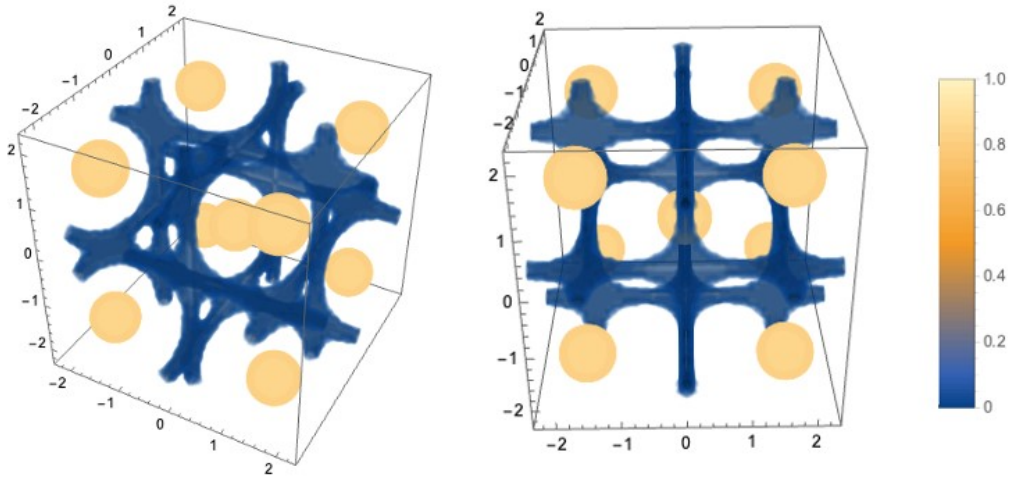


Figure 3.7: The A_3^* lattice [187].

corresponding coefficients. Due to the lattice symmetry, the structure around each lattice point is identical, meaning that the metric values near each lattice point are similar. Since there are infinitely many lattice points in space S, there are infinitely many combinations of mode coefficients that yield the same metric response. This phenomenon is analogous to phase wrapping of a single Walsh mode, as illustrated in Figure 3.4, where a specific metric value can correspond to infinitely many mode coefficients. Figure 3.8 further demonstrates that different combinations of mode coefficients can produce the same pupil phase function and metric value.

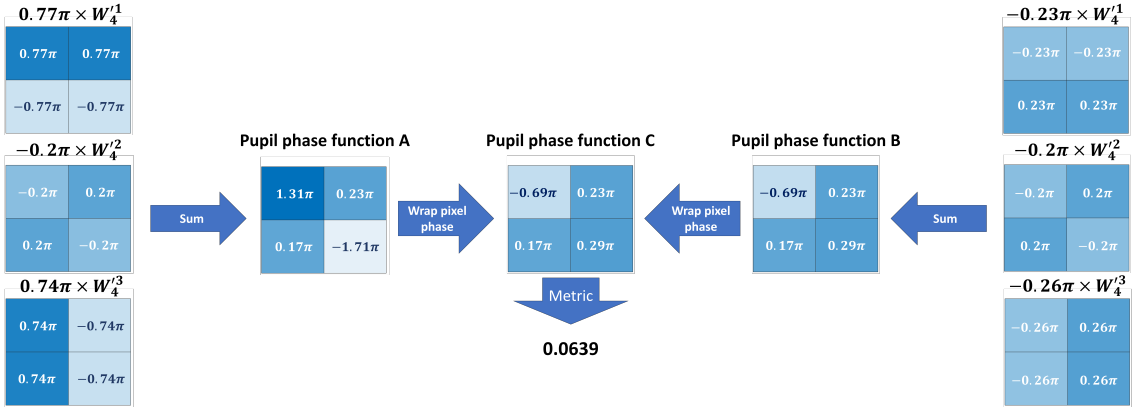


Figure 3.8: An example illustrating that different combinations of Walsh mode coefficients can produce the same pupil phase function and metric response. The phase value of the pixel on each Walsh mode or pupil phase function is indicated by the pixel colour and the label on it. Two combinations of three coefficients (numbers above each 2×2 matrices) are generated for three Walsh modes of W_4' (three 2×2 matrices on left and right), with the piston mode already removed. The three Walsh modes in one set are summed to compute the corresponding pupil phase functions A and B, and each pixel's phase on pupil phase functions A and B is wrapped to the range $[-\pi, \pi)$. After this process, two different combinations of mode coefficients yield the same pupil phase function — pupil phase function C, and the same metric value. The pupil phase function C is the more 'primary' one that should compensate for the wavefront in experiments.

3.4 Walsh Mode Metric Function Based Estimation Algorithm

In this section, we discuss Walsh-mode-metric-function-based estimation algorithms, including those for the single Walsh mode metric function and the Walsh mode metric space.

3.4.1 Single Walsh Mode Metric Function Based Estimation Algorithm

Based on the previous sinusoidal metric function, we derive this Walsh-mode-metric-function-based estimation algorithm.

According to Equation 3.8, we have:

$$\mathrm{SR} = |\cos\varphi|^2 = \frac{1}{2} + \frac{1}{2}\cos 2\varphi \tag{3.11}$$

A Walsh mode with coefficients: $-b$, b, and 0, are added to the adaptive element. Resulting metric values are:

$$\mathrm{SR}_+ = \mathrm{SR}(\varphi + b) = \frac{1}{2} + \frac{1}{2}\cos 2\varphi \cos 2b - \frac{1}{2}\sin 2\varphi \sin 2b \tag{3.12}$$

$$\mathrm{SR}_- = \mathrm{SR}(\varphi - b) = \frac{1}{2} + \frac{1}{2}\cos 2\varphi \cos 2b + \frac{1}{2}\sin 2\varphi \sin 2b \tag{3.13}$$

$$\mathrm{SR}_0 = \mathrm{SR}(\varphi) = \frac{1}{2} + \frac{1}{2}\cos 2\varphi \tag{3.14}$$

Then it can be inferred that:

$$\mathrm{SR}_+ - \mathrm{SR}_- = -\sin 2\varphi \sin 2b \tag{3.15}$$

$$\mathrm{SR}_+ + \mathrm{SR}_- - 2\mathrm{SR}_0 = \cos 2\varphi \cos 2b - \cos 2\varphi \tag{3.16}$$

So that the mode coefficient φ can be calculated as:

$$\varphi = \frac{1}{2}\arctan\left(\frac{1-\cos 2b}{\sin 2b}\frac{\mathrm{SR}_+ - \mathrm{SR}_-}{\mathrm{SR}_+ + \mathrm{SR}_- - 2\mathrm{SR}_0}\right) \tag{3.17}$$

This estimation algorithm is analogous to phase-shifting interferometry (PSI), which is introduced in Section 2.3.3. By measuring biased metric values with equal but opposite

mode coefficients (SR_+ and SR_-), along with the unbiased metric value SR_0, the optimal mode coefficient can be determined. In practice, when dealing with a set of Walsh modes, the CA and SA introduced in Section 2.4.3 use the PSI method to estimate the coefficient of each mode.

3.4.2 Walsh Mode Metric Space Based Estimation Algorithm

As discussed in Section 2.4.3, conventional estimation algorithms encounter difficulties in handling multidimensional metric spaces, often requiring excessive image acquisition and computation time. To overcome this limitation, we employ a neural network (NN) in the thesis, taking advantage of its ability to deal with multidimensional data.

However, as discussed in Section 3.3.2, multiple combinations of coefficients can correct the same aberration. This makes the optimization problem challenging when searching the space S with multiple maxima. NN training becomes much simpler if the search space contains only a single maximum, improving estimation accuracy and speed and significantly simplifying NN architecture.

If we define the search space as the same fixed range for each Walsh mode coefficient (for example, from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, forming the cube cell in Figure 3.9), multiple maxima may occur. In this case, a training dataset can correspond to multiple labels, thereby complicating and confusing NN training. To address this limitation, the search space is confined to a fundamental unit of the lattice known as a Voronoi cell. A Voronoi cell partitions space S such that all points within it are closer to a given lattice point than to any other lattice point [189]. Voronoi cells are convex polytopes whose union covers the

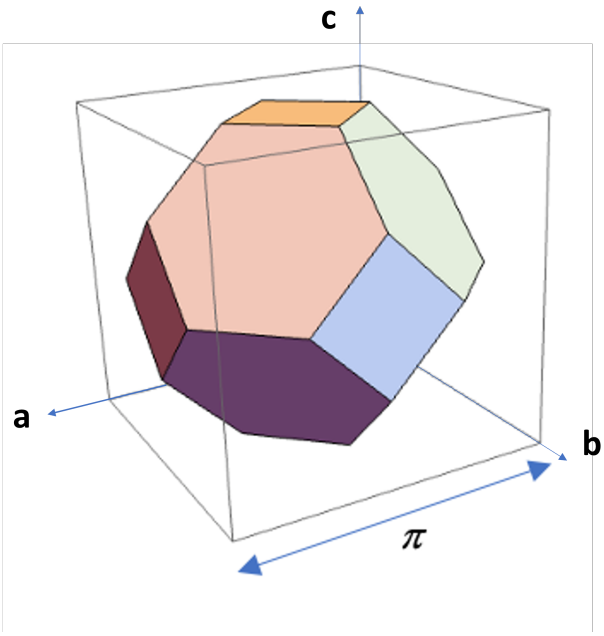


Figure 3.9: Example of a Voronoi cell corresponding to the A_3^* lattice in Figure 3.7. a, b, and c are the coefficients of the three reduced Walsh modes. The Voronoi cell is a truncated octahedron enclosed by a cube with side length π in radian, representing the originally assumed search space [187].

entire space S [188]. This Voronoi-cell-based search space is smaller than the initially assumed cube cell [187], yet it includes all possible states that the NN needs to study, containing a single maximum within it. An example of the Voronoi cell of the A_3^* lattice and the outer cube cell is illustrated in Figure 3.9.

3.5 Summary

In this chapter, we introduced Walsh modes and derived metric functions for both single and multiple modes. We showed that the single-Walsh-mode metric function is

sinusoidal and the multi-dimensional Walsh-mode metric space exhibits a lattice-like structure. For estimation, a one-dimensional algorithm can be applied to the single-mode metric function, while a NN is used for the multi-dimensional lattice-like metric space. To enable effective NN estimation, the lattice-like metric space should be confined to a finite region containing all possible states with a single maximum, the Voronoi cell. Further demonstrations of the NN approach will be presented in the next chapter.

Chapter 4

Neural Network Design for Walsh Mode Correction

Using Walsh modes for sensorless AO is challenging due to the complexity of the metric space resulting from the discrete nature. This can lead to slow convergence of the optimisation process if conventional algorithms, such as the CA and SA mentioned in the Section 2.4.3, are used. Due to the limitations of conventional algorithms, we developed a neural network (NN) to approach this optimisation problem.

4.1 Training Data Generation and Neural Network Model Construction

As mentioned in Section 2.5.1, supervised learning is ideal for online AO correction and hence is more often used for live imaging by AO microscopes. Supervised learning

normally requires a large number of labeled training datasets, which are not available if all are collected by microscope imaging. We thus developed a digital twin of a physical microscope to numerically generate the training dataset.

4.1.1 Training Data Generation Procedure

The complete data generation process is presented as a flow chart in Figure 4.1. Each training dataset consists of input-output pairs: signal readings and rough estimates are the input part, and labels are the output part.

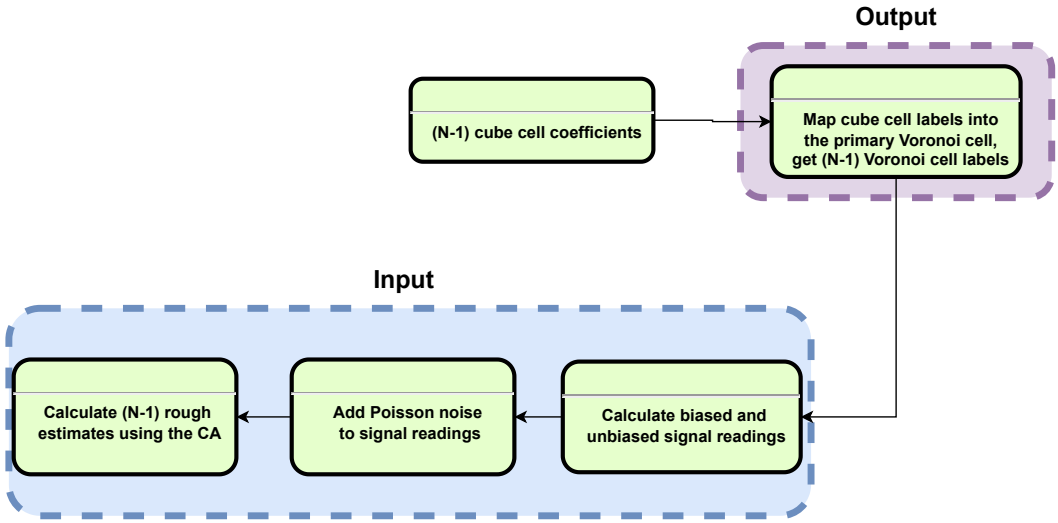


Figure 4.1: A flow chart of the training dataset generation procedure. The generation process of the output training dataset, the mode coefficients inside the primary Voronoi cell, is indicated by a purple block. The generation process of the paired input training dataset, corresponding signal readings added with Poisson noise, and CA calculated rough estimates of coefficients, is indicated by a blue block.

The two-photon simplified system model we use is the same as Figure 3.3, except the

metric function used here is the square of the Strehl ratio (SR):

$$\mathrm{SR}_0^2 = \left| \frac{\sum_{l=1}^{N} e^{j\phi_l}}{N} \right|^4 \tag{4.1}$$

N is the pixel number of the pupil phase; we choose 8 here as an example, which will be justified in Section 5. ϕ_l is the phase of each pixel. The pupil phase Ψ can be summed by Walsh modes:

$$\Psi = \sum_{k=1}^{N-1} \beta_k W_N'^k \tag{4.2}$$

$W_N'^k$ is the Walsh mode, the k^{th} row of W_N' (excluding the piston mode). β_k is the corresponding cube cell coefficient (see Figure 3.9), each of which is uniformly chosen within the range $[-\frac{\pi}{2}, \frac{\pi}{2})$. This leads to the first step of the training data generation process:

Step 1: generate $(N-1)$ cube cell coefficients β_k, $k = 1, 2, \ldots, 7$.

These cube cell coefficients are mapped into the primary Voronoi cell to get coefficient β_k', which serve as labels for training. As shown in [187], by subtracting the angle of the mean phasor of the original pupil phase Ψ and wrapping each pixel phase back within the range of $[-\pi, \pi)$, we obtain an equivalent pupil phase Ψ' but with coefficients constrained to the primary Voronoi cell. The mean phasor is calculated by averaging the complex field over the pupil phase. The Ψ' is given by:

$$\Psi' = \arg\left(e^{j\left[\Psi - \arg\left(\frac{\sum_{l=1}^{N} e^{j\phi_l}}{N}\right)\right]} \right) \tag{4.3}$$

Here, the $\arg$ operation calculates the angle of the complex field. We can then decompose the new pupil phase Ψ' and obtain Voronoi cell labels based on the orthogonality of the Walsh modes:

$$\beta_k' = \frac{\sum \Psi' * W_N'^k}{\sum W_N'^k * W_N'^k} \tag{4.4}$$

Here, $*$ denotes elementwise multiplication. Consequently, the Voronoi cell labels for NN training are obtained. The above process consists of the second step of the training data generation procedure:

Step 2: map $(N-1)$ cube cell coefficients into the primary Voronoi cell β'_k, $k = 1, 2, \ldots, 7$.

Step 3: calculate biased and unbiased signal readings. Each biased signal reading is calculated as the metric value obtained by adding a certain Walsh mode, with a coefficient of either $\pm\frac{\pi}{3}$, to the original pupil phase. Equation 4.5 illustrates this biasing process:

$$\Psi'_{k\pm} = \Psi' \pm \frac{\pi}{3} W_N^{'k} \tag{4.5}$$

$\Psi'_{k\pm}$ denotes the biased pupil phase, where Ψ'_{k+} is the mapped pupil phase with the k^{th} Walsh mode added with a coefficient of $\frac{\pi}{3}$, and Ψ'_{k-} is the mapped pupil phase with the k^{th} Walsh mode added with a coefficient of $-\frac{\pi}{3}$.

In this example with $N = 8$, each training dataset contains 14 biased signal readings (piston mode ignored):

$$\text{SR}^2_{k\pm} = \left| \frac{\sum_{l=1}^{N} e^{j\phi_l^{\pm k}}}{N} \right|^4 \tag{4.6}$$

$\phi_l^{\pm k}$ represents the phase of each pixel of the biased pupil phase, $+k$ denotes the positive biased pupil phase of Walsh mode k, and $-k$ denotes the negative biased pupil phase of Walsh mode k. $\text{SR}^2_{k\pm}$ represents the biased measurements. The unbiased measurement SR^2_0 is shown in Equation 4.1.

Step 4: add noise to signal readings. To better simulate real experiments, Poisson noise is added to the signal readings. Poisson noise represents a fundamental source

of uncertainty in signal detection, arising from the independent detection of individual photons, and its magnitude is signal-dependent [190]. In this data generation procedure, a signal level μ is randomly selected within a given interval: $\mu \in [10, 10^6]$, where μ can be interpreted as the photon number. The SNR of a Poisson-limited measurement is proportional to $\sqrt{\mu}$. Hence, choosing the photon number across such a wide range provides the NN with examples that reflect both high and low SNR experimental conditions. The noise addition is then performed as follows:

$$S_{k\pm/0} = \frac{\mathrm{Poisson}[\mathrm{round}(\mathrm{SR}^2_{k\pm/0} \times \mu)]}{\mu} \tag{4.7}$$

round operation rounds the result to the nearest integer, $\mathrm{Poisson}$ indicates the generation of Poisson-distributed random variables, and $S_{k\pm/0}$ denotes the new squared SR with noise.

Step 5: calculate rough estimates using the CA. $(N-1)$ rough estimates are also calculated to serve as a seed for the NN's learning. By learning these rough estimates obtained from the conventionally used CA (described in Section 2.4.3) together with metric-space-structure information, the NN is expected to outperform conventional algorithms. This step utilizes the signal readings from **Step 4**. For each Walsh mode, an approximate coefficient is calculated using a PSI method, as introduced in Section 3.4.1:

$$\beta''_k = -\frac{1}{2}\arctan\left(\frac{-\sqrt{3}(\sqrt{S_{k+}} - \sqrt{S_{k-}})}{-\left(\sqrt{S_{k+}} + \sqrt{S_{k-}} - 2\sqrt{S_0}\right)}\right) \tag{4.8}$$

β''_k is the calculated approximated coefficient of Walsh mode k. Then those rough estimates are also mapped into the primary Voronoi cell, following the **Step 2**. Eventually, these $(N-1)$ rough estimates also serve as inputs for NN training.

Following the steps above, a single training dataset can be created. For effective training, a large number of such datasets should be generated.

4.1.2 Neural Network Construction

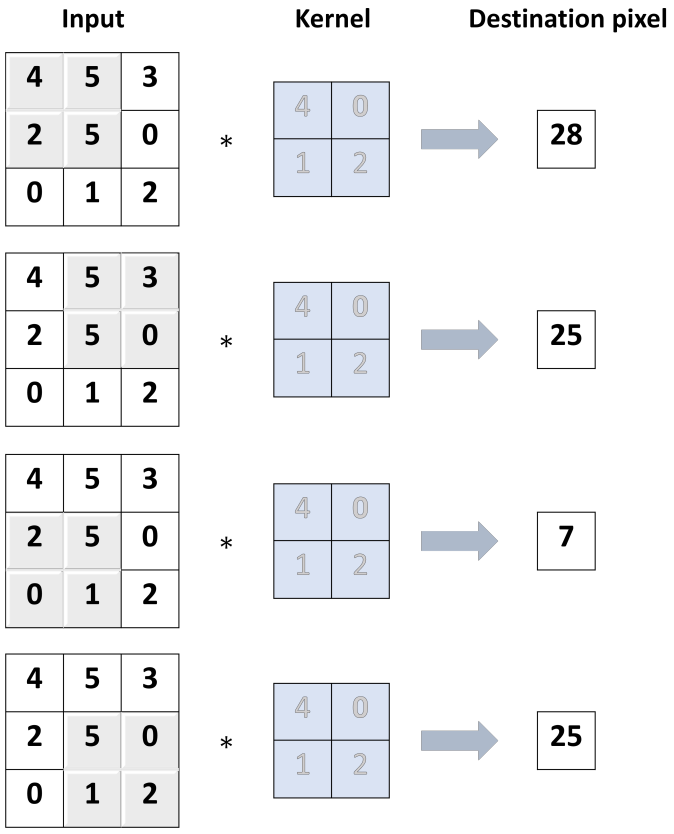
Convolutional Neural Network

Before describing the NN model construction, we first introduce an important NN architecture: the convolutional neural network.

The convolutional neural network (CNN) is one of the most widely used NNs for image processing. Its key advantage is the reduction in the number of parameters compared to traditional NNs, allowing the construction of deeper and larger models to tackle more complex problems. Three types of layers can be included in CNNs: convolutional layers (CLs), pooling layers, and fully connected layers (FCLs).

In a CL, trainable parameters are updated through learnable kernels, which significantly reduces the number of connections in the CNN and simplifies model training. A kernel can be viewed as a 'window' that slides across the input image, performing convolution with the corresponding input pixels to produce a weighted sum at each output location. During training, the kernel learns to activate in response to specific features, regardless of their spatial location in the image. Additional methods to simplify the CNN include adjusting the kernel depth, stride, and applying zero-padding [191]. An example of this convolution procedure is shown in Figure 4.2.

For the pooling layer, the primary purpose is to reduce the input dimensions of the CL (down-sampling), thereby decreasing the complexity of the CNN. Max-pooling is the



(a)

28	25
7	25

(b)

Figure 4.2: The convolution procedure in a CL. (a) A 3×3 matrix represents the input, and a 2×2 matrix represents the kernel. The kernel first convolves with the corresponding input pixels at the top-left corner to produce an output pixel. With a stride of 1, it then moves one pixel to the right and performs convolution again, followed by convolutions with the bottom-left and bottom-right input pixels. (b) Output of the convolutional process shown in (a).

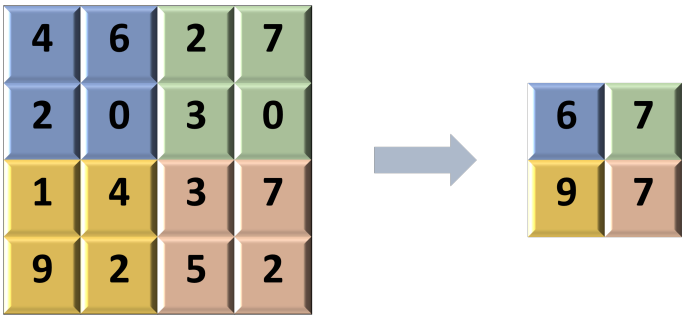


Figure 4.3: Max-pooling procedure. The blue, green, yellow, and pink blocks in the left image represent the max-pooling 'window', which selects the maximum value within each sub-region using a stride of 2. A 4×4 input is down-sampled to a 2×2 output.

most commonly used type of pooling, returning only the maximum value within the sub-region defined by the 'window'. Figure 4.3 illustrates an example of max-pooling.

FCLs are analogous to the layers in a traditional NN. A CNN architecture is illustrated in Figure 4.4. Other widely used CNN architectures, such as LeNet [192] and AlexNet [193], have also been developed and applied.

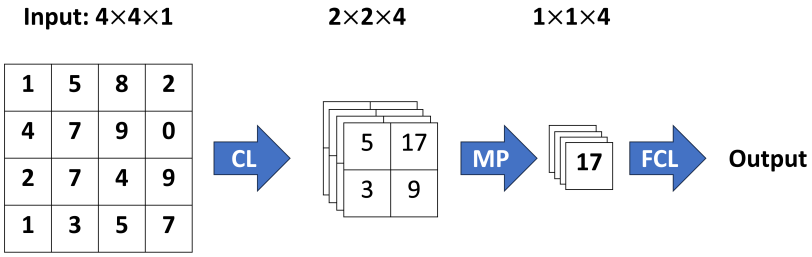


Figure 4.4: A CNN architecture. MP denotes the max-pooling layer. The input matrix is first convolved to a $2 \times 2 \times 4$ matrix, then down-sampled by the max-pooling layer to a $1 \times 1 \times 4$ matrix. The FCL then produces a single output.

Neural Network Based on Walsh Mode Metric Space

The overall architecture of the NN is shown in Figure 4.5. The network was imple-

mented using TensorFlow Keras. The input signal readings were arranged in a specific pattern to enhance feature extraction (see Figure 4.6). The CLs were specially designed, with each kernel in the first CL spanning a single row. This design allows the network to capture information from one row of the input data frame, which includes one unbiased signal reading and two biased signal readings for a given mode, to extract features from each mode's sinusoidal metric function. FCLs were employed to capture higher-order relationships by integrating information across multiple modes.

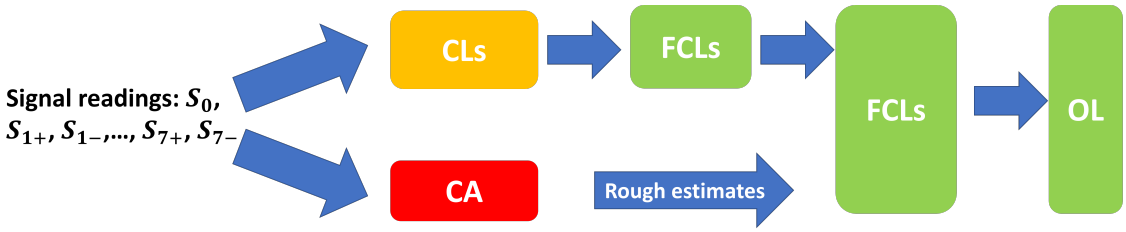


Figure 4.5: The overall architecture of the NN. CLs: convolutional layers; FCLs: fully connected layers; OL: output layer.

The detailed structure of the NN is shown in Figure 4.6. In the network, the first CL used 'valid' padding, a stride of $(1, 1)$, and the tanh activation function. The following three CLs used the same activation and stride, but their padding was set to 'same'. From the second to the fourth CL, each layer was followed by a max-pooling layer with a pool size of $(2, 1)$, which halved the data size along the rows. When both the row and column dimensions reached 1, a flattening layer was applied. The output of the CNN module used a linear activation and was then concatenated with the rough estimates before being input into the FCLs. All FCLs used tanh activations, except for the output layer, which used linear activation. We used the Glorot Uniform initialiser and Adam optimiser. The loss function used was mean squared error (MSE).

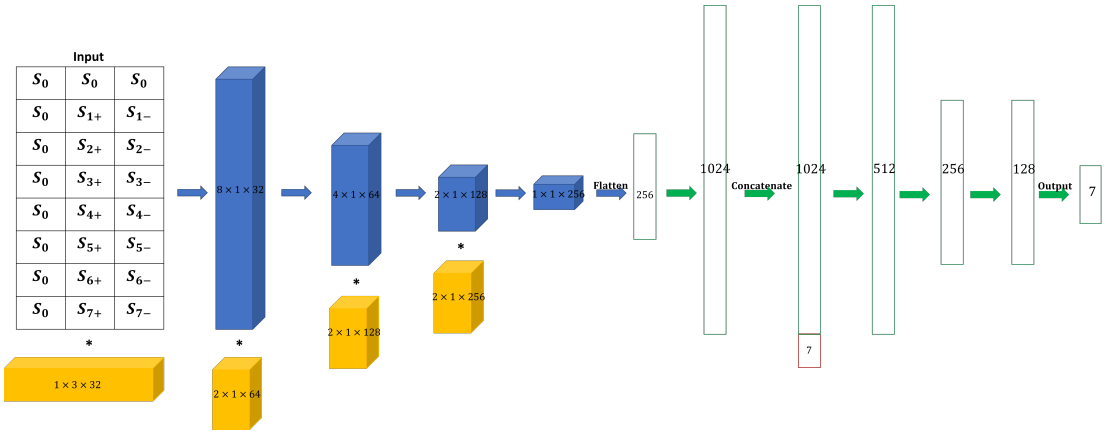


Figure 4.6: The detailed structure of the NN. Input signal readings are arranged into an $8 \times 3 \times 1$ matrix. Yellow 3D blocks represent the trainable kernels of the CLs, and blue 3D blocks represent the data after each CL. Green 2D blocks represent the data used by the FCLs, while the red 2D block represents the rough estimates. Blue arrows indicate the flow through the CLs, and green arrows indicate the flow through the FCLs.

4.2 Optimization of Neural Network Structure

The performance of a NN depends on many factors, and one key aspect is the NN architecture. Network architecture, especially the number of trainable parameters and the number of layers, is typically considered to affect the network's capability to map a complex relationship. Fewer layers or parameters simplify the network, facilitating training and reducing the risk of overfitting, whereas too many layers or parameters can introduce unnecessary complexity, increase computational cost, and potentially lead to overfitting without significant performance gains. However, an overly simple structure may limit the network's ability to model a complex relationship, resulting in large errors between the outputs and ground truths. Therefore, we investigated the trade-off between simplicity and effectiveness.

First, the number of parameters was investigated. The NN was constructed on an 8-dimensional Walsh mode set. The network architecture consisted of 4 CLs and 5 FCLs, while the number of parameters was varied. The network was trained with 2^{20} training datasets and validated with 2^{10} datasets. The signal level was chosen within the range $\mu \in [10, 10^6]$.

It's worth noting that the training process was designed to place greater emphasis on minimizing errors for small aberrations. The NN demanded a more accurate estimation for small aberrations than for larger ones, since a larger error in correcting a large aberration could be compensated by a smaller error in the next correction step. It was achieved by ensuring a higher density of training data in that region: 30% of the datasets had their 7 cube cell coefficients randomly selected within the interval $[-0.15\pi, 0.15\pi]$, while the remaining 70% had their 7 cube cell coefficients generated within the whole cube cell range $[-0.5\pi, 0.5\pi)$. Figure 4.7 shows an example comparing a NN trained with data concentrated near the cube cell center and one trained without such concentration. In this test, multiple correction rounds were applied, with all modes corrected once per round. The NN trained with this concentration performed better than the one trained without it. The result illustrates the importance of this operation for improving NN performance.

During training, we ran a total of 10 epochs with a learning rate of 10^{-3} and 256 steps per epoch. The trained NN was then tested using 2^{10} testing datasets, with the signal level identical to that used in training but without central concentration. The training loss, validation loss, and testing results are shown in Figure 4.8. After each correction round, the averaged SRs of the test datasets were improved; at the third round

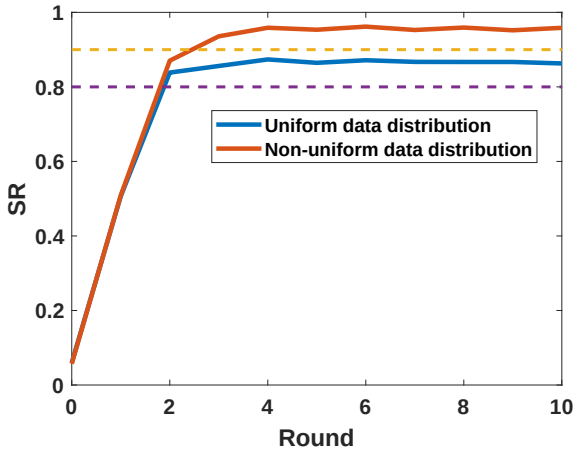


Figure 4.7: Improvement in NN performance following the redistribution of training data to concentrate on smaller aberrations. Corrected SR is plotted against correction rounds for NNs trained with and without data concentration. The NN trained with this concentration performs better than the one trained without it.

of correction, the averaged SRs were increased to over 0.9. More importantly, after each correction round, the averaged SRs for each of the NNs with specific numbers of trainable parameters were similar, and even the error bars resembled each other. These results showed that, with the same number of layers, increasing or decreasing the number of NN parameters slightly affected performance, but the improvement was not significant. A parameter count of 10^5 is selected for the NN in this study.

Next, the effect of the number of layers was examined. First, the number of CLs was varied while keeping the total number of parameters approximately constant at 10^5. The training and testing procedures were the same as described above. The results are shown in Figure 4.9. The averaged SRs were increased after each correction round for all the NNs. The averaged SRs and error bars for four and six CLs were similar across correction rounds, and at the third correction round, the averaged SRs exceeded

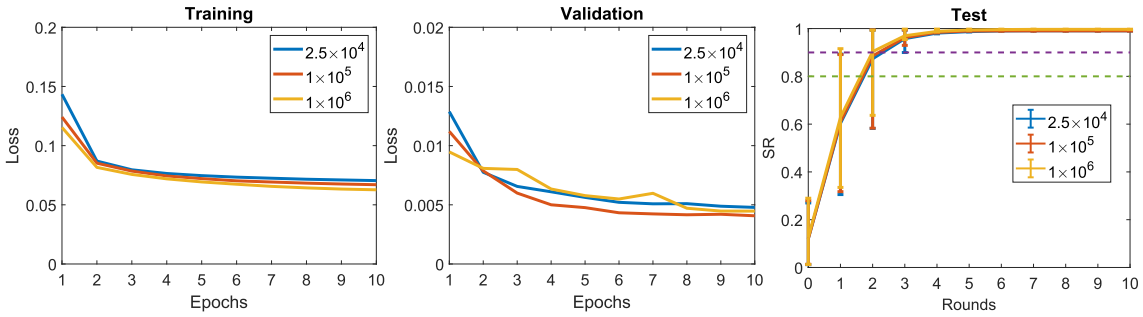


Figure 4.8: NN performance for different numbers of trainable parameters. From left to right: training loss versus epochs, validation loss versus epochs, and mean SR versus correction rounds. Labels at the top right indicate the parameter count of each NN. In the test figure, the bars represent the 10^{th} and 90^{th} percentiles of the data.

0.9. However, either the averaged SRs or the error bars after each correction round for two CLs were worse than the other two NNs: averaged SRs were lower, and error bars were larger. These results showed that adding CLs did not necessarily improve NN performance, whereas removing them reduced it. Based on this analysis, four CLs are chosen as the optimal balance between complexity and performance.

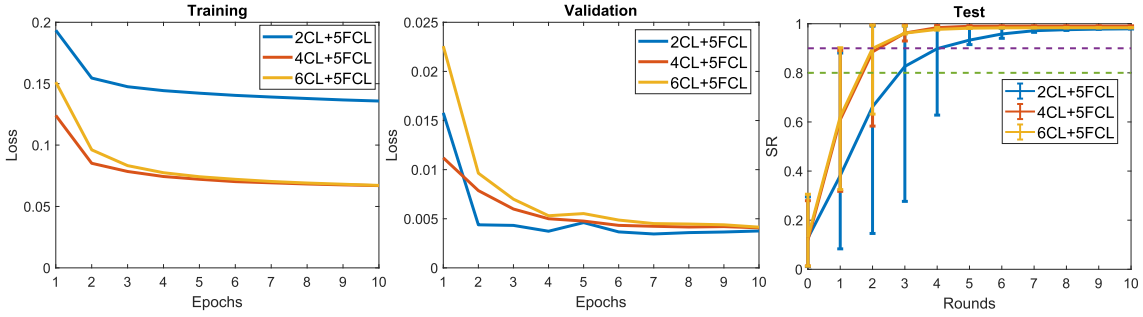


Figure 4.9: NN performance for different numbers of CLs. From left to right: training loss versus epochs, validation loss versus epochs, and mean SR versus correction rounds. Labels at the top right indicate the CLs count of each NN. In the test figure, the bars represent the 10^{th} and 90^{th} percentiles of the data.

Subsequently, the effect of the number of FCLs was tested. The training and testing

procedures remained the same, with only the number of FCLs in the NN structure being varied. The averaged SRs were increased after each correction round for all the NNs. For NNs with five and seven CLs, both the averaged SRs and the error bars were similar after each correction round, and at the third correction round, the averaged SRs exceeded 0.9 and reached nearly 1. However, the NN with three CLs, after the third correction round, the averaged SRs can merely exceed 0.9, with slightly larger error bars than the other two NNs. The results were similar to those observed for CLs (see Figure 4.10): adding FCLs did not significantly improve NN performance, whereas removing them reduced it. Based on this analysis, five FCLs are chosen as the optimal configuration for this scenario.

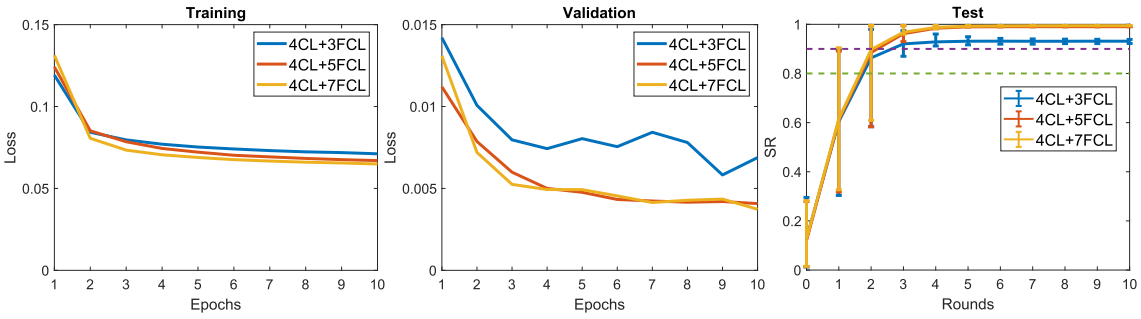


Figure 4.10: NN performance for different numbers of FCLs. From left to right: training loss versus epochs, validation loss versus epochs, and mean SR versus correction rounds. Labels at the top right indicate the FCLs count of each NN. In the test figure, the bars represent the 10^{th} and 90^{th} percentiles of the data.

During training, we observed that the validation loss consistently remained lower than the training loss. Although both losses decreased steadily and eventually stabilized—indicating proper model convergence—the validation loss persistently exhibited smaller values. We carefully examined the training procedure, datasets, and implemen-

tation details, and did not identify any anomalies or technical issues that could account for this behavior.

Given that the model demonstrated stable convergence and satisfactory performance on the test set, this discrepancy does not appear to adversely affect its overall effectiveness. Therefore, it is not considered a critical concern in the present study.

4.3 Neural Network Simplification via Subset Training

This section introduces a method to reduce the complexity of the NN structure when handling a high-dimensional Walsh mode set.

As shown in Section 4.2, a NN based on an 8-pixel Walsh mode set performs well with approximately 10^5 parameters, 4 CLs, and 5 FCLs. However, in many practical scenarios, the Walsh mode dimension N can be much higher—for example, $N = 256, 512, 1024, \ldots$ Building a NN directly on such a high-dimensional Walsh mode set results in a complex structure, computationally intensive training, and slower training speed. Table 4.1 compares several CL/FCL configurations and parameter counts for the optimal NN design at different values of N, all the configurations and parameter numbers were determined by comparing the corrected SRs for each NN configuration and deciding the optimal one. As N increases, both the number of layers and parameters required grow, the network design becomes increasingly complex, and training speed decreases. For instance, when $N = 4$, one epoch requires approximately 9 s for training,

whereas for $N = 16$, one epoch requires about 230 s.

N	CL	FCL	Parameters
4	3	2	9716
8	4	5	99336
16	6	5	3051760

Table 4.1: Number of CLs/FCLs and parameters of the optimal NN under different Walsh mode dimension N.

To address this, a practical solution is to train a low-dimensional NN on a subset of Walsh modes and iteratively correct the remaining modes using the same NN, as illustrated in Figure 4.11. The detailed working process can be seen in Section 5.2.

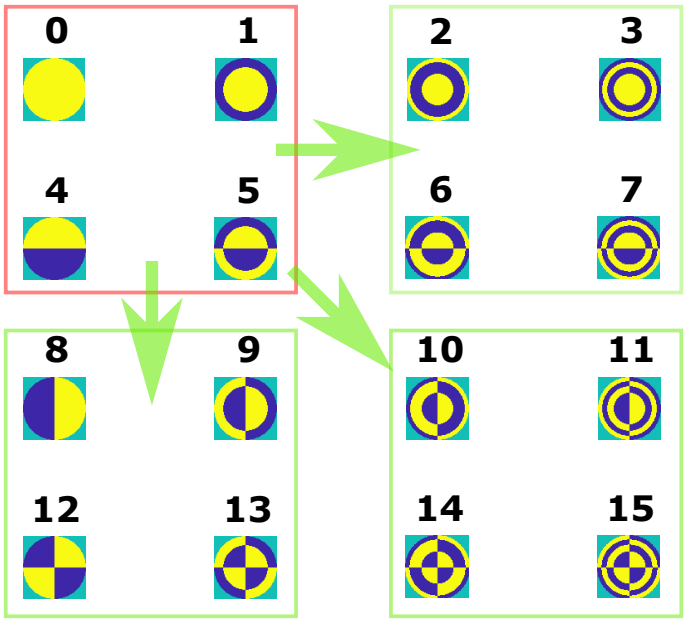


Figure 4.11: Simplified NN working strategy using subset training. For $N = 16$, the NN is trained on a polar Walsh mode subset within the red box. Once trained, the NN can be applied to the subsets indicated by the green boxes as well as to itself. By sequentially applying the trained NN to all subsets, the entire Walsh mode set is corrected.

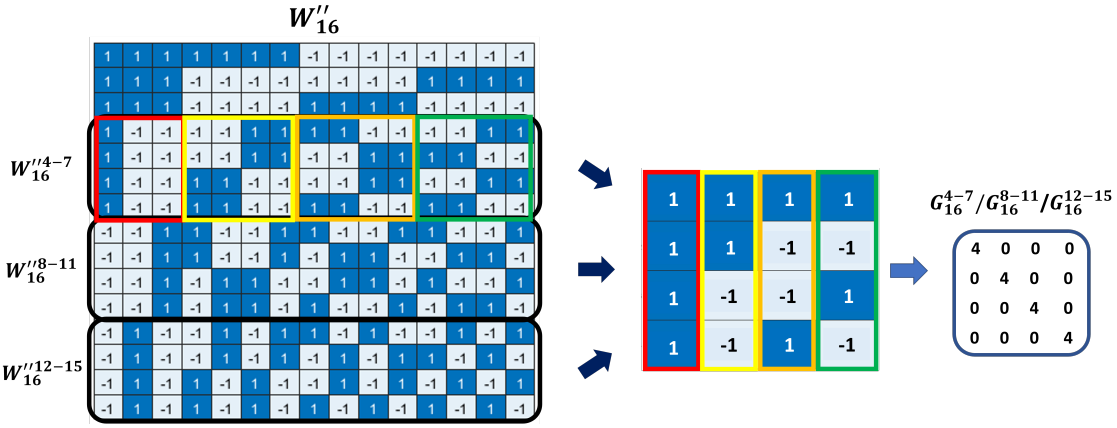


Figure 4.12: Generation process of Gram matrices for Walsh mode subsets. The left matrix is divided into three subset matrices (divided by the black boxes). One basis vector is selected from each category (indicated by red, yellow, orange, and green) to form a 4×4 generator matrix (the middle matrix), and this is done iteratively for all three subset matrices. Using these generator matrices, the Gram matrices G_{16}^{4-7}, G_{16}^{8-11}, and G_{16}^{12-15} are calculated (right matrix). The fact that all Gram matrices are identical indicates that the lattices generated by these Walsh mode subsets are the same.

This can be illustrated with the following example. A W''_{16} matrix is generated, and subsets W''^{4-7}_{16}, W''^{8-11}_{16}, and W''^{12-15}_{16} are used to construct lattices, from which the Gram matrices G_{16}^{4-7}, G_{16}^{8-11}, and G_{16}^{12-15} are obtained. Here, W''^{n-m}_{16} refers to the Walsh mode subset matrix containing rows $(n-1)$ to $(m-1)$ of W''_{16}, as indicated by the black boxes dividing W''_{16} in Figure 4.12. Since the subset matrix has more columns than its rank, some redundant columns are removed. For example, in Figure 4.12, the columns of the subset matrix represent the lattice basis vectors. Each colored square corresponds to a category of basis vectors. Vectors within the same category are either identical or reflections (reflection denotes identical values with opposite signs) [194]. However, the basis vectors are defined to be orthogonal to each other, which is achieved between

vectors from different categories. In this case, one vector from each category can be selected, others can be removed, and the four chosen vectors form a new generator matrix, see the middle matrix of Figure 4.12. The Gram matrix is then calculated from this generator matrix. This Gram matrix generating process is repeated for each Walsh mode subset. Consequently, G_{16}^{4-7}, G_{16}^{8-11}, and G_{16}^{12-15} are obtained.

As shown in Figure 4.12, G_{16}^{4-7}, G_{16}^{8-11}, and G_{16}^{12-15} are all identical, indicating that the lattice structures are the same. This means that a NN can be trained on one Walsh mode subset (for example, $W_{16}^{\prime\prime 4-7}$, which relates to the Gram matrix G_{16}^{4-7}) and applied to the other subsets (correcting $W_{16}^{\prime\prime 4-7}/W_{16}^{\prime\prime 8-11}/W_{16}^{\prime\prime 12-15}$ of the whole $W_{16}^{\prime\prime}$ set, which are related to Gram matrices $G_{16}^{4-7}/G_{16}^{8-11}/G_{16}^{12-15}$). This strategy will be extensively used in the simulations presented in the next chapter.

4.4 Summary

In this chapter, the complete data generation procedure was described in detail. The construction of the NN for an 8-pixel Walsh mode set was presented, and the preferred network structure for this application was identified: approximately 10^5 parameters, 4 CLs, and 5 FCLs. Finally, a method to reduce NN complexity was proposed by iteratively correcting high-dimensional Walsh mode sets using a NN trained on a subset. Further simulations using this NN will be discussed in the following chapter.

Chapter 5

Simulations of Neural-Network-Based Correction of Walsh Modes

Neural networks (NNs) need to be trained with sufficient training datasets that, ideally, encompass the whole searching space. This means that the amount of training dataset increases exponentially with the Walsh mode dimensions that were involved in the discussion: if too many Walsh modes are approached and corrected altogether by one NN, the data generation and training process will become unmanageable.

To understand how to divide the Walsh mode set into subsets to ensure an efficient and effective correction, we systematically discussed by simulating different NNs, each trained with a different subset, and compared their performances. We also compared, through simulation, the optimised NN setting with commonly used CA and SA.

5.1 Optimal Walsh Mode Subset Dimension for Neural Network Design

When dealing with a high-dimensional Walsh mode set, as discussed in Section 4.3, a smaller Walsh mode subset is required for both training and correction. In this section, we investigate the optimal subset size for achieving effective NN-based estimation.

To address this, NNs were trained with Walsh mode dimensions $K = 4, 8, 16$, referred to as 4NN, 8NN, and 16NN, respectively. The solution used here will be introduced in detail in Section 5.3.3. The signal level was defined within the interval $\mu \in [10, 10^6]$ (see Section 4.1.1). All NNs were trained on 2^{23} training datasets and validated on 2^{10} datasets, with central concentration applied to improve correction of small aberrations, as discussed in Section 4.2.

The trained NNs were subsequently tested on Walsh mode sets with dimensions $N = 64, 128$, which were sufficiently high to evaluate performance differences among the NNs. The testing dataset consisted of 2^{10} samples with the same signal level range as the training data but without central concentration. For a correction subset of dimension K, $\frac{N}{K}$ iterations (indicated by i) were required to complete one full correction round of all modes. A total of $D = 10$ correction rounds were performed in these demonstrations. The overall correction process is illustrated in the flow chart shown in Figure 5.1, using $K = 16$ and $N = 64$ as an example for clarification.

Figure 5.2 shows how the average SR increases with the number of images collected for different NNs under different Walsh mode dimensions. As benchmarks, corrected

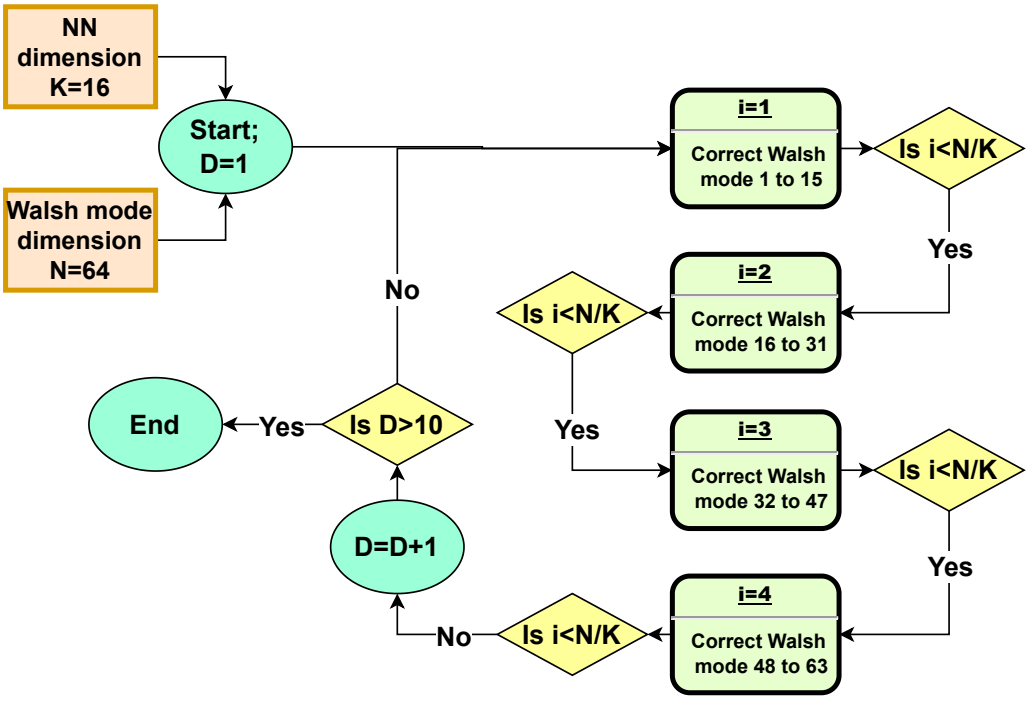


Figure 5.1: Correction process for NN with dimension $K = 16$ and Walsh mode dimension $N = 64$. A total of $D = 10$ correction rounds are performed. Each correction round consists of $\frac{N}{K}$ iterations, which is indicated by i.

SR values of 0.6 and 0.8 were chosen. When $N = 64$, 8NN required a similar or smaller number of measurements than 16NN to reach an average SR of 0.6 or 0.8, while 4NN could not achieve these values. The final average SR was the highest for 8NN. When $N = 128$, 8NN required a similar number of measurements as 16NN to reach an average SR of 0.6; at an average SR of 0.8, only 8NN reached this benchmark. Again, the final average SR was the highest for 8NN. These results indicated that 8NN was the optimal choice for these scenarios.

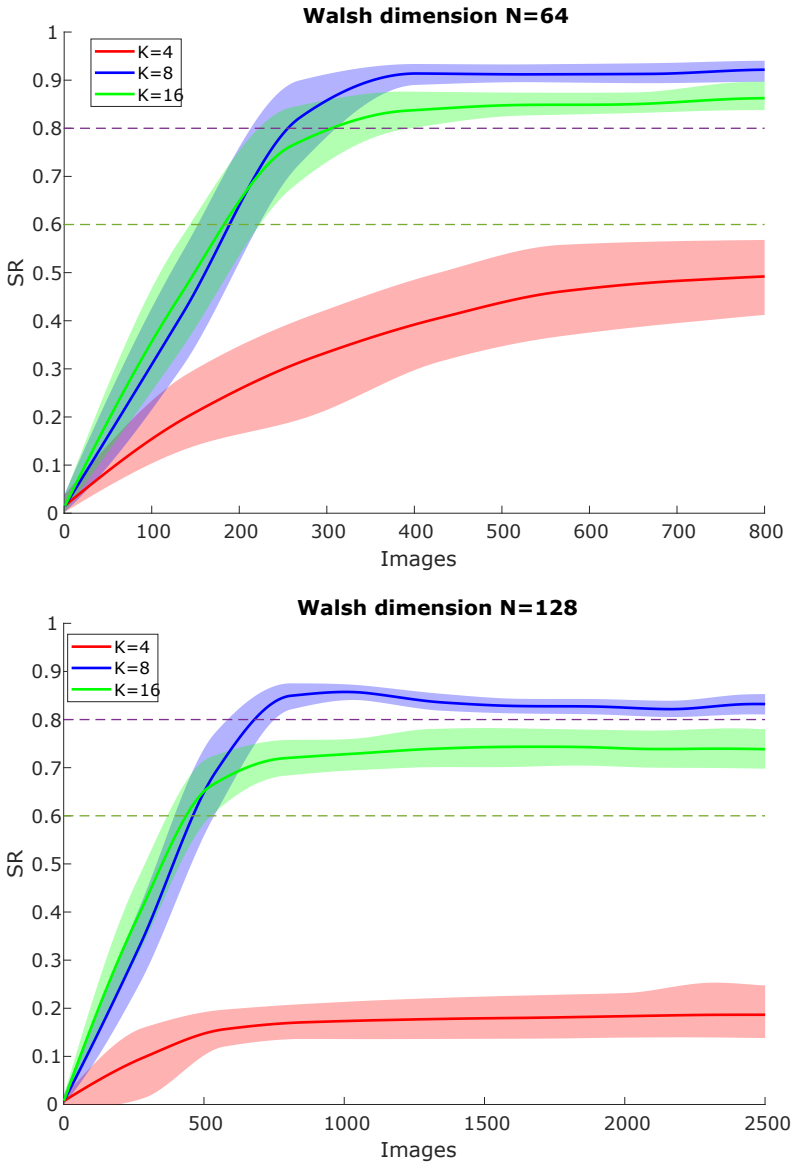


Figure 5.2: Comparison of Walsh mode subset dimensions (K) for NN design. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different NN dimensions K (labels in the top left) and Walsh mode dimensions N (indicated in the panel headings). Shaded regions represent the $10^{\text{th}} - 90^{\text{th}}$ percentiles of SR across samples.

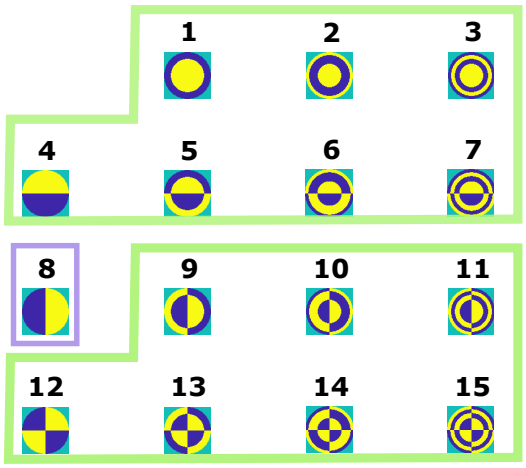
A subset dimension of $K = 4$ is not chosen because it represents too small a fraction of the overall N modes to be effective. In contrast, a NN trained with $K = 8$ provides suitable performance without excessive computational cost. Although 16NN can also work, its high subset dimension makes the network construction computationally demanding and the training time long. Therefore, since $K = 8$ is the optimal Walsh mode subset dimension, all the NNs in the subsequent simulations are trained on this dimension.

5.2 Neural Network Working Procedure

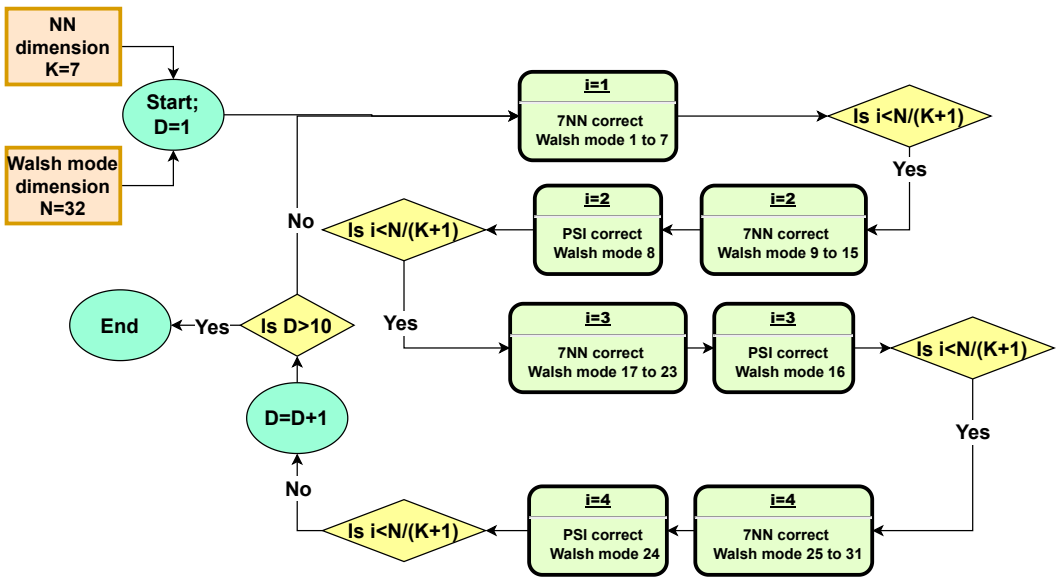
This section describes the detailed working procedure of the NN we built in Chapter 4.

The NN was originally designed to consider the lowest-order Walsh modes. Since the piston does not affect the measurement, the network was trained on the next seven modes, from index 1 to 7 (hereafter referred to as $W1$–7 for clarity; all subsequent Walsh mode subsets follow the same notation). However, the presence of the piston creates an anomalous range: for higher-order groups of eight modes (for example, $W8-15$), the lowest-order mode in the group, Walsh mode 8 (hereafter referred to as $W8$ for clarity; all subsequent Walsh modes follow the same notation), does influence the signal. This process is illustrated in Figure 5.3(a). Therefore, an additional step is required in the operation of this 7NN to make it effective for higher-order mode groups.

For a Walsh mode set of dimension N, in each correction round D, there are $\frac{N}{8} - 1$ modes that are not corrected by the 7NN. For example, when $N = 32$, three modes are



(a)



(b)

Figure 5.3: (a) The anomalous phenomenon: the NN is trained on $W1-7$, with the piston mode removed. When applied to the subset of $W8-15$, $W8$ is omitted, requiring an additional estimation algorithm. (b) The workflow of 7NN + PSI corrects the Walsh mode set with dimension $N = 32$. Total correction rounds $D = 10$.

omitted in each round.

In this case, a PSI method correction step is added to correct the omitted mode in each iteration, as described in Equation 3.17 in Section 3.4.1. Since the ignored mode during 7NN training is the piston (lowest-order) mode, we apply the PSI measurement to the lowest-order Walsh mode in each group, namely $W8$, $W16$, and $W24$. In every iteration, after the standard 7NN prediction, the PSI step is applied to the designated modes. This ensures that no uncorrected modes remain in the Walsh mode set. This approach is referred to as 7NN + PSI. At the end of the correction round, 7NN + PSI collects $15 + 18 \times (\frac{N}{8} - 1)$ images, 15 images in the first iteration when correcting the first 7 Walsh modes without piston and 18 images in the rest of the iterations when correcting 7 Walsh modes (15 images) and one extra mode with PSI (3 images). The above 7NN + PSI application example is illustrated as a workflow in Figure 5.3(b).

5.3 Simulations

This section presents simulation results of NNs tested under varying SNR conditions and Walsh mode set dimensions, with comparisons to CA and SA introduced in Section 2.4.3.

5.3.1 Comparison of 7NN + PSI and CA under Varying Signal-to-Noise Ratios

To evaluate the effect of 7NN + PSI under different SNR conditions, its performance on $W8 - 15$ was tested at various SNRs and compared with that of the CA.

Testing datasets consisted of 2^{10} samples, not centrally concentrated. The photon number μ was selected from $\mu \in [10, 10^6)$, corresponding to the training SNR range of SNR $\in [3, 1000)$. High, medium, and low SNR conditions were defined as SNR $\in [300, 1000)$, SNR $\in [30, 100)$, and SNR $\in [3, 10)$, respectively. Figure 5.4 shows

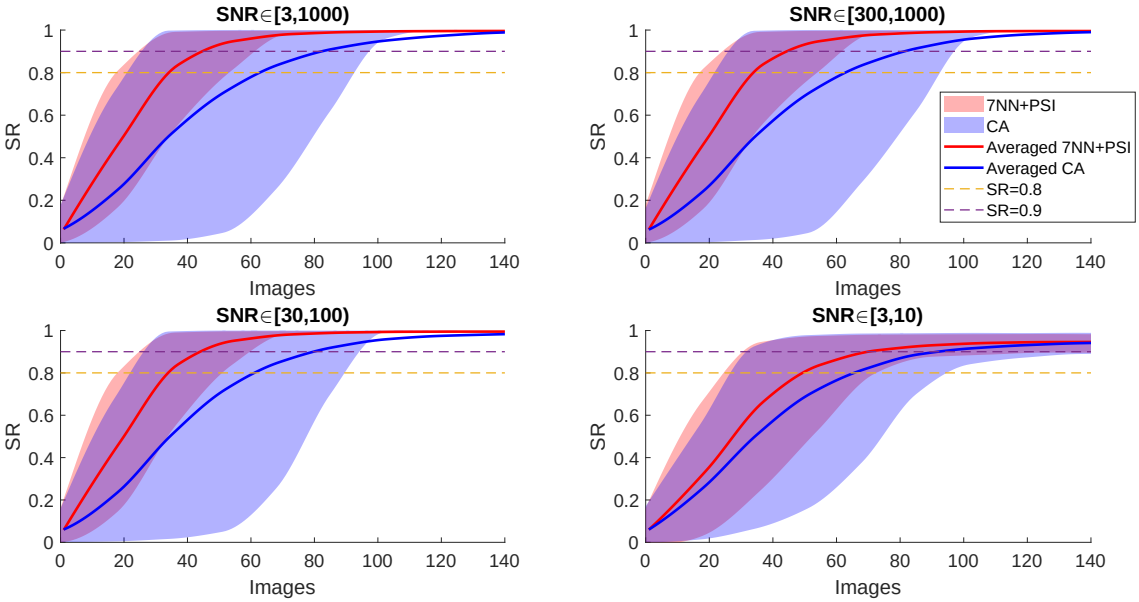


Figure 5.4: Comparison of the performance of 7NN + PSI and CA correcting $W8 - 15$ under various SNR conditions. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms (labels in the top right) and SNRs (indicated in the panel headings). Shaded regions represent the $10^{\text{th}} - 90^{\text{th}}$ percentiles of SR across samples.

the average corrected SR as a function of the number of images collected, with a comparison to the CA. The CA was applied to the same datasets as the NN method, measuring all 8 Walsh modes' metric values and using PSI to calculate the target mode coefficients (see Section 2.4.3). In Figure 5.4, 7NN + PSI consistently required the fewest images to reach benchmarks—average SR values of 0.8 and 0.9. The number of images needed for 7NN + PSI remained similar across SNR conditions. For the final average SR, even with low SNR, the 7NN + PSI could always reach final average SRs over 0.9.

The above analysis indicates that 7NN + PSI performs better than CA in efficiency on each SNR occasion, and 7NN + PSI is robust among SNRs.

5.3.2 Comparison of 7NN + PSI and CA under Varying Walsh Mode Dimensions

To evaluate the performance of 7NN + PSI across different Walsh mode set dimensions, both 7NN + PSI and CA were applied to Walsh mode sets with $N = 32, 64, 128$ with the same SNR during the training process. It is worth mentioning that 7NN + PSI collected and corrected subsets of the full Walsh mode set, whereas CA processed the entire set without updating the 'zero-bias' pupil (see Section 2.4.3).

The testing dataset consisted of 2^{10} samples, with SNR matching the training range (SNR $\in [3, 1000)$) and without central concentration. The resulting average corrected SR curves are shown in Figure 5.5. The results indicate that 7NN + PSI outperforms CA. This is because CA attempts to correct too many modes at once without updating the pupil function.

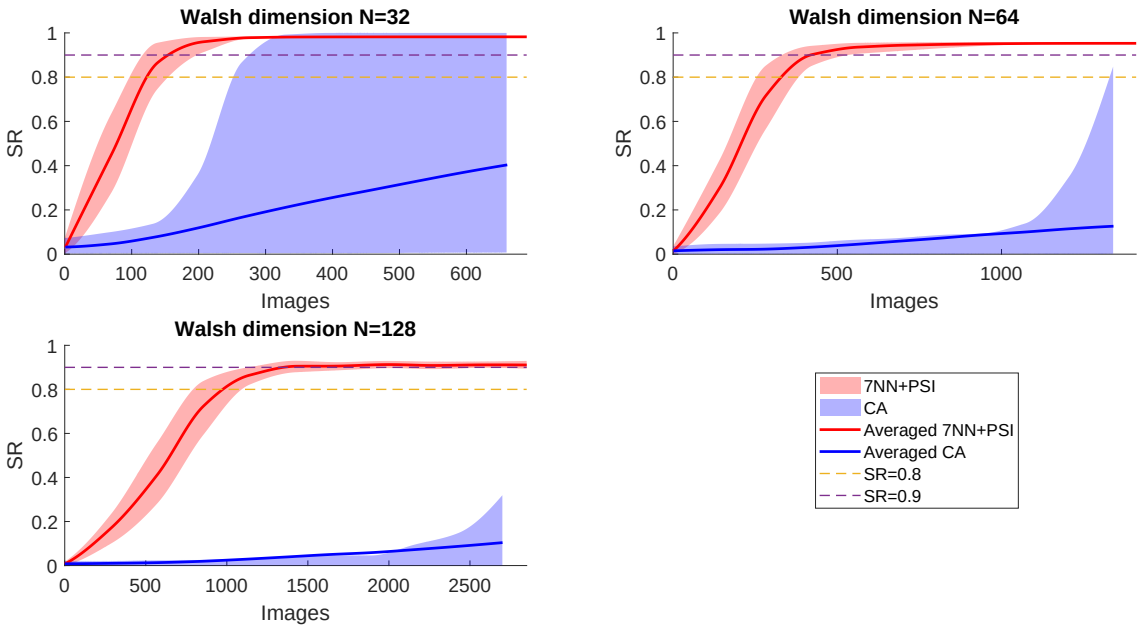


Figure 5.5: Comparison of the performance of 7NN + PSI and CA correcting various Walsh mode dimensions N under the same SNR condition (CA applied to the whole Walsh mode set of $N = 32/64/128$). The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms (indicated in labels) and dimension N (indicated in the panel headings). Shaded regions represent the $10^{\text{th}} - 90^{\text{th}}$ percentiles of SR across samples.

Therefore, the CA was modified to follow the same correction process as 7NN + PSI: modes were grouped into subsets of 8, and in each iteration, 8 modes (7 in the first iteration) were corrected. A total of $\frac{N}{8}$ iterations completed one correction round.

The resulting average SR curves versus the number of collected images are shown in Figure 5.6. When $N = 32$, 7NN + PSI required a smaller number of images than CA to meet the criteria with a final average SR result near 1. When $N = 64$, either the number of images required to meet the SR criteria or the final average metric values were similar for 7NN + PSI and CA. However, when $N = 128$, the number of images needed to meet

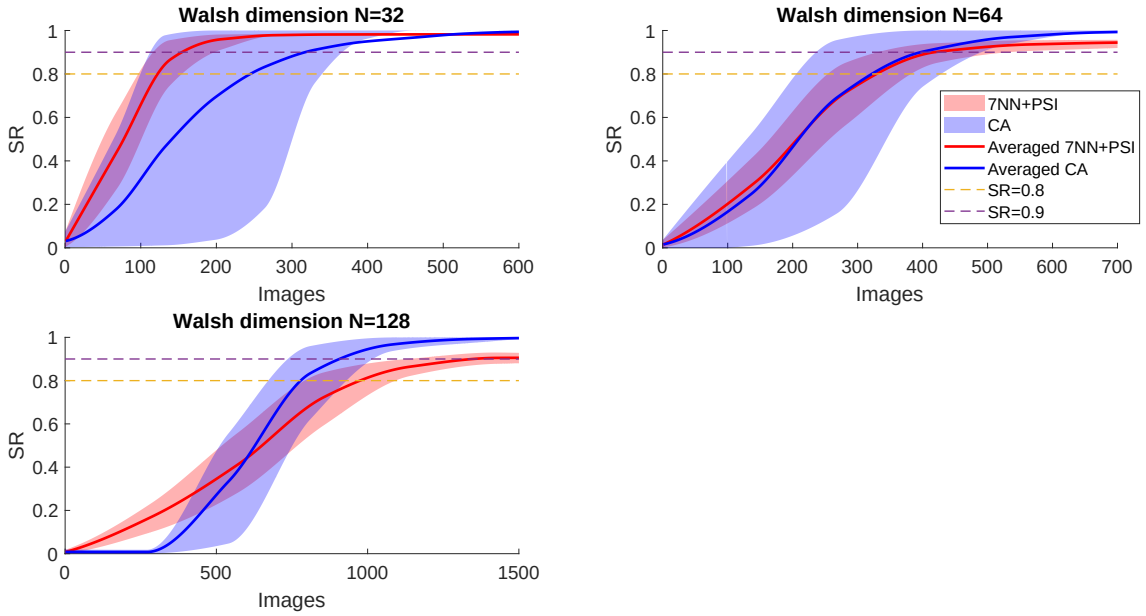


Figure 5.6: Comparison of the performance of 7NN + PSI and CA correcting various Walsh mode dimensions N under the same SNR condition (CA applied to the Walsh mode subsets of 8 modes in each correction iteration). The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms (labels in the top right) and dimension N (indicated in the panel headings). Shaded regions represent the $10^{\text{th}} - 90^{\text{th}}$ percentiles of SR across samples.

the SR criteria for 7NN + PSI was larger than that of CA, and the final average SR value was smaller; here, the CA showed better performance. One possible explanation for the poor performance of the 7NN + PSI is that the extra PSI method limited its efficiency, which will be further explored in Section 5.3.3.

In conclusion, when the Walsh dimension N is low (lower than 128), 7NN + PSI outperforms (or is similar to) CA in efficiency. The performance of CA is strongly influenced by its dimensionality: a low-dimensional CA can achieve better performance than a high-dimensional CA. Moreover, the Walsh mode dimension is an important property

that should be considered when applying 7NN + PSI: when the Walsh mode dimension is low, 7NN + PSI can perform better than a same-dimensional CA (which can correct the same number of Walsh modes as 7NN + PSI in one iteration) in efficiency.

5.3.3 Modification of Neural Network Structure

The NN's input data frame structure is modified here, which improves the NN's working efficiency.

The NN was initially designed with a dimension of 7 to correct the mode set $W1-7$, excluding $W0$ (piston), since it does not affect the system. However, when applied to higher-order groups of eight modes (for example, $W8-15$, $W16-23$), the 7NN requires an additional PSI step to handle the lowest-order mode in each group, which may limit its working efficiency. To address this, we investigated the use of a NN trained directly on a complete set of eight modes.

An 8NN was therefore trained and tested on the mode set $W8-15$. The training data generation procedure was adapted from the 7NN described in Section 4.1.1, with the pupil phase generation Walsh mode set being modified accordingly. Each training dataset contained 17 input signal readings, 8 rough estimates, and 8 labels. The 17 input signals were arranged in the same manner as in the 7NN (see Figure 5.7). The 8NN architecture consisted of 4 CLs and 5 FCLs, with 99336 trainable parameters, mirroring the structure of the optimized 7NN from the previous chapter. The training set contained 2^{23} samples, while the validation set contained 2^{10} samples. Both were centrally concentrated, with SNR $\in [3, 1000)$.

After training, the 8NN was tested on 2^{10} datasets generated in the same way as the training data but without central concentration. As shown in Figure 5.8, while tested on Walsh mode set $W8 - 15$, the 8NN performs effectively, achieving a final average SR close to 1.

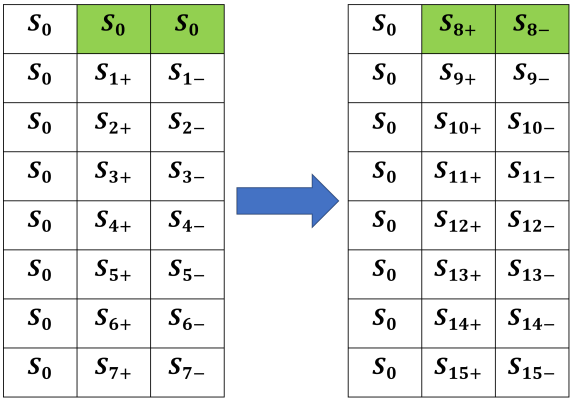


Figure 5.7: Training data signal reading arrangement for 8NN (right), adapted from the arrangement shown in Figure 4.6 for 7NN (left). The similar pattern is applied to all other Walsh mode subsets when estimating their coefficients, except the lowest-order subset.

A special case arises when applying this NN to the correction mode set $W0-7$. Since the piston mode does not influence the signal readings, the NN will otherwise attempt to estimate a value even though the piston coefficient should be arbitrary. To address this, we force the input measurements for S_{0+} and S_{0-} to be $0.0625S_0$, see Figure 5.9. These values are chosen such that, according to Equation 3.17, the NN may return a coefficient of 0 for the piston mode. These 'phantom' signal readings are incorporated into the input signal readings' data frame during the correction of $W0 - 7$. In this way, the NN predicts the piston coefficient as 0 (or negligible) while estimating the other Walsh modes normally.

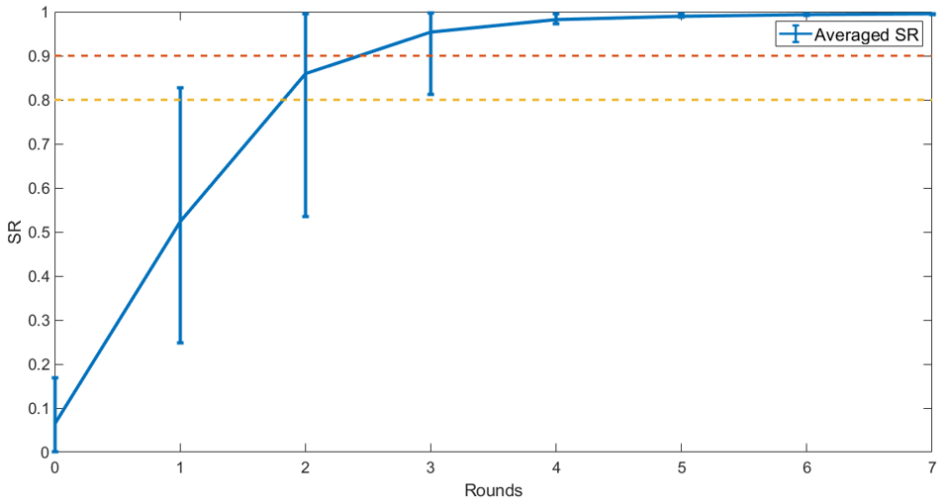


Figure 5.8: Correction results of 8NN on $W8 - 15$ with SNR matching the training range. The plot shows the evolution of the average SR over correction rounds. Error bars indicate the 10^{th}–90^{th} percentiles of the samples. Yellow and red dashed lines mark SR $= 0.8$ and SR $= 0.9$, respectively.

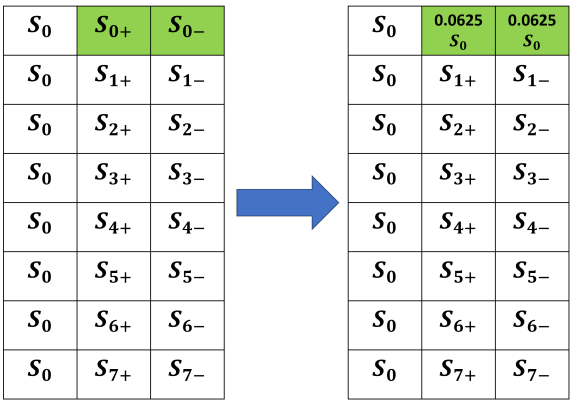


Figure 5.9: Input signal reading arrangement for 8NN (right) while correcting the lowest order Walsh mode subset $W0 - 7$.

5.3.4 Comparison of 7NN + PSI with 8NN

To compare the working efficiency of 7NN + PSI and 8NN, two sets of tests were conducted, focusing on variations in SNR and Walsh mode dimension.

First, 7NN + PSI and 8NN were tested on $W8-15$ using 2^{10} non–centrally concentrated testing datasets across different SNRs. The results, shown in Figure 5.10, indicate that 8NN outperforms 7NN + PSI: under every SNR condition, while achieving similar final average SRs, 8NN consistently required fewer images to meet the correction criteria.

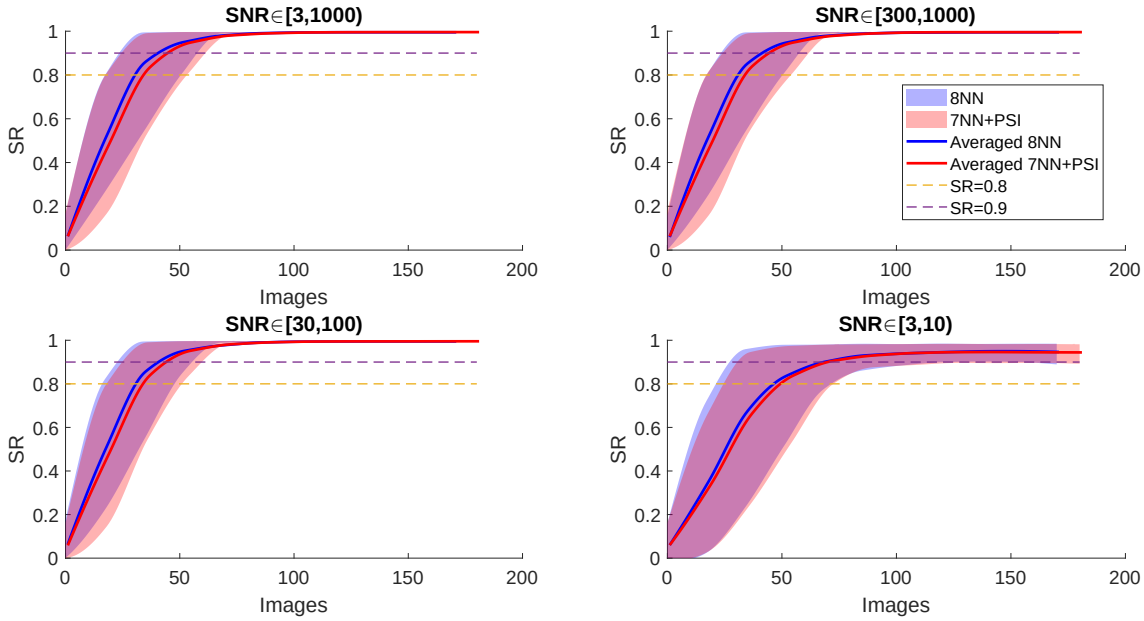


Figure 5.10: Comparison of the performance of 7NN + PSI and 8NN correcting $W8-15$ under various SNR conditions. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms (labels in the top right) and SNRs (indicated in the panel headings). Shaded regions represent the $10^{\text{th}}-90^{\text{th}}$ percentiles of SR across samples.

Next, 7NN + PSI and 8NN were tested on Walsh mode sets of varying dimensions $N = 32, 64, 128$, under the same SNR during the training process. The testing datasets contained 2^{10} samples, were not centrally concentrated. The results, shown in Figure 5.11, demonstrated that across all Walsh mode dimensions, while achieving similar final

average SRs, 8NN consistently required fewer images to meet the correction criteria.

These findings confirm that 8NN achieves higher working efficiency than 7NN+PSI. Consequently, all subsequent comparison simulations are conducted using 8NN.

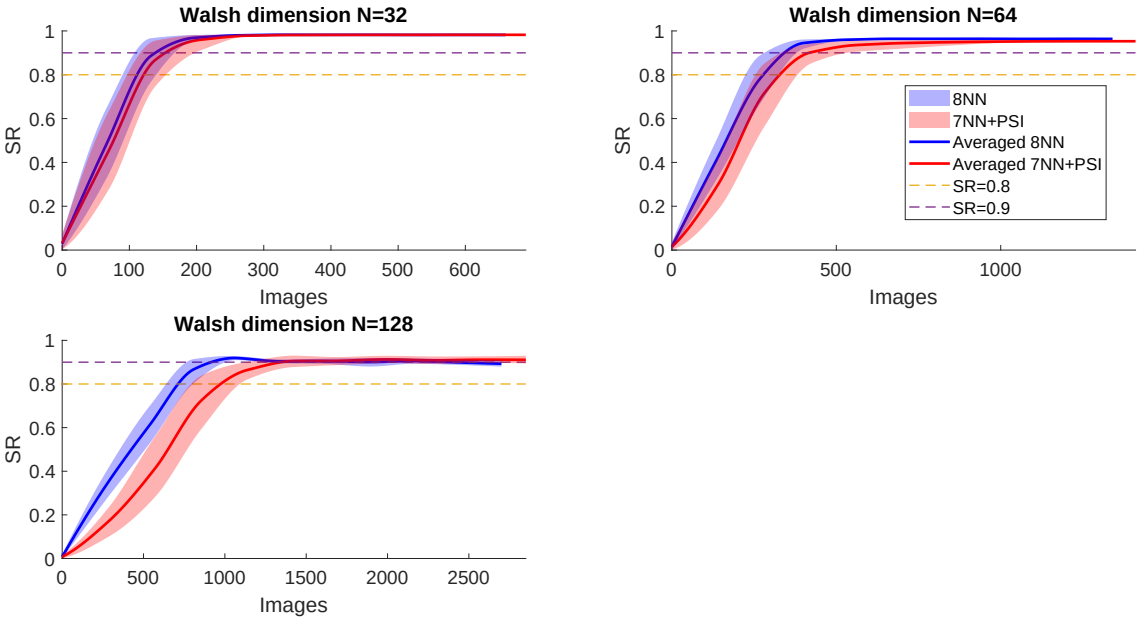


Figure 5.11: Comparison of the performance of 7NN + PSI and 8NN correcting various Walsh mode dimensions N under the same SNR condition. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms (labels in the top right) and dimension N (indicated in the panel headings). Shaded regions represent the $10^{\text{th}} - 90^{\text{th}}$ percentiles of SR across samples.

5.3.5 Comparison of 8NN with SA

Previous tests showed that lower-dimensional CAs achieve better working efficiency. Accordingly, the SA introduced in Section 2.4.3 represents the optimal CA in image quality improvement efficiency. In SA, each Walsh mode is measured and corrected sequentially, with the pupil function updated after correcting each mode. To provide a

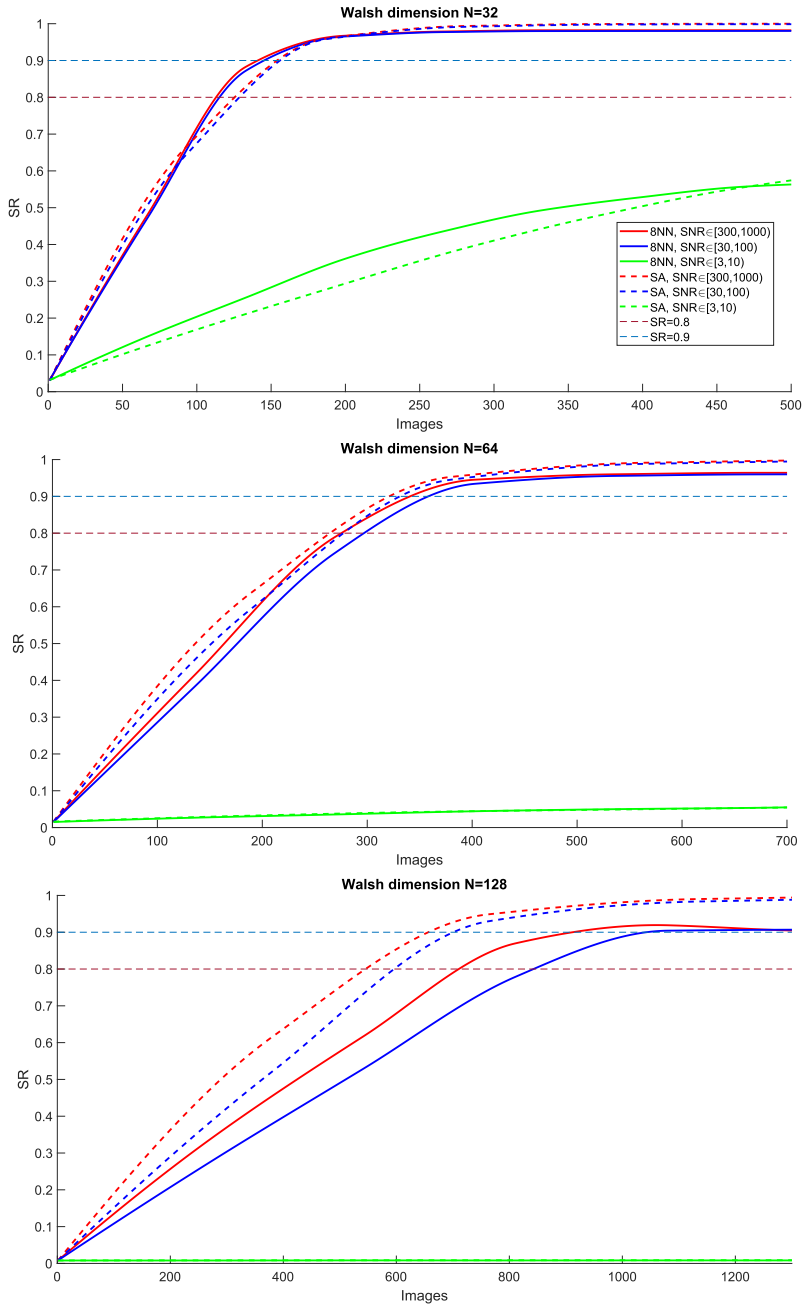


Figure 5.12: Comparison of the performances of 8NN and SA correcting various Walsh mode dimensions N under varying SNRs. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms, SNRs (labels in the top right), and dimension N (indicated in the panel headings).

more rigorous comparison, 8NN was evaluated against SA across different SNR levels and Walsh mode dimensions.

The testing dataset size was 2^{10} (not centrally concentrated); other parameters are indicated in Figure 5.12 as legends and panel headings. For high and medium SNRs (SNR $\in [300, 1000)$/SNR $\in [30, 100)$) and low Walsh dimensions ($N = 32, 64$), 8NN performs similarly to SA. Both algorithms reached benchmarks SR $= 0.8/0.9$ with nearly the same number of images (with 8NN reaching the criteria with slightly fewer images for $N = 32$), and the final average SRs were comparable.

The reason 8NN can only perform to a similar standard as SA is that the formula generating testing datasets, which is described in Section 4.1.1 Equation 4.8, is based on the same calculation procedure to the calculating formula used by SA (see Equation 3.17), so SA is more suitable in this condition. Therefore, we should not expect a NN estimator to perform better than an optimal analytic model.

5.3.6 Summary

The SA is nearly optimal in idealized simulation conditions, making it difficult for the NN to outperform. However, there are scenarios in which 8NN demonstrates superior performance compared to SA within simulations.

We employed a modified NN structure, denoted as 8NN*. This 8NN* consists of 4 CLs and 4 FCLs with 385480 parameters, simplifying the network by reducing layers and parameters to improve performance under low SNR conditions. It was trained on $W8-15$ with SNR $\in [3, 1000)$, and then tested on $W1-31$ across the SNR $\in [300, 1000)$,

SNR $\in [30, 100)$ and SNR $\in [3, 10)$, as shown in Figure 5.13. At high and medium SNR conditions (SNR $\in [300, 1000)$/SNR $\in [30, 100)$), 8NN* reached SR $= 0.8/0.9$ requiring a smaller number of images than or similar to that of SA. Under low SNR conditions (SNR $\in [3, 10)$), 8NN* corrected faster and achieved a higher final average SR than SA. This demonstrates that a NN can potentially surpass a nearly optimal algorithm in simulations like SA.

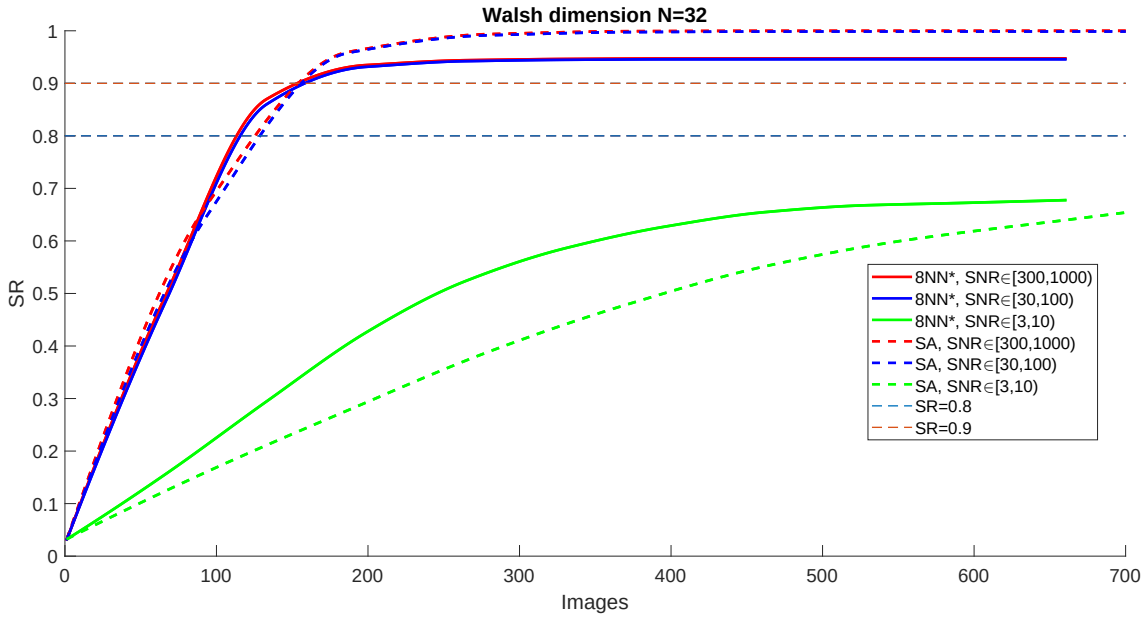


Figure 5.13: 8NN* (simplified 8NN) outperforms the SA in low SNR conditions, correcting Walsh mode set with dimension $N = 32$. The corrected average SR is shown as a function of the number of collected images (interpolated curves) for different algorithms and SNRs (indicated in labels).

Chapter 6

Two-Photon Microscopy Demonstrations of Neural-Network-Based Correction of Walsh Modes

In this chapter, the trained 8NN was implemented in a two-photon microscope to evaluate its efficiency in comparison with SA. Results verify the simulation results in Chapter 5 that the NN we trained can outperform SA in working efficiency.

6.1 System Configuration

This section describes the two-photon microscope (2PLSM) setup used in experiments. The setup consists of three main components: a two-photon illumination path, a signal detection path, and an adaptive optics path for aberration correction. The full

setup schematic plot is shown in Figure 6.1.

6.1.1 Two-Photon Illumination Path

This path is indicated in red in Figure 6.1. The 2PLSM system uses a femtosecond pulsed laser source. The pump laser is a 532 nm continuous-wave diode-pumped solid-state (DPSS) laser (Millennia eV^{TM}, Spectra-Physics), which pumps a Ti:sapphire oscillator (Tsunami, Spectra-Physics). The Ti:sapphire laser output is linearly polarized, tuned to a central wavelength of 870 nm, with a pulse width of 182 fs and a beam diameter of 1.3 mm.

Along this path, the main optical axis is redirected by four mirrors, M1–M4 (BB1-E03, Thorlabs), which will not be restated in the subsequent descriptions. For laser safety, a shutter S (SH05/M, Thorlabs) is placed immediately after M1. Following this, a power-adjusting module is implemented, consisting of a half-wave plate HWP (AHWP10M-980, Thorlabs), a polarizing beam splitter PBS (GL5-B, Thorlabs), and a beam trap (BT610/M, Thorlabs). The rotating HWP adjusts the polarization of the incident laser, and together with the PBS, enables control of the transmitted light intensity. The portion of light reflected by PBS is absorbed by the beam trap.

After passing through the power-adjusting module, the attenuated beam enters the first 4f system, which consists of two lenses, L1 and L2 (45806 and 49362, Edmund Optics), arranged with coincident focal planes. This relay setup expands the beam diameter to 1.95 mm.

Next is the scanning module consisting of two conjugated galvanometer mirrors,

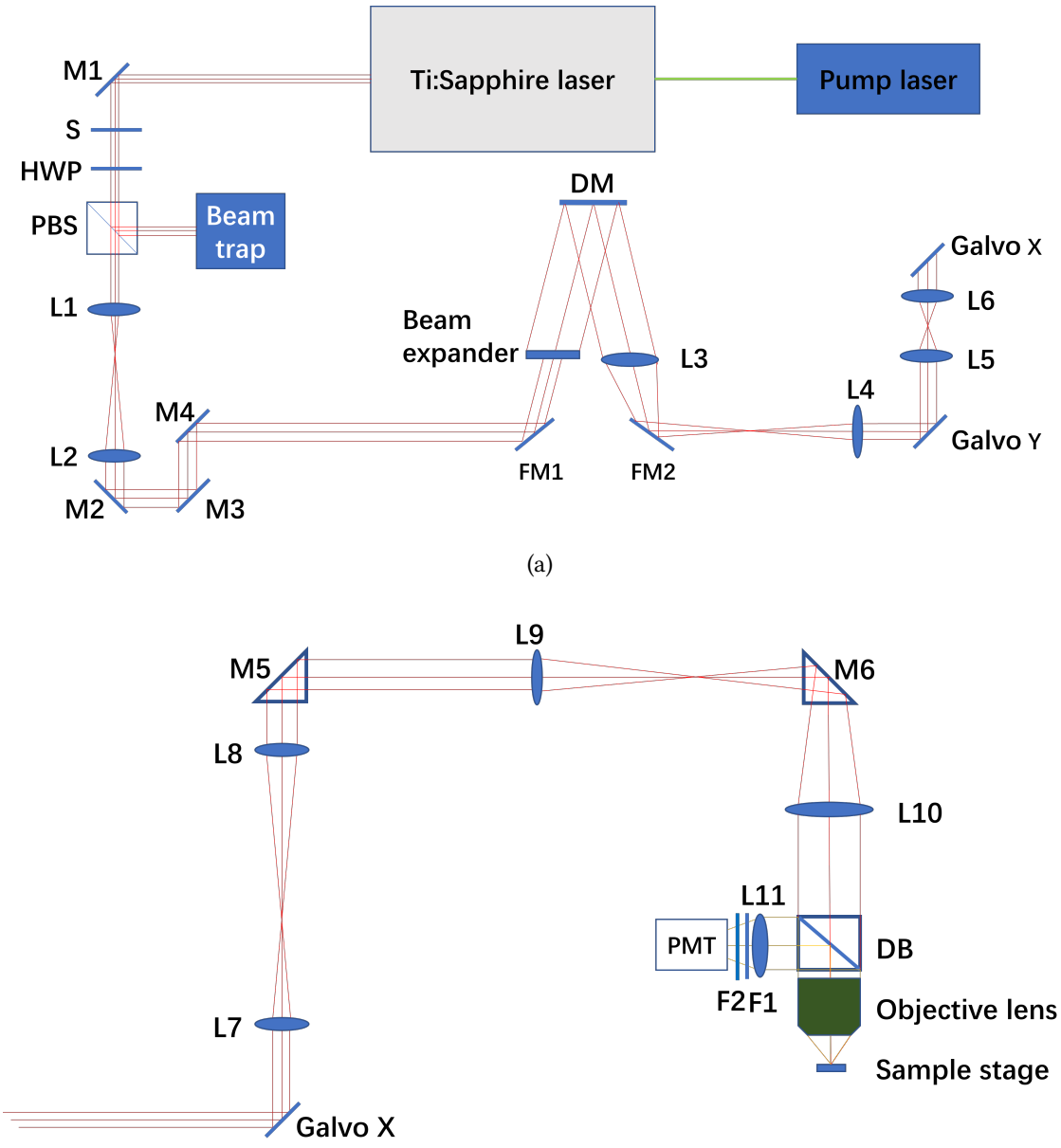


Figure 6.1: System configuration. (a) Two-photon illumination path with the adaptive optics path. (b) Two-photon illumination path and the signal detection path.

Galvo X and Galvo Y (GVS001, Thorlabs), together with a 4f system L5 and L6 (Edmund Optics, 49-357-INK). Driven by control signals from the FPGA, the galvanometer mirrors rotate around their respective axes at defined rates, producing the desired scanning pattern of the focal spot across the sample. In particular, Galvo X redirects the main optical axis, enabling the downstream optical components mounted over the optical air table as shown in Figure 6.1(b).

After the scanning module, the beam passes through two pairs of 4f systems: L7 (49357, Edmund Optics), L8 and L9 (49-361-INK, Edmund Optics), and L10 (49-364-INK, Edmund Optics). These systems expand the beam diameter to match the objective back aperture, enlarging it first to 4 mm and then to 6.4 mm. Two mirrors, M5 and M6 (BBE1-E03, Thorlabs), are used to redirect the optical axis.

The illumination path terminates at the objective lens (LUMPLFLN60XW, Olympus), a $60\times$ water immersion lens with a 6 mm pupil diameter, which is matched by the expanded beam diameter. The collimated laser beam is focused by the objective in the sample, exciting two-photon fluorescence.

6.1.2 Signal Detection Path

This path is indicated in yellow in Figure 6.1(b). The excited fluorescence is first collected by the objective lens and then reflected by a dichroic beamsplitter DB (FF735-Di02-25x36, Semrock) with a transmission band of 750 nm-1600 nm. A single lens L11 is placed before the detector to collect the photons. To further block unwanted wavelengths, two bandpass filters are added: F1 (FF01-527/70-25, Semrock) with a 492 nm-

562 nm transmission band, and F2 (FF01-750/SP-25, Semrock) with a 380 nm-720 nm transmission band. Together with the DB, these filters ensure that only the fluorescence signal reaches the detector. A photomultiplier tube PMT (Thorlabs PMT2101/M) is used for signal detection.

6.1.3 Adaptive Optics Path

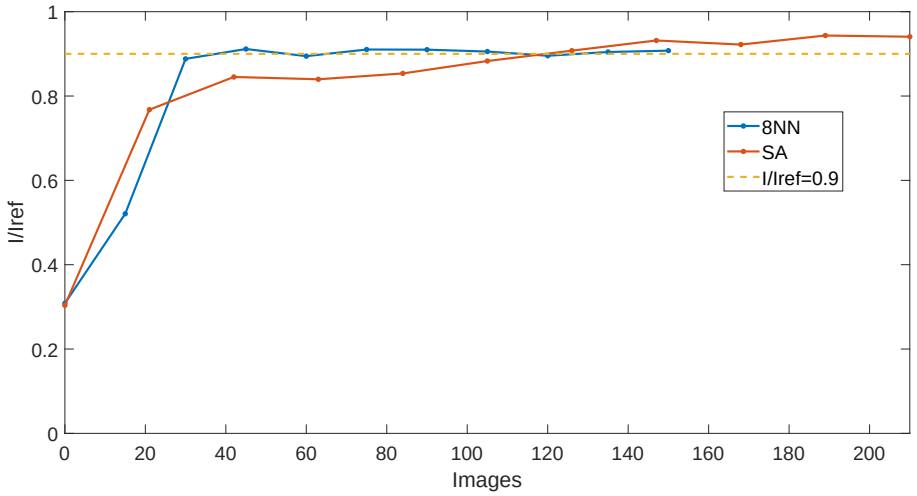
The AO path is an optional path, controlled by three components: flip mirror FM1 and FM2 (BB1-E03, Thorlabs), and a flip lens L4 (49356, Edmund Optics). Collimated light reflected by FM1 enters a $5\times$ beam expander (GBE05-B, Thorlabs), which enlarges the beam diameter from 1.95 mm to 9.75 mm to fill the surface of the deformable mirror DM (Kilo Circular SLM, Boston Micromachines Corporation). The DM has 952 active aluminum-coated actuators with a 10.2 mm clear aperture. The actuator's maximum stroke is 800 nm with piston precision below 1 nm. A 4f system L3 (49366, Edmund Optics) and L4 then conjugates the DM surface to Galvo Y while reducing the beam diameter back to 1.95 mm. Ultimately, the DM surface is conjugated to the objective's back pupil plane by the downstream optical components. The AO path is illustrated in Figure 6.1(a).

6.2 Comparison of 8NN and SA in Experimental Results

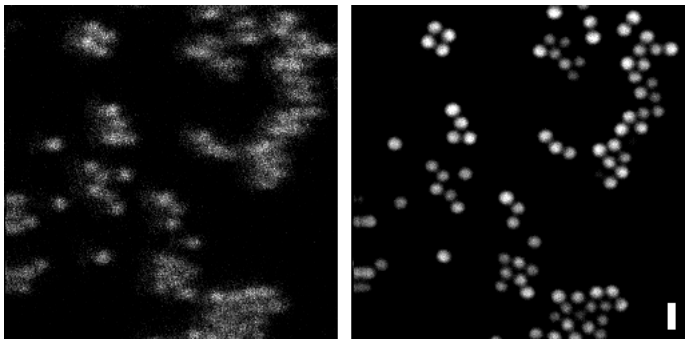
In the following experiments, 8NN and the SA were tested to compare their performance. The specimen was a single layer of 1 µm beads (FluoSpheres Size Kit #2, carboxylate-modified microspheres, Thermo Fisher) with yellow-green fluorescence emission wavelength 515 nm. The laser power at the sample stage was 2.5 mW.

At the beginning of the experiment, the SA was applied to correct the system aberration. This step served two purposes: first, to remove the intrinsic system aberration, enabling statistical analysis of experimental results with known input aberrations manually introduced by the adaptive element; and second, to establish an intensity reference for 8NN, since all prior training and testing were based on Strehl ratios (SRs). Applying SA first, therefore ensured that 8NN operated with a reliable reference for accurate correction.

SA was applied to polar Walsh 1 to 15 (hereafter referred to as $PW1$–15 for clarity; all subsequent polar Walsh mode subsets follow the same notation) for 10 rounds. After this procedure, a system-aberration–corrected pupil phase pattern was obtained, which served as the zero-bias pupil phase pattern for the following experiments, along with an aberration-free intensity as the intensity reference, referred to I_{ref}. Next, $PW1-7$ (excluding the piston mode) were introduced onto the DM as aberrations, with amplitudes randomly selected within the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. Both 8NN and SA were then applied to the same aberrated pupil, and ten rounds of corrections were performed. The results are



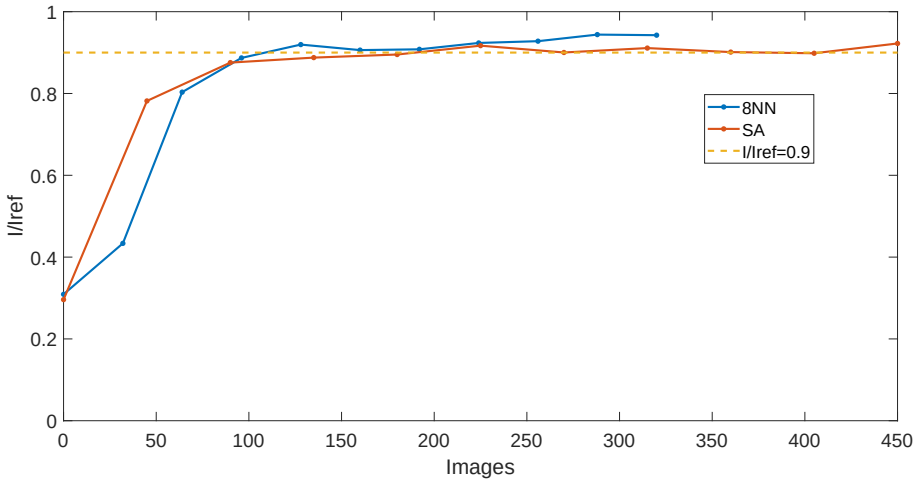
(a)



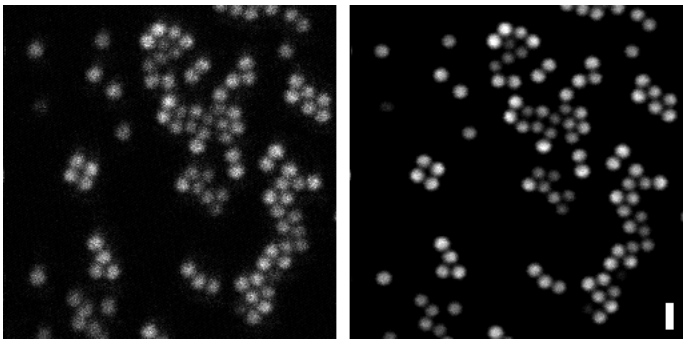
(b)

Figure 6.2: Comparison between 8NN and SA while correcting $PW1-7$. (a) Experimental intensity ratios $\frac{I}{I_{ref}}$ versus collected images, comparing 8NN and SA. (b) One micron bead images before (left) and after (right) 8NN correction. The scale bar: 2 µm.

shown in Figure 6.2. It can be observed that 8NN reached the correction criterion, the intensity ratio $\frac{I}{I_{ref}} = 0.9$, with fewer images, while both 8NN and SA achieved comparable final intensity ratios. Following aberration correction with 8NN, the image intensity and resolution were markedly improved. These results experimentally validate the conclusion drawn in the previous chapter that 8NN can outperform SA.



(a)



(b)

Figure 6.3: Comparison between 8NN and SA while correcting $PW1 - 15$. (a) Experimental intensity ratios $\frac{I}{I_{ref}}$ versus collected images, comparing 8NN and SA. (b) One micron bead images before (left) and after (right) 8NN correction. The scale bar: 2 µm.

In the second experiment, $PW1-15$ were manually applied to the DM surface, with amplitudes randomly selected within the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. The comparison of intensity ratios versus the number of collected images after correction with 8NN and SA is shown in Figure 6.3. The two algorithms showed similar performance, with a similar required number of images to meet the $\frac{I}{I_{ref}} = 0.9$ criterion and similar final $\frac{I}{I_{ref}}$. Following aberration correction with 8NN, both image intensity and resolution were significantly improved as well.

Despite these promising results, 8NN is not practically feasible for experimental applications. It requires a separate aberration correction procedure to generate the aberration-free reference so that the SR can be computed. While this work serves as a promising demonstration and useful exploration of using a NN-based algorithm for Walsh mode correction, to derive a practically feasible method that can be used in realistic imaging scenarios, we further developed an image-based NN method, which will be introduced in the next chapter.

Chapter 7

Machine Learning Assisted Wavefront-Sensorless Adaptive Optics Control Method

In this chapter, a physics-based machine learning assisted wavefront-sensorless AO control method (MLAO) is presented. The approach is designed to fully exploit the information contained in images without relying on a condensed scalar image quality metric. The data pre-processing method (pseudo-PSF generation), MLAO training data generation, network architecture, and training procedure all support this goal. Furthermore, analyzing MLAO parameters provides insights into the information content of different input types and guides performance improvements.

7.1 Introduction to the Pseudo-PSF

In previous chapters, neural networks (NNs) were all trained based on the Strehl ratio (SR). This implies that an aberration-free intensity should always be provided as a reference to compute the SR before using a NN to derive the wavefront correction. In practical applications, such an aberration-free reference is not available before the aberration is fully corrected. In previous experiments, for demonstration purposes, the aberrations were introduced using adaptive elements to analyze the developed NN-based method in a microscope. Such demonstrations, although effective for analyzing the method's performance, have little direct practical application.

In this chapter, we explore a more practical aberration correction strategy that does not rely on the SR as the network input. Instead, we use extended images to generate images of point-like objects, which are equivalent to the microscope's PSF. PSF deformations contain distinct information about aberrations and can be directly captured in microscope images. To extract this information, we implement a pre-processing step: pseudo-PSF computation [146].

An introduction to the PSF is provided in Section 2.2.5. The PSF can be calculated by Fourier transforming the objective pupil function of the fluorescence microscope under the paraxial approximation, and it carries information about aberrations. In a fluorescence microscope, the acquired image is mathematically expressed as the convolution of the PSF with a function representing the fluorophore distribution in the specimen. To better condition the input for NN processing and simplify the MLAO structure, the input (pre-processed microscope images) should ideally contain only aberration infor-

mation, independent of specimen structure. To achieve this, the pseudo-PSF is defined to provide specimen-independent information.

Mathematically, image formation in a 2PLSM is given by:

$$\mathcal{P} = P(\vec{r})e^{j\phi(\vec{r})} \tag{7.1}$$

$$P(\vec{r}) = \begin{cases} 1, & \text{if } |\vec{r}| \leq 1 \\ 0, & \text{if } |\vec{r}| > 1 \end{cases} \tag{7.2}$$

$$Img_0 = |\mathcal{FT}(\mathcal{P})|^4 * Obj \tag{7.3}$$

Equation 7.1 shows the formation of the pupil function. $\vec{r}$ denotes the position vector on the pupil, and $\phi(\vec{r})$ denotes the pupil phase. Equation 7.2 indicates the pupil function amplitude. In Equation 7.3, Img_0 denotes the acquired image, $\mathcal{FT}(\mathcal{P})$ shows the Fourier transform of the pupil function, $*$ represents the convolution operation, and Obj represents the function of the fluorophore distribution in the specimen.

If we use the adaptive element to add a bias mode to the pupil and image the same FOV, the image acquired will be:

$$Img_b = |\mathcal{FT}\{P(\vec{r})e^{j[\phi(\vec{r})+b(\vec{r})]}\}|^4 * Obj \tag{7.4}$$

$b(\vec{r})$ represents the bias mode we add, which should be the aberrating mode we aim to

capture information from in the pupil phase. Now, if we calculate:

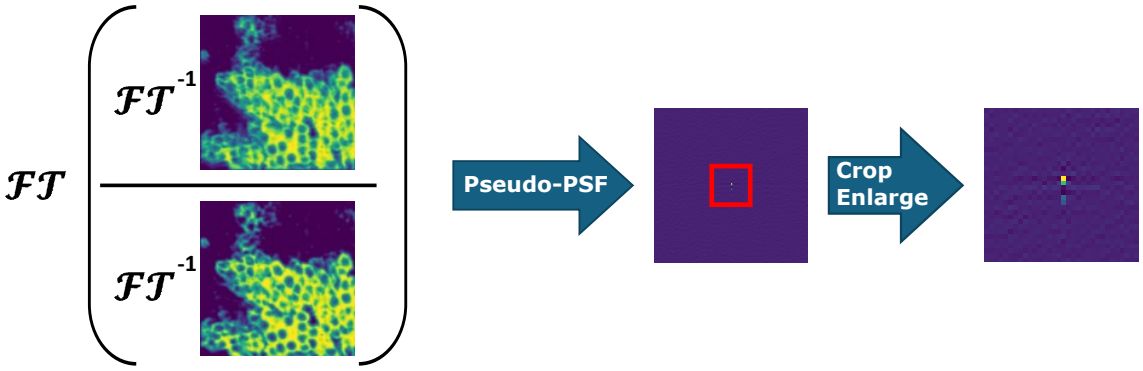
$$\begin{aligned}\frac{\mathcal{FT}^{-1}(Img_b)}{\mathcal{FT}^{-1}(Img_0)} &= \frac{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j[\phi(\vec{r})+b(\vec{r})]}]|^4 * Obj\}}{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j\phi(\vec{r})}]|^4 * Obj\}} \\ &= \frac{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j[\phi(\vec{r})+b(\vec{r})]}]|^4\} \times \mathcal{FT}^{-1}(Obj)}{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j\phi(\vec{r})}]|^4\} \times \mathcal{FT}^{-1}(Obj)} \\ &= \frac{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j[\phi(\vec{r})+b(\vec{r})]}]|^4\}}{\mathcal{FT}^{-1}\{|\mathcal{FT}[P(\vec{r})e^{j\phi(\vec{r})}]|^4\}}\end{aligned} \tag{7.5}$$

$\mathcal{FT}^{-1}$ denotes the inverse Fourier transform. In this formulation, most of the specimen information Obj is removed (see Equation 7.5); the specimen structure is suppressed, though small residual contributions may remain in practice, because all the functions are digitised and pixelated. The aberrated pupil function and the known bias mode are preserved. We then proceed as follows:

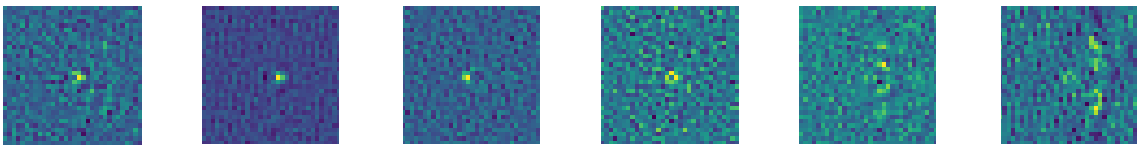
$$\text{PSF}^{\text{pseusdo}} = \mathcal{FT}[\frac{\mathcal{FT}^{-1}(Img_b)}{\mathcal{FT}^{-1}(Img_0)}] \tag{7.6}$$

The resultant pseudo-PSF is concentrated near the center; therefore, we crop the central patch of the 2D function for further processing. This helps simplify the NN architecture. An illustration of the pseudo-PSF generation process is shown in Figure 7.1, including some examples of the pseudo-PSFs.

It is important to note that zeros and noise in the denominator of Equation 7.6 can influence the resulting values. This issue is addressed during the subsequent MLAO training process: zeros are handled as described in Section 7.2.1 **Step 4** 'Pseudo-PSF structure 3D array', and noise is incorporated into the training datasets during data generation.



(a)



(b)

Figure 7.1: (a) The pseudo-PSF generation process. The cropped pseudo-PSF image is enlarged to see the detailed structure. (b) Some examples of the pseudo-PSFs.

7.2 Introduction to MLAO

The MLAO method is a physics-based, machine-learning-assisted, wavefront sensorless AO control approach [146]. It processes multiple types of input: intensity ratios, which provide the NN with information about the Voronoi cell structure; the pseudo-PSF structure 3D array, which reflects the structure of the aberrated pupil; and the signal reading matrix and rough estimates inherited from the 8NN designed in Chapter 4.

For the MLAO application, unbiased and biased images are collected, processed through the image pre-processing module, and used as input to the specially designed MLAO. The estimated aberration correction pupil phase is then applied to the original pupil. This procedure can be repeated iteratively. The overall workflow is illustrated in

Flowchart 7.2.

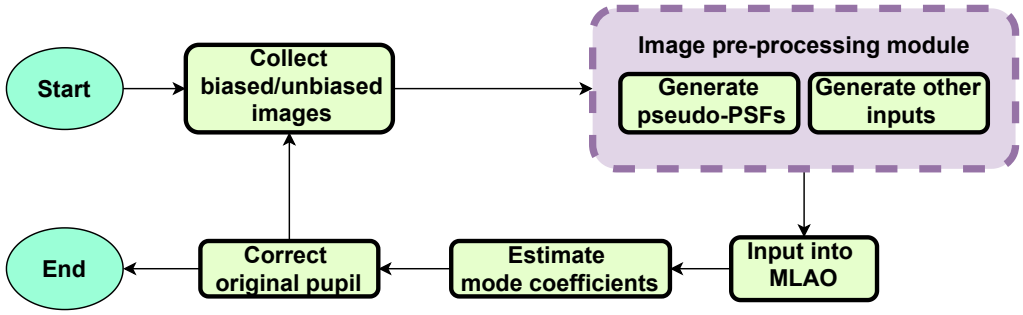


Figure 7.2: MLAO correction procedure.

The MLAO method is effective under a range of working conditions, including varying SNRs and sample compositions, and can substantially improve image quality within a few correction iterations. Furthermore, its architecture is grounded in a physical understanding of the imaging process. This is reflected in the relationship between the aberration information content of each input data category and the variance of the corresponding trained MLAO parameters. These insights offer a deeper understanding of the imaging process and can inform future research.

7.2.1 MLAO Training Data Generation

The complete procedure is illustrated in Flowchart 7.3. In this polar-Walsh-mode-based MLAO, synthesized images are used to generate two divisions of input data: intensity-reading-based and pseudo-PSF-based. These inputs allow the MLAO to capture more information about aberrations, improving its correction efficiency.

As illustrated in Section 5.1, we select a subset dimension of 8 and specifically trained

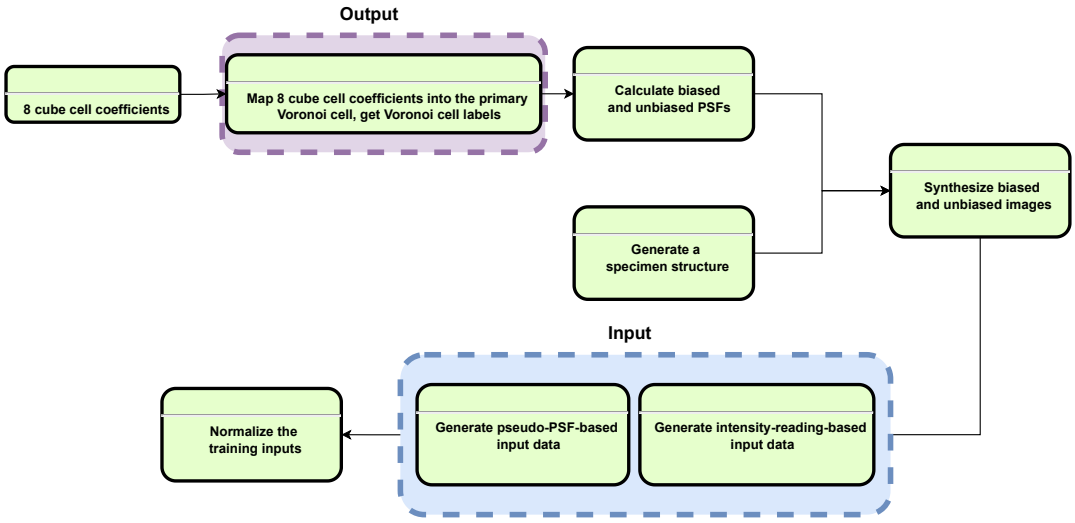


Figure 7.3: The MLAO training dataset generation procedure. The input-output data pair of the training dataset is indicated by two blocks.

on the polar Walsh modes $PW8-15$. All polar Walsh modes are generated based on the layout of the real SLM (X10468-02, Hamamatsu) with 422 pixels in diameter. The high pixel count of the SLM improves the resolution of the Walsh modes and enhances their orthogonality, which largely mitigates mode crosstalk. The 16 polar Walsh modes based on SLM layout are illustrated in Figure 7.4, with their 1D indices labeled.

Step 1.1: generate 8 cube cell coefficients $\beta_k, k = 8, 9, ..., 15$**.** This step corresponds to randomly choosing aberration modes' coefficients in the cube cell (see Figure 3.9 for the cube cell). Each one of them is uniformly chosen in the range $[-\frac{\pi}{2}, \frac{\pi}{2})$.

Step 1.2: map 8 cube cell coefficients into the primary Voronoi cell $\beta'_k, k = 8, 9, ...15$**.** Details are provided in Section 4.1.1 **Step 2**. In this step, we obtain the labels for training.

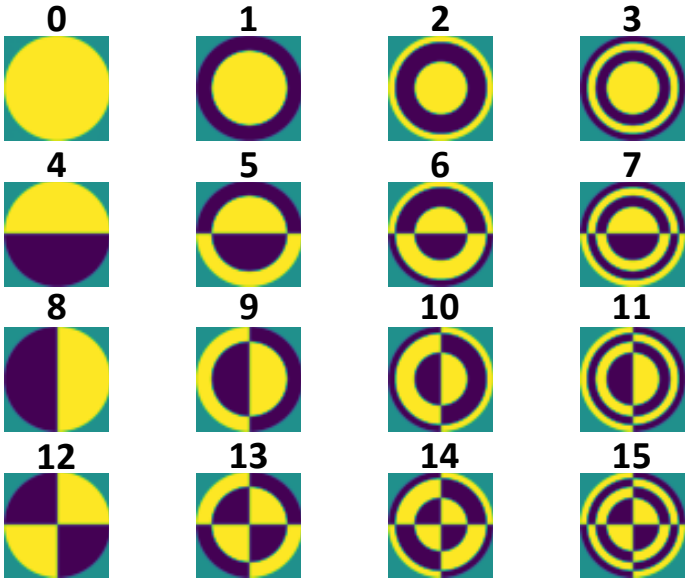


Figure 7.4: The polar Walsh modes based on SLM. Yellow here represents $+1$ and blue here represents -1.

Step 1.3: generate unbiased/biased PSFs. We generate the unbiased/biased pupil phases with Equation 4.5. The mode bias amplitudes are $\pm\frac{\pi}{4}$. Before we conduct the Fast Fourier Transform (FFT) of the resulting pupil, the pupil function array is first padded with zeros (add zeros outside the pupil function) until the sampling rate is $1.2\times$ the base sampling rate. The base sampling rate is defined as: using two pixels to sample the full width half maximum (FWHM) of the PSF of the aberration-free system [146].

Then we calculate:

$$psf(\vec{r}) = \left|\mathcal{FT}[P(\vec{r})e^{j\phi'(\vec{r})}]\right|^2 \tag{7.7}$$

$\phi'(\vec{r})$ denotes the padded phase. The $psf(\vec{r})$ is then normalized to preserve power consistency, followed by squaring to account for the two-photon microscope response:

$$psf'(\vec{r}) = psf(\vec{r}) - \min[psf(\vec{r})] \tag{7.8}$$

$$\mathrm{PSF}(\vec{r}) = \left[\frac{psf'(\vec{r})}{\sum_{\vec{r}} psf'(\vec{r})}\right]^2 \tag{7.9}$$

The vector $\vec{r}$ represents the coordinate of each pixel in $psf'(\vec{r})$. The $psf'(\vec{r})$ is obtained by subtracting the minimum pixel value of $psf(\vec{r})$, which is determined by the $\min$ operation, at each corresponding location. The $\mathrm{PSF}(\vec{r})$ is then obtained and cropped to the central 32×32 pixels, yielding the PSF used for subsequent calculations.

In this step, one unbiased PSF and two biased PSFs for each polar Walsh mode $PW8$–15 are calculated, resulting in 17 PSFs in total.

Step 2: generate a specimen structure. We combine artificial and real microscope images [146]. A pool of realistic sample structures is applied with random shapes such as dots, rings, circular shapes, and curved or straight lines. Rotations of realistic sample structures are also applied. This helps MLAO generalize for application to various specimens. Examples are shown in Figure 7.5.

Figure 7.5: Synthesized specimen structures. Structures at the bottom of the left image are real microscope-collected structures. The other shapes indicated by the red arrows are artificially computed. Small dots in the middle and the right images are also artificially computed.

Step 3: synthesize the unbiased/biased images. The specimen structure computed in **Step 2** is convolved with the PSFs obtained in **Step 1.3**. For each dataset, a

single specimen structure is convolved with both unbiased and biased PSFs to generate the corresponding images. Then the photon number is randomly selected from the range $\mu \in [10^3, 10^5)$ and applied to all 17 images, and Poisson noise is added to simulate real experimental conditions. These processes are inherited from the training data generation **Steps 3** and **Step 4** described in Section 4.1.1. After generating all images, the minimum pixel value across the entire image set is identified and subtracted, following the same procedure used in experiments to remove the background signal.

Step 4: generate each part of the input dataset. A visualization of the input data generation procedure and the arrangement of input data is shown in Figure 7.6.

Signal reading matrix: We sum each of the 17 images:

$$I_{k\pm} = \sum Img_{k\pm} \tag{7.10}$$

$$I_0 = \sum Img_0 \tag{7.11}$$

Here, $Img_{k\pm}$ is a biased image corresponding to polar Walsh mode k, with $\pm$ denoting positive or negative bias, $I_{k\pm}$ represents its image total intensity, while Img_0 and I_0 correspond to the unbiased image and its total intensity.

Subsequently, the calculated total intensities are arranged in the corresponding place following the pattern in Figure 5.7. Finally, we obtain the matrix **I**, which can be seen in Figure 7.6, denoted as 'signal reading matrix'.

Pseudo-PSF structure 3D array: Following Equation 7.6, we get the pseudo-PSF image: $\mathrm{PSF}_{\pm k}^{\mathrm{pseusdo}}$. Each polar Walsh mode k generates two pseudo-PSFs with positive and negative bias amplitudes. Notice that after performing the inverse Fourier transform, we add a tiny epsilon value ($\epsilon \approx 2.22 \times 10^{-16}$) to all pixels to prevent the denominator

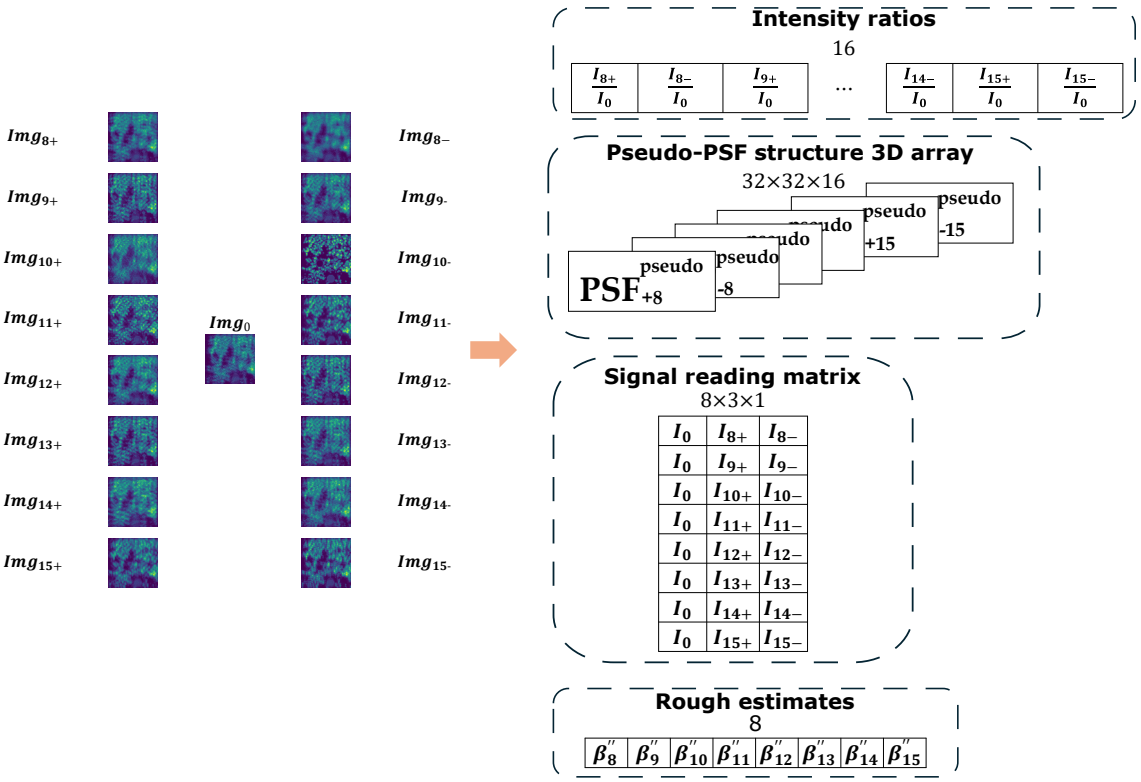


Figure 7.6: Training data generation procedure. Images are arranged with eight positively biased images and eight negatively biased images aligned in two columns, denoted as Img_{k+} and Img_{k-}, where k is the mode index. The unbiased image is Img_0. In total, 17 images are computed for the four types of input data, grouped within dotted blocks. Each block is labeled with the corresponding input data type, and the data size is indicated.

from being zero. Then we crop the pseudo-PSF frame to 32×32 pixels. The pseudo-PSF frame's center pixel value is set to zero because the center part is too bright and can obscure information from dimmer peripheral structures. Finally, the 16 pseudo-PSF frames are stacked together to form the pseudo-PSF structure 3D array: $\mathbf{PSF}^{\text{pseudo}}$. This can also be seen in Figure 7.6.

Intensity ratios: The intensity ratio is calculated as $\frac{I_{k\pm}}{I_0}$. It is the biased image total intensity of mode k with positive/negative bias amplitude over the unbiased image total intensity. Sixteen of these ratios form the intensity ratio vector: $\vec{I}$.

Rough estimates: Following Equation 3.17:

$$\beta_k'' = -0.5 \arctan\left(\frac{-(\sqrt{I_{k+}} - \sqrt{I_{k-}})}{-\left(\sqrt{I_{k+}} + \sqrt{I_{k-}} - 2\sqrt{I_0}\right)}\right) \tag{7.12}$$

The 8 rough estimates are also mapped into the primary Voronoi cell, following **Step 2** from Section 4.1.1.

At this point, a full set of the MLAO training dataset is ready for normalization.

Step 5: input dataset normalization. Normalization of the input training data is beneficial to the training procedure. For the signal reading matrix, we calculate:

$$\mathbf{I}^* = \frac{\mathbf{I} - \min(\mathbf{I})}{\max(\mathbf{I}) - \min(\mathbf{I})} \tag{7.13}$$

The $\min$ function extracts the minimum signal reading in the matrix, and the $\max$ function extracts the maximum. In this way, each $\mathbf{I}$ is normalized to the range $[0, 1]$.

For the pseudo-PSF structure 3D array, which has size $32 \times 32 \times 16$, we also perform:

$$\mathbf{PSF}^{\mathbf{pseudo}} = \frac{\mathbf{PSF}^{\mathbf{pseudo}} - \min(\mathbf{PSF}^{\mathbf{pseudo}})}{\max(\mathbf{PSF}^{\mathbf{pseudo}}) - \min(\mathbf{PSF}^{\mathbf{pseudo}})} \tag{7.14}$$

This entire $32\times32\times16$ array is scaled to the range $[0, 1]$.

Similarly, for the intensity ratios, which form a 1×16 vector, we perform:

$$\vec{I}^* = \frac{\vec{I} - \min(\vec{I})}{\max(\vec{I}) - \min(\vec{I})} \tag{7.15}$$

Eventually, the input dataset is ready for the training process.

7.2.2 MLAO Architecture

The architecture of the MLAO is illustrated in Figure 7.7. It was implemented in TensorFlow Keras. The pseudo-PSF structure 3D array was processed by a convolutional layers (CLs) module to extract information from different spatial scales of the aberration effects on the pseudo-PSFs and hence the imaging process. Another CLs module, identical to that described in Section 4.1.2, extracted features from the signal reading matrix. The outputs of the two CLs were concatenated with the intensity ratios and the rough estimates. The concatenated data was processed by fully connected layers (FCLs), and the output layer generated the final prediction. The detailed MLAO architecture is shown in Appendix A.

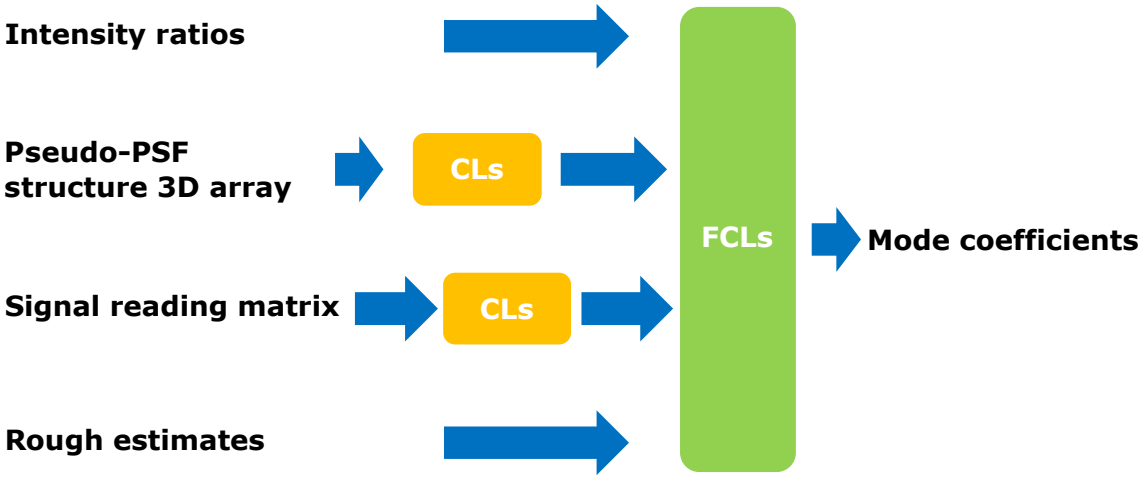


Figure 7.7: Simplified MLAO structure. CLs: convolutional layers; FCLs: fully connected layers.

We adopted this MLAO architecture to process different input types in parallel, which can also provide insight into how information is utilized within the network. In

the FCL after the concatenate layer, the trainable parameters are updated during training. After training, the parameters corresponding to each input type are extracted, and their variances are calculated. A larger variance indicates that the associated parameter group is adjusted more, suggesting that the MLAO absorbs more information from that input type. Such inputs, therefore, provide more information about the aberrations to the MLAO. Further details are provided in Section 7.3.

7.2.3 MLAO Training Process

The MLAO was trained with 8000 batches, each containing 128 datasets, for a total of 16000 epochs. The validation dataset consisted of an additional batch of 128 datasets generated using the same procedure as the training data. During training, the step per epoch was set to 32, meaning the network parameters were updated after every 32 batches. This design increases the frequency of loss updates, facilitating comparison of loss trends across different learning rates and trainable parameters. At the end of each epoch, the batches were shuffled. The initial learning rate was 8.5×10^{-3} and was maintained for 2000 epochs; thereafter, it was reduced by a factor of 0.95 every 200 epochs. Training and validation loss curves were monitored using TensorBoard. The training results are shown in Figure 7.8.

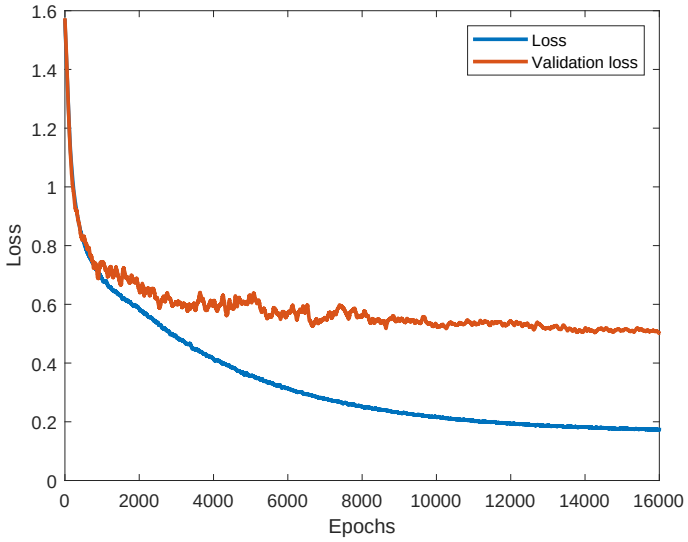


Figure 7.8: The training loss and validation loss of MLAO, both are smoothed with a 50 data span.

7.3 Parameter Variance Analysis for Different Input Types

In our MLAO architecture, extracting the NN parameters from the first FCL after the concatenate layer allows us to assess the relative importance of each input. Larger parameter variances indicate that the MLAO learned more information from the corresponding input type.

In this experiment, we first tested the MLAO trained without noise. The network has 16 parameters for intensity ratios, 9088 for the pseudo-PSF structure 3D array, 512 for the signal reading matrix, and 8 for rough estimates. The corresponding loss curves are shown in Figure 7.9 (light green and pink curves), and the parameter variances are summarized in Table 7.1:

Input	Parameter Variance
Intensity ratios	0.3490
Pseudo-PSF structure 3D array	0.0064
Signal reading matrix	0.0273
Rough estimates	0.0366

Table 7.1: Parameter variances of corresponding inputs.

As shown in Table 7.1, the parameter variance of the pseudo-PSF 3D array was much smaller than that of the other inputs, while MLAO still performed effectively. This observation suggests that this particular input could potentially be removed. Consequently, we removed the pseudo-PSF 3D array input and trained the MLAO with noise, comparing models trained with and without this input. As shown in Figure 7.9, the training and validation loss curves (dark green and brown) were worse without the pseudo-PSF input, indicating that the MLAO performance deteriorated when this input was omitted.

One possible explanation for this behavior is that the number of trainable parameters associated with the pseudo-PSF 3D array was too large, causing some parameters to remain effectively unmodulated and near zero. This masked the true variance of the useful parameters. To address this, we reduced the number of trainable parameters for the pseudo-PSF input from 9088 to 1136 and 284, as shown in Figure 7.9 and Table 7.2. We can observe that the parameter variance of the pseudo-PSF 3D array increased while maintaining similar loss curves, indicating that a larger proportion of parameters were effectively modulated and fewer remained untrained. This reduction can also improve computational efficiency by decreasing the total number of trainable parameters.

The discussion above leads to two main conclusions. First, the parameter variances of inputs are valuable indicators of the information each input provides to the NN. In

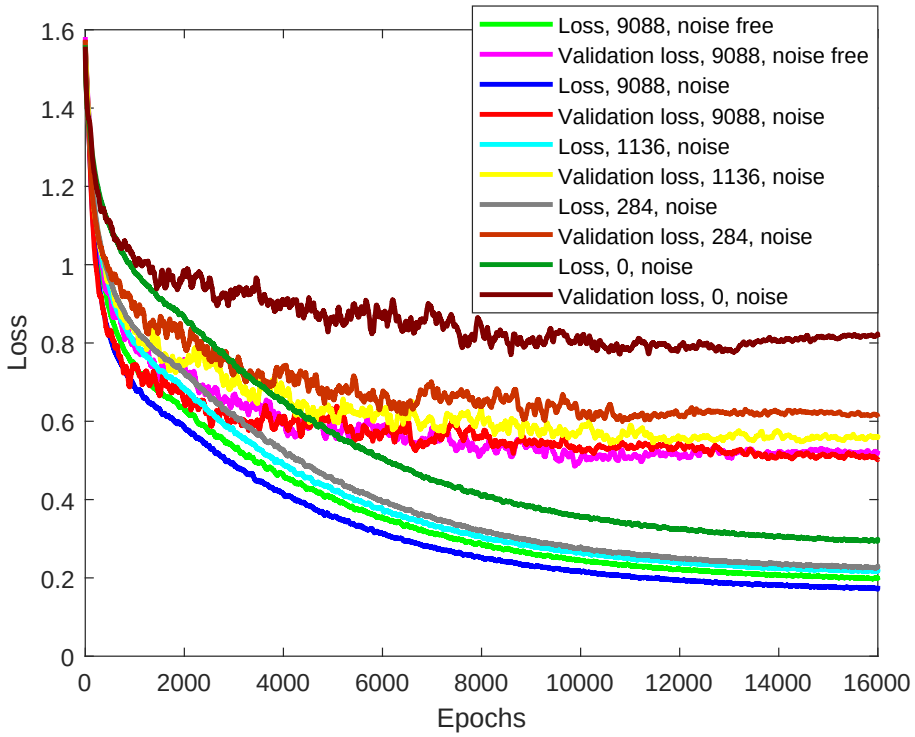


Figure 7.9: Training and validation loss of the MLAO under different pseudo-PSF-related parameter counts and noise conditions, all are smoothed with a 50 data span. The number of parameters and noise conditions are indicated in the top-right legend. The loss curves remain similar as the number of parameters decreases, until the pseudo-PSF input is completely removed, the loss curves become worse.

Input	Variance (9088)	Variance (1136)	Variance (284)
Intensity ratios	0.2599	0.3862	0.3897
Pseudo-PSF structure 3D array	0.0182	0.0552	0.0929
Signal reading matrix	0.0270	0.0410	0.0625
Rough estimates	0.0443	0.0479	0.0608

Table 7.2: Parameter variances of corresponding inputs. Numbers on the titles indicate the number of parameters corresponding to the pseudo-PSF structure 3D array.

the MLAO, this analysis demonstrates that the pseudo-PSF input contains useful aberration information. This solution offers deeper insight into the imaging process, which can guide further research. Second, parameter variances can reveal overfitting: if an important input exhibits a small variance, the number of trainable parameters for that input can be reduced until the loss curves degrade and variance no longer increases. This strategy reduces the total number of trainable parameters, improving the computational efficiency of the MLAO.

7.4 Summary

This chapter presented four key components of the MLAO: data pre-processing (pseudo-PSF computation), training data generation, network architecture, and the training process. These components collectively enhance the MLAO's efficiency and robustness, which will be simulated in subsequent chapters. In addition, the 'relevance' of each input type was evaluated to quantify its information content, offering deeper insight into the imaging process and guiding further research. The MLAO architecture was modified accordingly to improve performance as well.

Chapter 8

Simulations of MLAO-Based Correction of Walsh Modes

8.1 Comparison of MLAO, CA, and SA under Varying Signal-to-Noise Ratios

In the following tests, polar Walsh modes $PW8$–15 were corrected under different SNR conditions to evaluate the robustness and efficiency of MLAO and to compare its performance with SA and CA. Each test used 256 datasets within a specified SNR range. The training SNR range was SNR $\in [10, 100)$. High, low, and extra-low SNR conditions were defined as SNR $\in [30, 100)$, SNR $\in [10, 30)$, and SNR $\in [3, 10)$, respectively.

For each test dataset, a specimen structure was synthesized, and the coefficients of $PW8$–15 were randomly selected within the range $[-\frac{\pi}{2}, \frac{\pi}{2})$. Subsequently, $2N+1$ simulated images (with $N = 8$) were generated using bias mode amplitudes of $\pm\frac{\pi}{4}$, and the

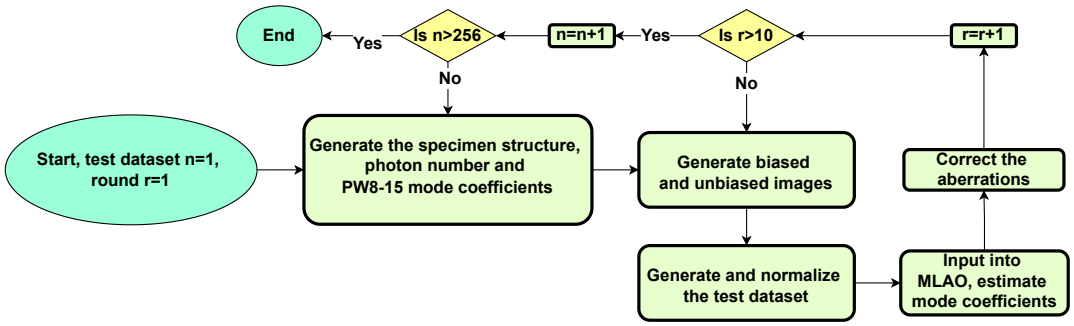


Figure 8.1: Workflow of MLAO correcting $PW8$–15. Each test includes 256 datasets, with 10 correction rounds per dataset.

test dataset was computed and normalized. Further details of the data generation procedure are provided in Section 7.2.1. Using MLAO, the predicted mode coefficients were obtained. The predicted pupil phase was then applied to the original pupil to prepare it for the next correction round. Each test dataset underwent 10 correction rounds. The complete workflow of this procedure is illustrated in Figure 8.1.

Figure 8.2 shows the average corrected SR as a function of the number of images collected, comparing MLAO with CA and SA on the same datasets. Under SNR $\in [10, 100)$, SNR $\in [30, 100)$, and SNR $\in [10, 30)$, MLAO reached both SR $= 0.8$ and SR $= 0.9$ benchmarks with the minimum number of images, while achieving final average SR values close to 1. Even in the SNR $\in [3, 10)$ range, MLAO still reached the benchmarks with fewer images than SA and achieved a final average SR over 0.9. Overall, MLAO demonstrates robustness across SNR conditions, with consistently higher correction efficiency than conventional algorithms.

Moreover, SA was observed to be less robust than MLAO: approximately 5.4% of datasets remained 'stuck' at low SR values (SR < 0.2) and could not be corrected. This

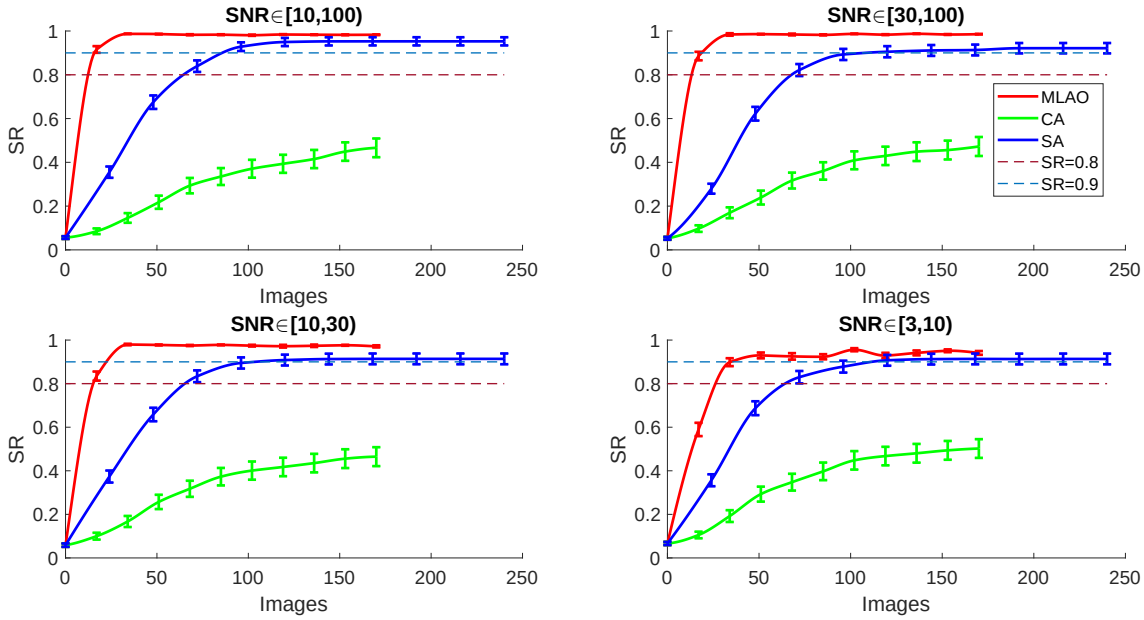


Figure 8.2: Comparison of MLAO, CA, and SA in correcting $PW8$–15 under varying SNR conditions. Corrected average SR versus the number of collected images is shown for different algorithms and SNRs (interpolated curves). Algorithm colors are indicated in the top-right legend, and SNRs are labeled in the panel headings. Error bars indicate the standard error of the SR values. The bump of the MLAO curve when SNR $\in [3, 10)$ is due to the limited size of the testing dataset.

phenomenon can be explained using Figure 3.6. When two correctable polar Walsh modes are present, the metric value can become trapped within the 'zero scaffold' (the dark blue skeleton in the lattice figure). In such cases, a third polar Walsh mode is required to resolve the stagnation. This third mode is obtained through pairwise multiplication of the two modes (see Appendix B, Figure B.1). However, in our experiments with $PW8$–15, the corresponding third modes are $PW1$–7 (see Figure B.2), which were excluded from the mode set. Consequently, some aberrations constructed from a 'semi-complete' polar Walsh mode set (where mode multiplication yields modes outside the

chosen set) can not be corrected with SA, which processes each mode sequentially. In contrast, MLAO directly learns the structure of the high-dimensional metric space, enabling it to largely overcome this limitation. From this perspective, MLAO demonstrates greater robustness than SA.

In the above tests, MLAO proves more efficient than SA and demonstrates robustness to noise. In contrast, SA lacks robustness when confronted with the high-dimensional metric structure, a limitation that MLAO is able to overcome.

8.2 Comparison of MLAO, CA, and SA under Untrained Polar Walsh Mode Subset

From Section 4.3, different Walsh mode subsets are known to generate the same lattice structure. This allows a NN to be trained on one subset and applied iteratively to others, simplifying training. However, this approach forces the network to rely solely on the lattice-like metric space. In contrast, the MLAO is designed to utilize both the metric space and the aberration structure itself. To evaluate this simplification capability, we tested MLAO on an untrained Walsh mode subset.

In the following tests, $PW1$–15 were corrected under different SNRs, with $PW1$–7 serving as the untrained subset, to assess whether MLAO performed equally well on previously unseen modes. Each test was performed with 256 datasets for each SNR range, identical to those used in the previous section. In each dataset, coefficients of $PW1$–15 were randomly selected within the range $[-\frac{\pi}{2}, \frac{\pi}{2})$. For the first part of each correction

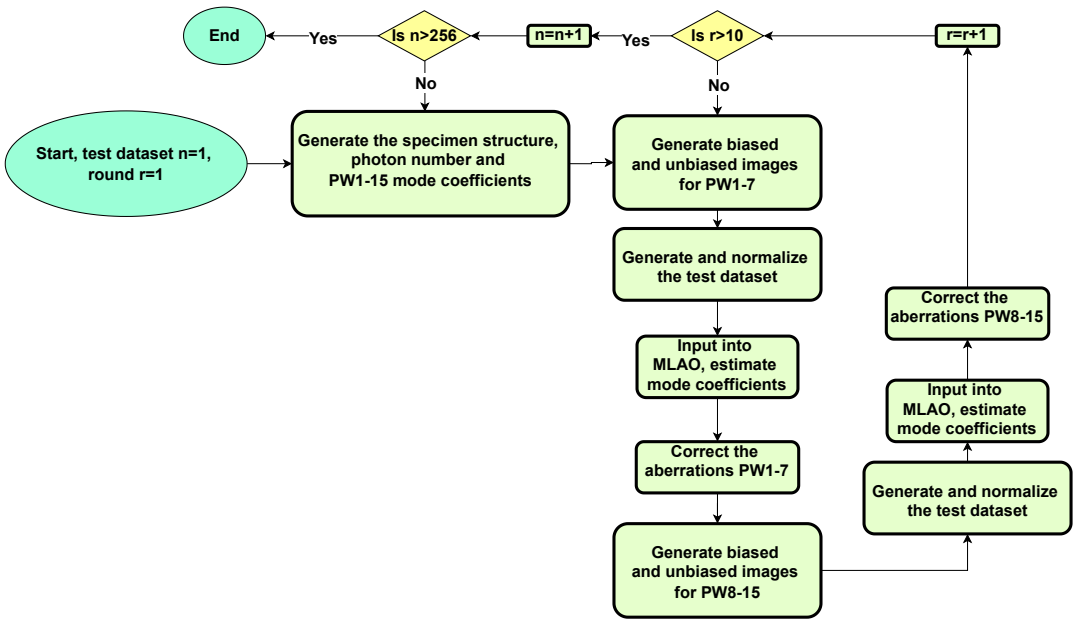


Figure 8.3: Workflow of MLAO correcting $PW1$–15. Each test includes 256 datasets, with 10 correction rounds per dataset.

round, $2N+1$ images were generated for $N = 7$ using bias mode amplitudes of $\pm\frac{\pi}{4}$ with $PW1$–7. Since the piston mode was ignored, the corresponding signal reading matrix inputs for I_{0+} and I_{0-} were set to $0.0625I_0$ (referring to Figure 5.9). The intensity ratios followed the same constraint, and the rough estimate of the piston mode was set to zero (see Section 7.2.1, Figure 7.6 for the input data arrangement). MLAO then predicted the seven mode coefficients, which were applied to update the pupil.

In the second part of the correction round, $2N + 1$ images were generated for $N = 8$ with the same bias amplitudes $\pm\frac{\pi}{4}$ using $PW8$–15. After updating the pupil with these predictions, one full correction round was completed. Each dataset underwent 10 correction rounds. The complete workflow of this testing procedure is illustrated in

Figure 8.3.

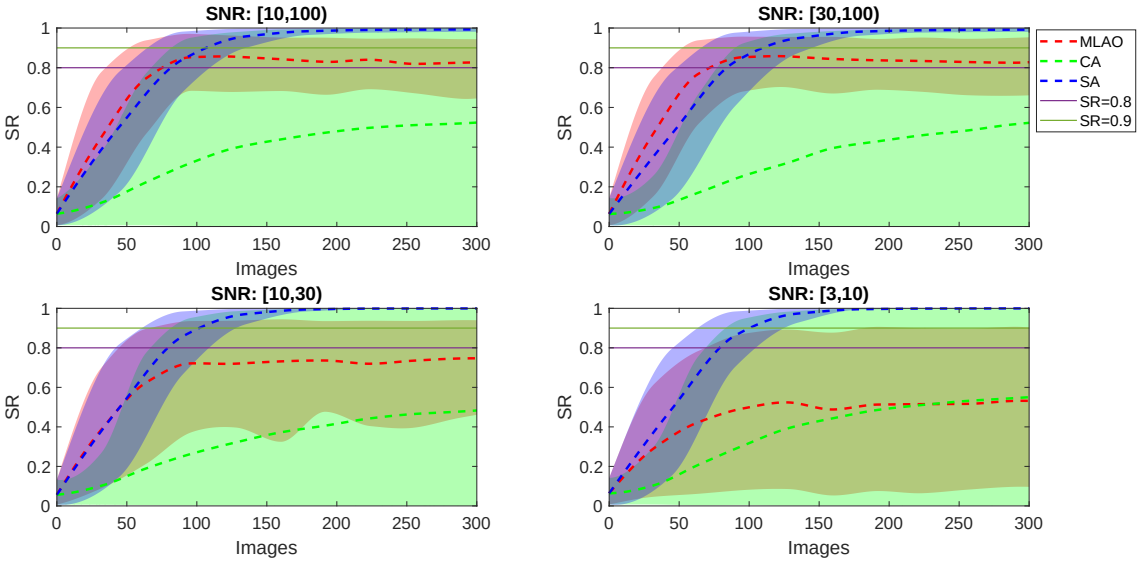


Figure 8.4: Comparison of MLAO (trained on $PW8$–15), CA, and SA in correcting $PW1$–15 under varying SNR conditions. The corrected average SR as a function of the number of collected images is shown in each panel (interpolated curves) for different correction algorithms and SNRs. Algorithm colors are indicated in the top-right legend, and SNRs are shown in the panel headings. Shaded areas represent the 10^{th} to 90^{th} percentiles of all SR samples.

Figure 8.4 shows the average corrected SR as a function of the number of images collected, comparing MLAO with CA and SA on the same datasets. We observed that SA achieved good results because it operated on a complete polar Walsh mode set and therefore did not suffer from the stuck-zero issue (see Section 8.1). For the benchmark SR $= 0.8$, MLAO reached the target under SNR $\in [10, 100)$ and SNR $\in [30, 100)$ using a similar number of images as SA. However, the average SR values only rose to between 0.8 and 0.9, after which the curves declined slightly and exhibited minor oscillations. Under other SNR conditions, MLAO's performance was closer to that of CA, with a broad

spread of SR values at each correction round, and it was unable to reach the higher SR benchmarks. These results indicate that MLAO is not well-suited for untrained polar Walsh mode subsets.

In this case, we trained and tested a new MLAO based on $PW1$–7 (with the piston mode removed). This enables the use of two separate MLAOs: one trained on $PW1$–7 and another trained on $PW8$–15, rather than relying on a single MLAO trained on $PW8$–15 to handle all 15 polar Walsh modes ($PW1$–15). The training procedure was identical to that of the MLAO based on $PW8$–15, except that the number of modes was reduced from eight to seven.

To test its performance, polar Walsh modes $PW1$–7 were corrected under different SNR conditions, and the newly trained MLAO's performance was compared with that of SA and CA. Each test used 256 datasets within a specified SNR range, similar to the previous tests.

Figure 8.5 shows the average corrected SR as a function of the number of images collected, comparing MLAO with CA and SA on the same datasets. In all cases, MLAO met the correction criteria with the minimum number of images and reached final average SR values near 1. Moreover, both the required number of images and the final average SR values remained stable across all SNR conditions. These results indicate that the newly trained MLAO outperforms the other algorithms in efficiency and demonstrates robustness across SNR conditions.

We then applied the two MLAOs—one based on $PW1$–7 and the other on $PW8$–15—as a combined MLAO set to the full Walsh mode set ($PW1$–15) under different SNR conditions and compared its performance with SA and CA. Each test used 256 datasets

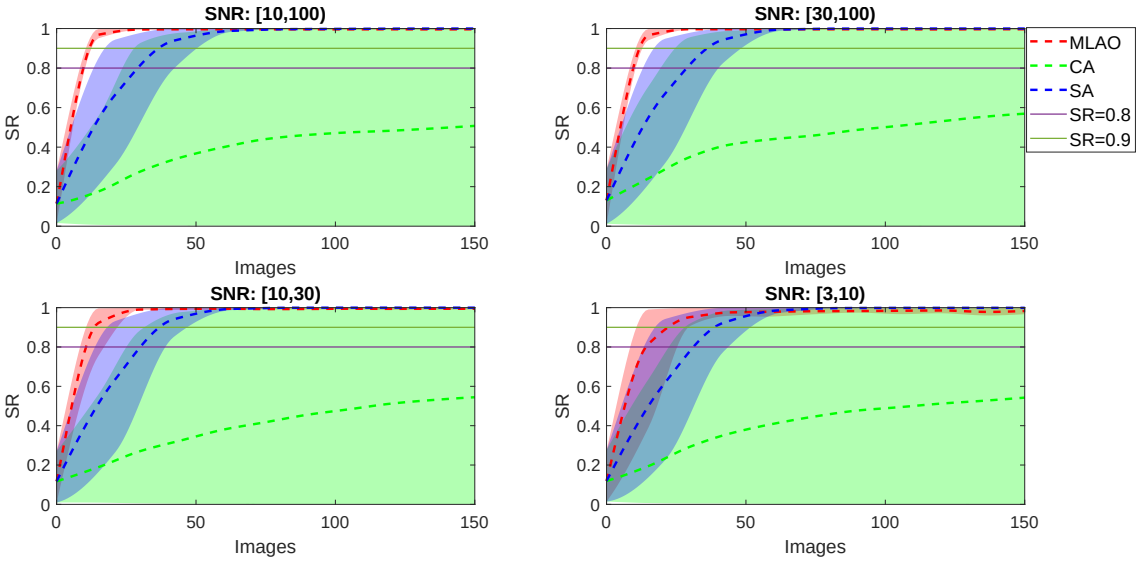


Figure 8.5: Comparison of MLAO (trained on $PW1-7$), CA, and SA in correcting $PW1$–7 under varying SNR conditions. The corrected average SR as a function of the number of collected images is shown in each panel (interpolated curves) for different correction algorithms and SNRs. Algorithm colors are indicated in the top-right legend, and SNRs are shown in the panel headings. Shaded areas represent the 10^{th} to 90^{th} percentiles of all SR samples.

per SNR range, consistent with the previous tests. Figures 8.6 shows that for SNR $\in [10, 100)$, SNR $\in [30, 100)$, and SNR $\in [10, 30)$, MLAO reached the benchmarks with the minimum number of images, achieving final average SR values near 1. Under SNR $\in [3, 10)$, both MLAO and SA achieved the SR $= 0.8$ benchmark with nearly the same number of images. Overall, the combined MLAO set demonstrates robustness across SNR conditions, with consistently higher correction efficiency.

In conclusion, an MLAO trained on a specific Walsh mode subset cannot be directly applied to another subset; however, a new MLAO can be trained for the alternative set. Compared with SA, the combined MLAO set shows robustness and higher efficiency

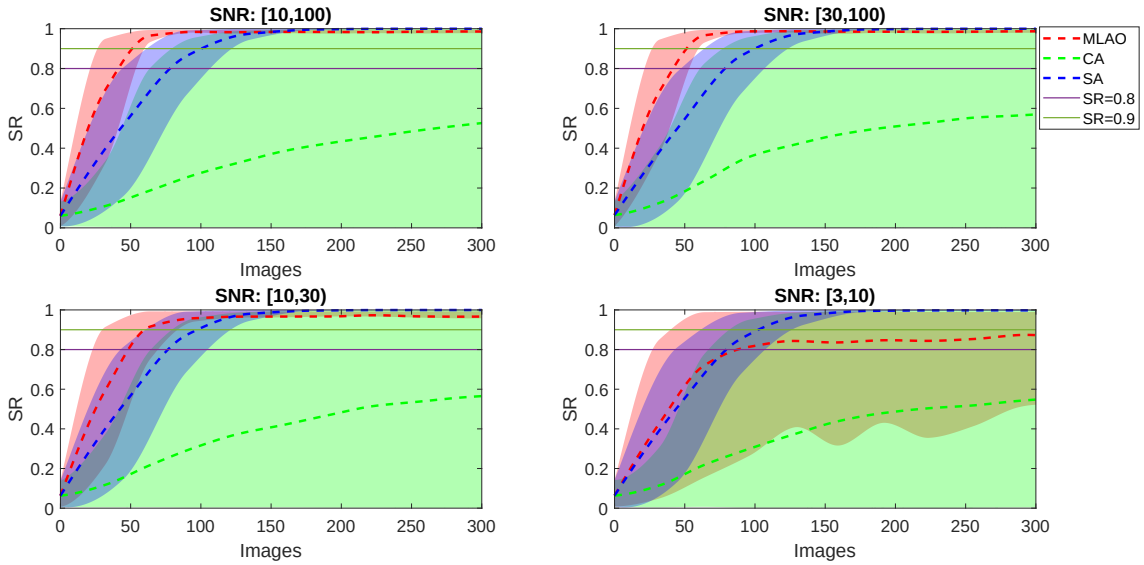


Figure 8.6: Comparison of combined MLAO set, CA, and SA in correcting $PW1$–15 under varying SNR conditions. The corrected average SR as a function of the number of collected images is shown in each panel (interpolated curves) for different correction algorithms and SNRs. Algorithm colors are indicated in the top-right legend, and SNRs are shown in the panel headings. Shaded areas represent the 10^{th} to 90^{th} percentiles of all SR samples.

across most SNR conditions.

8.3 Summary

In this chapter, MLAO was tested under various SNRs and compared with CA and SA. When applied to the trained subset $PW8$–15, MLAO demonstrates higher efficiency than SA and exhibits robustness to noise, whereas SA shows limited robustness when handling a high-dimensional metric structure with a semi-complete polar Walsh mode set. However, MLAO is not well-suited for untrained Walsh mode subsets. To address

this, a new MLAO was trained on $PW1$–7, and the two MLAOs ($PW1$–7 and $PW8$–15) were combined. This combined MLAO demonstrates robustness and superior efficiency across most SNR conditions compared with CA and SA. Further experiments will be conducted to validate these conclusions.

Chapter 9

Two-Photon Microscopy Demonstrations of MLAO-Based Correction of Walsh Modes

In this chapter, the trained MLAO introduced previously was applied to a two-photon microscope under varying working conditions to evaluate its efficiency and robustness, and was compared with CA and SA. Before presenting the experiments, the system configuration is described, followed by the aberration correction procedure and results.

9.1 System Configuration

The two-photon microscope can be divided into three main components: a two-photon illumination path, an adaptive optics path, and a signal detection path. A complete illustration of the system is shown in Figure 9.1, with the setup split into two figures for clarity.

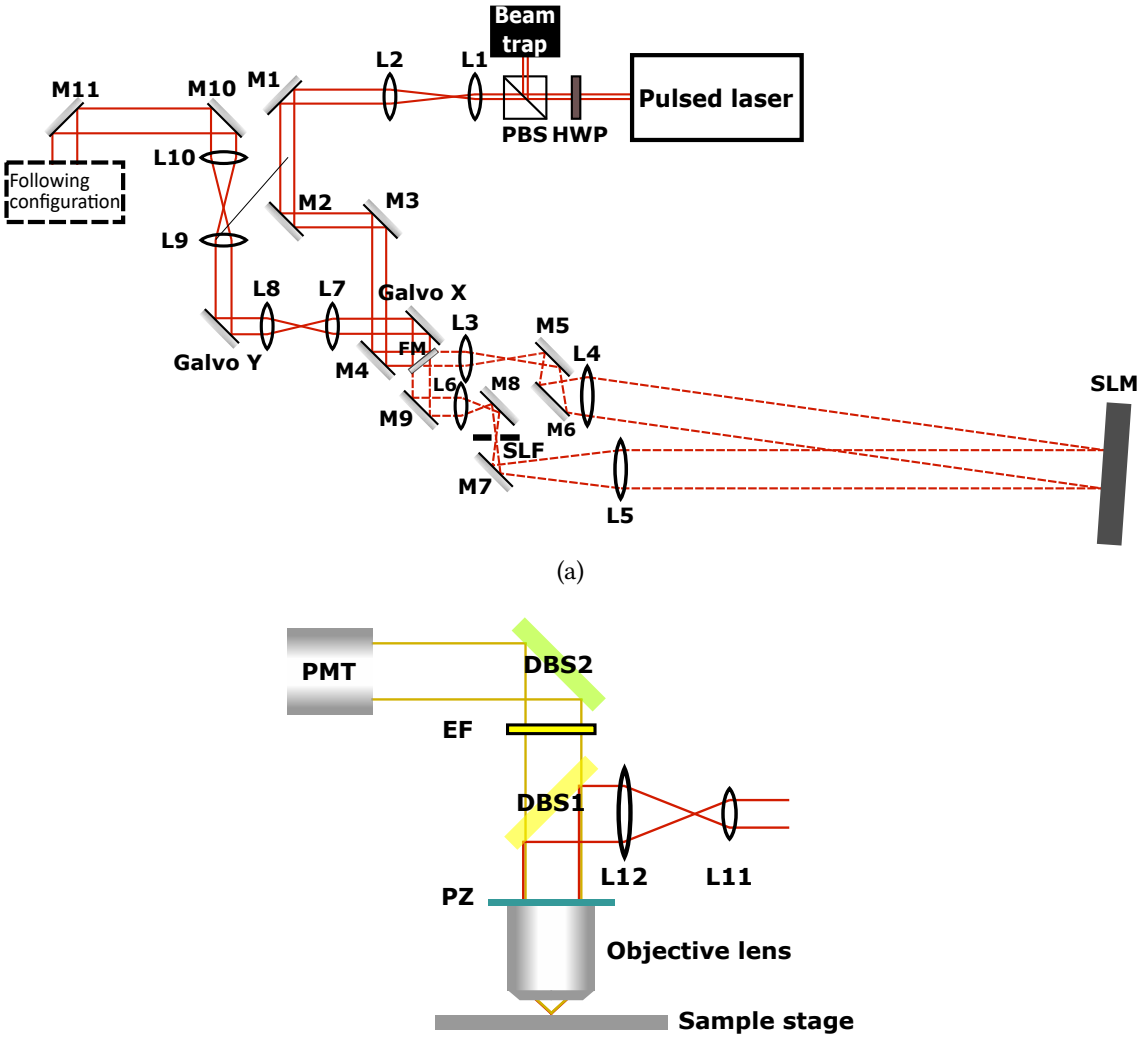


Figure 9.1: Configuration of the two-photon microscope system. (a) HWP: half-wave plate; PBS: polarizing beam splitter; L: lens; M: broadband dielectric mirror; FM: flip mirror; SLM: spatial light modulator; SLF: spatial light filter; Galvo X/Y: galvanometer mirrors. The dashed black box highlights the components detailed in (b). Red pathways indicate the illumination path, while red dashed pathways represent the adaptive optics path. (b) L11 is positioned immediately after M11. DBS1/2: dichroic beamsplitter; PZ: Z-piezo translation stage; EF: emission filter; PMT: photomultiplier tube. Red pathways indicate the illumination path, and yellow pathways denote the signal detection path.

9.1.1 Two-Photon Illumination Path

This path is shown in red in Figure 9.1. The excitation laser in the experimental setup is an InSight DS+ from Spectra-Physics, a femtosecond pulsed laser source. Its tuning range is from 680 nm to 1300 nm. In the experiments, we selected 820 nm as the excitation wavelength. The laser is linearly polarized, with a beam diameter of 1.1 mm and a pulse width of less than 120 fs. After leaving the laser source, the pulsed beam passes through a half-wave plate (HWP) mounted on a motorized cage rotator controlled by computer software. Rotation of the HWP adjusts the polarization direction of the incident beam. In combination with a polarizing beam splitter (PBS), the HWP regulates the intensity transmitted through the PBS, while the reflected portion is absorbed by a beam trap.

After the intensity control module, the beam enters the first 4f system, consisting of lenses L1 and L2 with focal lengths (denoted as f) of 50 mm and 150 mm, respectively, with coincident focal points. The beam diameter is expanded to 3.3 mm. It is then reflected by four mirrors (M1–M4), each redirecting the optical axis. After M4, a flip mirror (FM) can be inserted to direct the beam into the scanning module, which includes two conjugated galvanometer mirrors (Galvo X/Y) and a 4f relay system (L7 and L8) to control scanning. The galvanometers, driven by FPGA signals, rotate at designated rates to scan the focal point over the sample along a defined pattern. Galvo Y redirects the optical axis, fixing downstream optical elements over the air table. The subsequent path coincides with the adaptive optics path and is described in the next section.

9.1.2 Adaptive Optics Path

The adaptive optics path is indicated by the red dotted lines in Figure 9.1. When the FM is not inserted, the collimated beam reflected by M4 enters another 4f system composed of L3 ($f = 100\,\text{mm}$), M5, M6, and L4 ($f = 300\,\text{mm}$). The beam diameter is expanded from 3.3 mm to 9.9 mm to fully illuminate the SLM (X10468-02, Hamamatsu), a reflective type, pure-phase spatial light modulator with an effective area of 15.8 mm $\times$ 12 mm, 792×600 pixels, and 20 µm pixel pitch.

After modulation by the SLM, the beam passes through a 4f system consisting of L5 and L6 ($f = 400\,\text{mm}/150\,\text{mm}$), M7 and M8 (redirecting the optical axis), and a spatial light filter (SLF). The SLF, an adjustable aperture at the focal point of L5/L6, filters out all but the first-order (positive) beam (see Section 2.3.2). This 4f system reduces the beam diameter to approximately 3.7 mm. After reflection by M9, the beam enters the scanning module.

After the scanning module, the beam enters another 4f system composed of L9 ($f = 75\,\text{mm}$) and L10 ($f = 120\,\text{mm}$), expanding the beam diameter to 5.94 mm. Two mirrors, M10 and M11, redirect the beam into a subsequent 4f system L11 and L12 ($f = 75\,\text{mm}/150\,\text{mm}$), further expanding the diameter to 11.88 mm. The beam is then reflected by a short-pass dichroic beamsplitter DBS1 with an edge wavelength of 670 nm (FF670-SDi01-25×36, Semrock) onto the back aperture of the objective lens (UAPON40XW340, Olympus), a $40\times$ water-immersion lens with NA $= 1.15$. The objective converges the collimated laser beam into the specimen, exciting two-photon fluorescence. Vertical positioning of the objective is controlled by a Z-piezo translation stage (PZ), enabling 3D

imaging in combination with the galvanometers.

9.1.3 Signal Detection Path

Fluorescent light emitted from the specimen (yellow pathways in Figure 9.1(b)) passes through the objective and DBS1, is filtered by a short-pass emission filter EF (FESH0700, Thorlabs) with a cut-off wavelength of 700 nm, and is then reflected by a long-pass dichroic beamsplitter DBS2 with an edge wavelength of 568 nm (FF568-Di01-25×36, Semrock) before reaching the photomultiplier tube (PMT).

9.2 System Aberration Correction

Before conducting experiments, system aberrations should first be corrected to ensure consistent and quantifiable results. This was achieved using a single-layer bead sample and applying Zernike mode correction with a parabolic fitting algorithm.

9.2.1 Single Layer Bead Sample

The experimental sample consisted of a single layer of beads of different sizes: 0.5 µm, 1 µm, and 2 µm (FluoSpheres Size Kit #2, Carboxylate-modified Microspheres, Thermo Fisher). The beads and water were sealed between a microscope slide and a cover slip, with a drop of water on the cover slip for objective immersion. The cover slip thickness was 0.17 mm, introducing spherical aberration that was compensated using the objective's correction collar. Beads were excited with an 820 nm laser and emitted fluo-

rescence at 515 nm. An illustration of the sample is shown in Figure 9.2. Notice that all the aberrated wavefronts were manually introduced by the SLM.

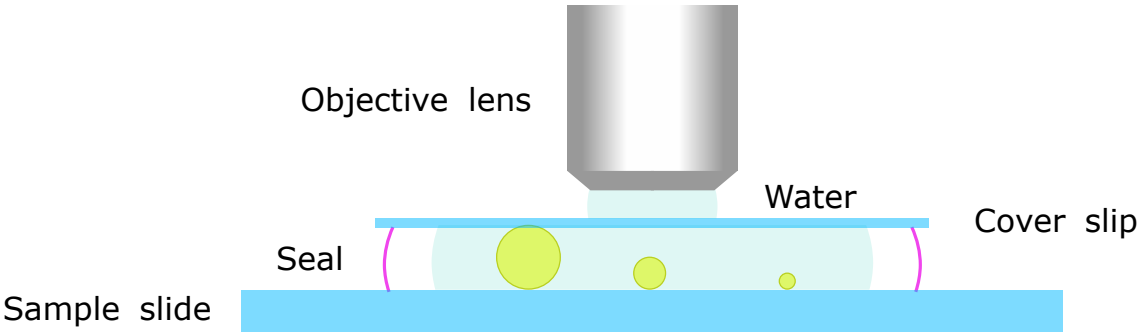


Figure 9.2: The bead sample with the objective lens. Three sizes of beads and water are sealed with transparent nail polish (indicated by purple curves) between the cover slip and the sample slide. The objective lens is immersed in a drop of water.

9.2.2 Implementation of System Aberration Correction

First, it was important to determine the optimal FOV with a dense specimen structure to achieve sufficient SNR, and to set the piezo position so that the beam was focused within the bead layer for the clearest visualization of bead structures. The program then ran 5 iterations, each estimating all Zernike modes from index 4 to 10 (see Figure 9.3). For each mode, 5 biased images were collected with bias amplitudes evenly spaced from -1 to 1 radian (−1, −0.5, 0, 0.5, 1 rad). The images were used to compute intensity metrics, which were fitted to determine the mode coefficient corresponding to the maximum intensity. The optimized coefficient was then applied to the SLM before proceeding to the next mode. The full workflow is illustrated in Figure 9.4.

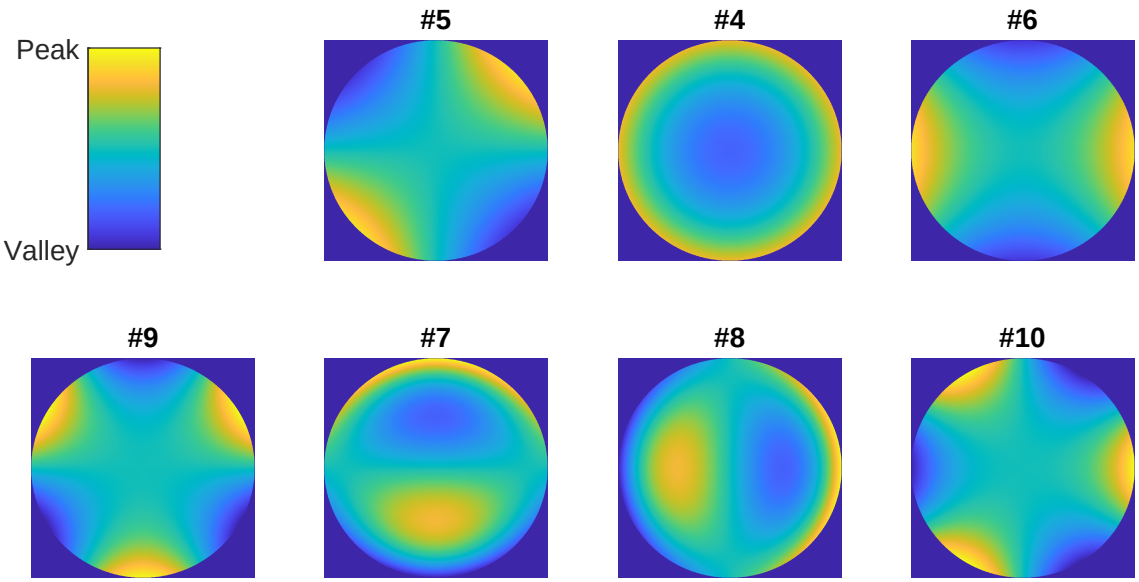


Figure 9.3: Zernike modes from index 4 to 10. Mode indices are labeled in the titles, and structures are displayed in different colors as indicated by the color bar.

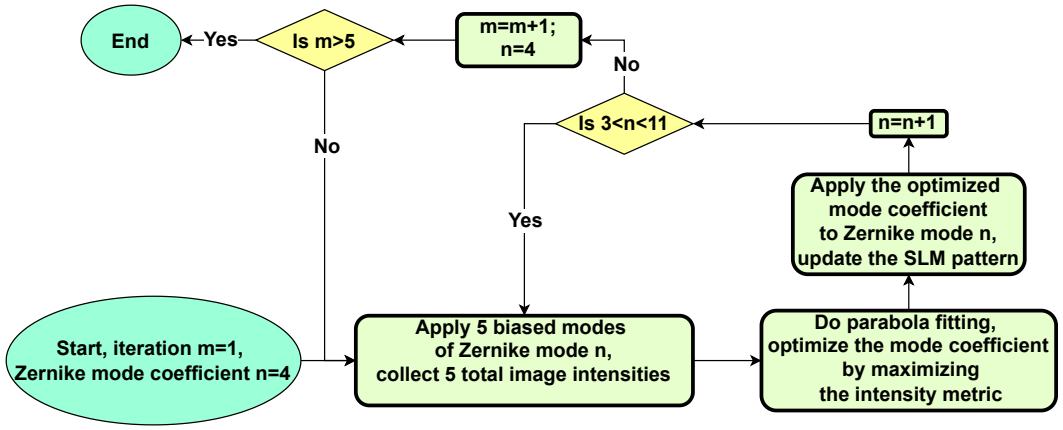
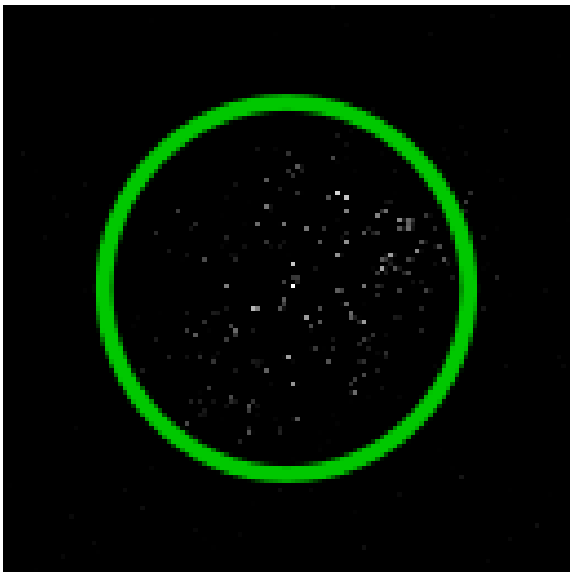


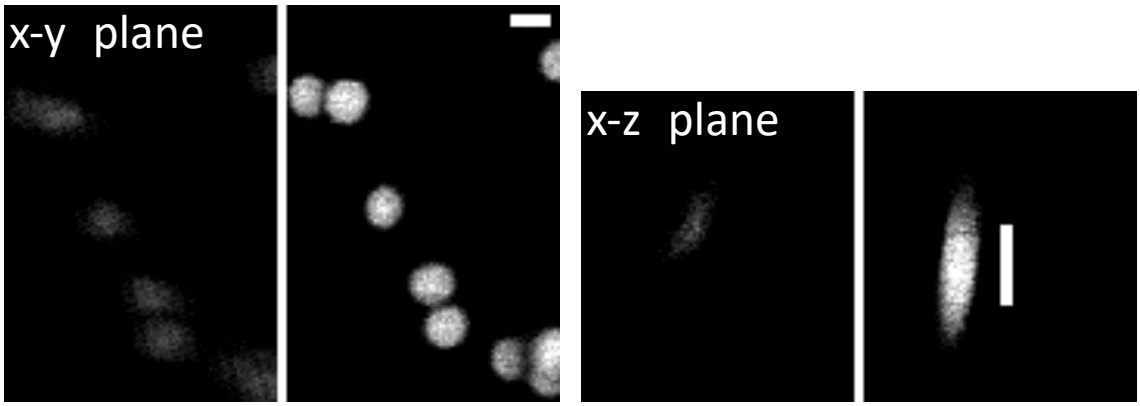
Figure 9.4: Working process of system aberration correction using parabolic fitting. Zernike modes 4–10 are corrected over five iterations.

9.2.3 Results of System Aberration Correction

After system aberration correction, the image intensity and resolution were largely improved. See Figure 9.5(b) and (c), both the x-y plane and x-z plane's image qualities



(a)



(b) (c)

Figure 9.5: (a) Largest FOV image after system aberration correction. The green circle marks the uniform, round illuminated region. Image size: 250 µm × 250 µm. (b) 2 µm bead x–y plane images before (left) and after (right) system aberration correction within the same FOV. Scale bar: 2 µm. Image size: 13.5 µm × 20 µm. (c) x–z plane image of the same 2 µm bead before (left) and after (right) system aberration correction. Scale bar: 4 µm. (The bead appears slightly tilted due to small residual system aberrations.)

were improved.

We imaged the largest FOV allowed by the galvanometer mirrors (250 µm × 250 µm) to verify that the illuminated area was uniform, see Figure 9.5(a) the green circled area.

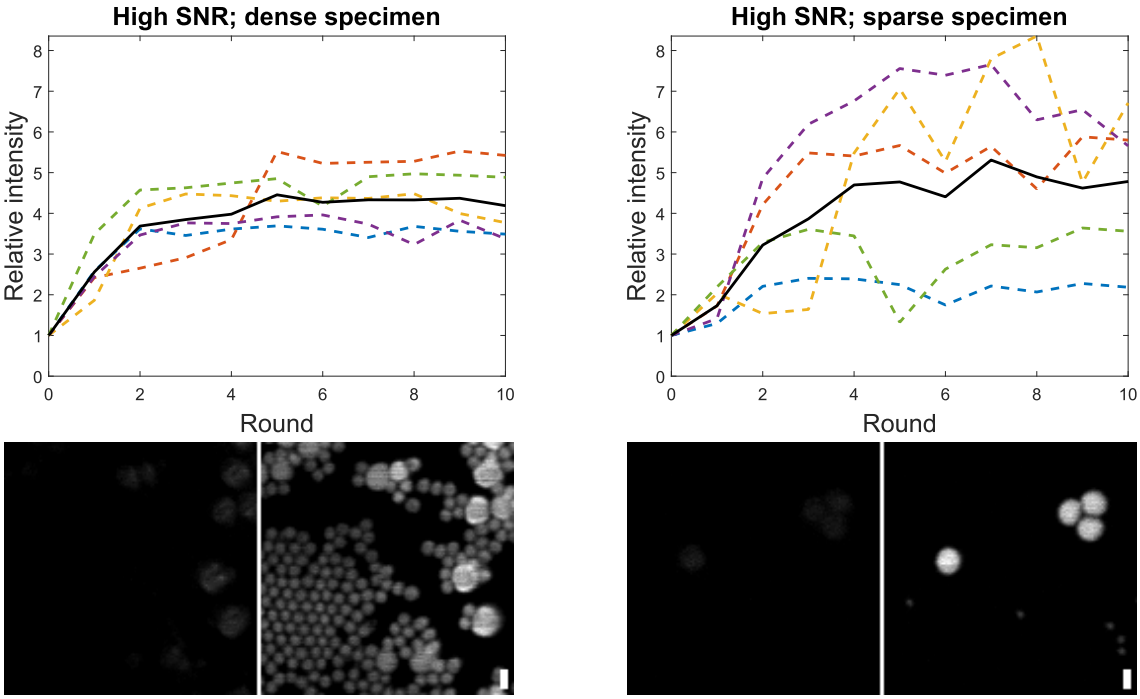
Axial bead images were then collected to verify that the microscope's PSF was near diffraction-limited. The axial full width at half maximum (FWHM) of the diffraction-limited PSF for this system is [195]:

$$\Delta z_{\text{FWHM}} \approx \frac{0.9\lambda}{n - \sqrt{n^2 - NA^2}} = \frac{0.9 \times 515 \text{ nm}}{1.33 - \sqrt{1.33^2 - 1.15^2}} \approx 700 \text{ nm} \tag{9.1}$$

In this case, the ideal axial length of the diffraction-limited pattern for a 2 µm bead was calculated by convolving the bead with Δz_{FWHM}, yielding a maximum of 4 µm. The measured axial image of the 2 µm bead was approximately 4 µm, indicating that system aberrations were largely corrected.

9.3 MLAO Aberration Correction under Varying Signal-to-Noise Ratios and Specimen Structures

After system aberration correction, six sets of experiments were performed under different SNR conditions. Low SNR corresponded to low laser power (∼ 4.7 mW), medium SNR to medium power (∼ 6.5 mW), and high SNR to high power (∼ 8.0 mW). All laser powers were measured at the sample stage. In addition, two specimen structures were tested: densely populated beads (multiple beads within the FOV) and sparsely populated beads (few beads within the FOV). These carefully controlled conditions al-

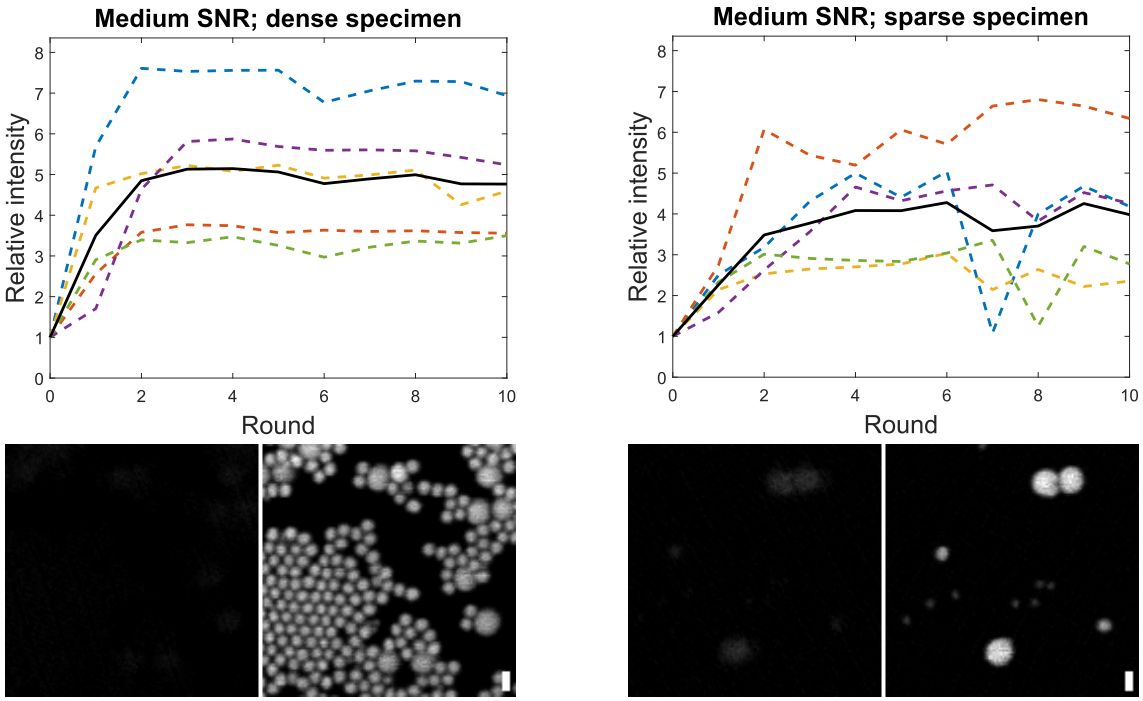


(a) (b)

Figure 9.6: Correction results under high-SNR conditions. Top: relative intensity as a function of correction rounds. Each panel corresponds to one experimental condition; five independent experiments are shown as dotted curves, with the average relative intensity shown in black. Bottom: example bead images before (left) and after (right) 10 rounds of aberration correction, each image pair corresponding to one of the experiments above. The scale bar: 2 μm.

lowed for a systematic evaluation of MLAO efficiency and robustness across different imaging scenarios.

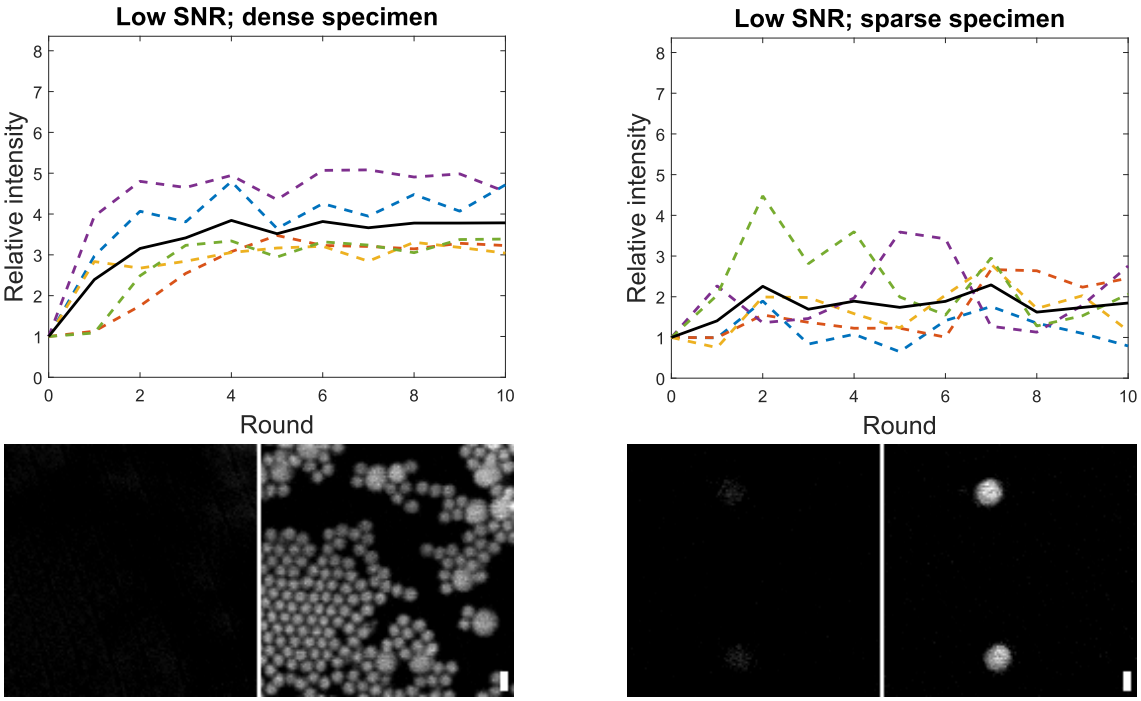
In each imaging scenario, five experiments were performed. For each experiment, aberrations were introduced by the SLM using randomly generated coefficients of PW8-15 within the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. This introduced aberration varied among all five experiments, and this pattern was maintained in the subsequent experiments in this chapter.



(a) (b)

Figure 9.7: Correction results under medium-SNR conditions. Top: relative intensity as a function of correction rounds. Each panel corresponds to one experimental condition; five independent experiments are shown as dotted curves, with the average relative intensity shown in black. Bottom: example bead images before (left) and after (right) 10 rounds of aberration correction, each image pair corresponding to one of the experiments above. The scale bar: 2 µm.

The two-photon microscope then collected both biased and unbiased images, where the bias amplitudes were set to $\pm\frac{\pi}{4}$. Note that the image FOVs were the same for all five experiments in each imaging scenario, and this was maintained in the subsequent experiments in this chapter, except where otherwise specified. These images were analyzed by the proposed MLAO method to estimate the aberration correction. The MLAO-derived correction was subsequently added to the original pattern on the SLM. Correction rounds



(a) (b)

Figure 9.8: Correction results under low-SNR conditions. Top: relative intensity as a function of correction rounds. Each panel corresponds to one experimental condition; five independent experiments are shown as dotted curves, with the average relative intensity shown in black. Bottom: example bead images before (left) and after (right) 10 rounds of aberration correction, each image pair corresponding to one of the experiments above. The scale bar: 2 µm.

were performed sequentially, with each round building upon the previous one. A total of ten correction rounds were conducted per experiment. The complete workflow is illustrated in Flowchart 9.9.

The results are shown in Figure 9.6, Figure 9.7, and Figure 9.8. The intensity metric was defined as the sum of the 30 brightest pixels in each collected image. For consistency and comparison across experiments, the intensity value at each correction round

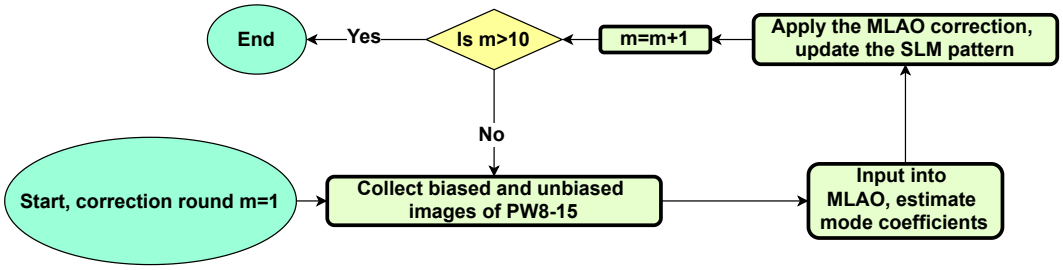


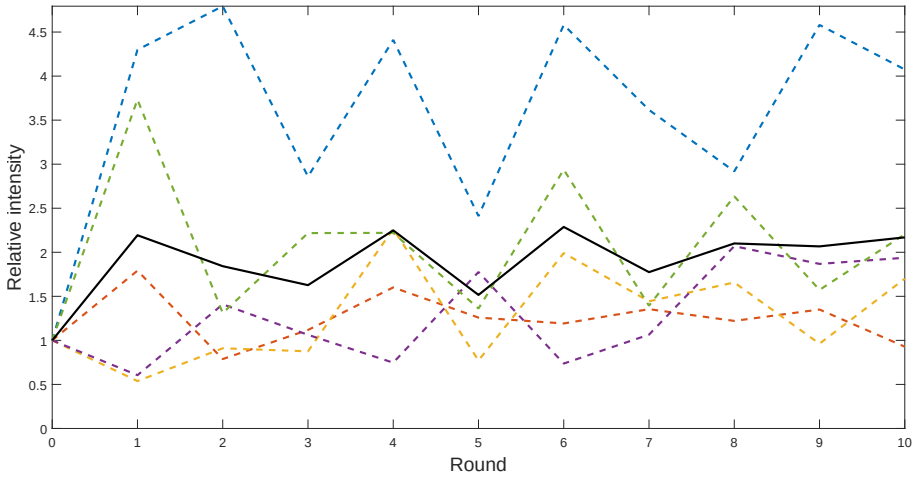
Figure 9.9: Working process of MLAO aberration correction. This process corrects $PW8-15$ in 10 rounds.

was normalized by dividing it by the initial intensity of that experiment. This normalized value is referred to as the relative intensity. We can see that for dense specimens, the average relative intensity reached its maximum within five correction rounds, the convergence curves were smooth, and the image qualities were largely improved. For sparse specimens, the final average relative intensities were not affected under high- and medium-SNR conditions, also reaching near-maximum relative intensities within the first five rounds with significant image quality improvements, albeit with slight fluctuations. However, under low-SNR conditions with sparse specimens, the final average relative intensity was limited, although the image quality can be improved. Overall, these results demonstrate that MLAO can efficiently and robustly improve imaging quality across most SNR levels and specimen densities.

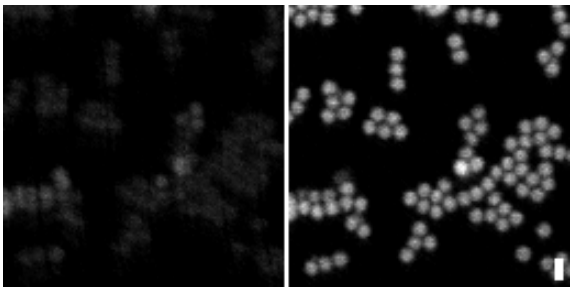
9.4 MLAO Correction of the Untrained Polar Walsh Mode Subset

This experiment relates to the Section 8.2. PW1-7 were corrected by MLAO under high-SNR and dense specimen conditions to evaluate MLAO's performance. Five experiments were conducted, each using an initial aberration applied to the SLM composed of $PW1-7$, with each coefficient randomly selected from the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. Biased and unbiased images were collected by applying $PW1-7$ with bias amplitudes of $\pm\frac{\pi}{4}$. During input dataset generation, 'phantom values' were assigned to the piston mode (see Section 8.2). After the MLAO prediction, the updated pattern displayed on the SLM for the next correction round was obtained by adding the predicted $PW1-7$ pattern to the current pattern. Ten correction rounds in total were conducted for each experiment. The overall workflow was similar to that shown in Figure 9.9, except changing the $PW8-15$ to $PW1-7$.

The results are shown in Figure 9.10. The final average relative intensity was approximately twofold, and the image quality can be improved, but the curve exhibited noticeable fluctuations. This outcome confirms that the MLAO trained on $PW8-15$ is not well-suited for correcting $PW1-7$.



(a)



(b)

Figure 9.10: Results of MLAO correcting untrained polar Walsh mode set under high-SNR and dense sample structure condition. (a) Relative intensity as a function of correction rounds. Five independent experiments are shown as dotted curves, with the average relative intensity shown in black. (b) Example bead images before (left) and after (right) 10 rounds of aberration correction, the image pair corresponding to one of the experiments above. The scale bar: 2 μm.

9.5 MLAO Correction of Trained Polar Walsh Modes with Untrained Modes Present

In practice, aberrations often contain modes beyond the chosen correction set. To evaluate the MLAO's robustness under such conditions, we tested its performance in

the presence of an untrained mode subset. The MLAO was trained on $PW8-15$, but in the experiments, aberrations included both the trained modes ($PW8-15$) and the untrained subset ($PW1-7$).

Five experiments were conducted under high SNR conditions with a dense specimen structure. In each experiment, the initial aberration included $PW1-15$, with each coefficient randomly chosen within the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. During each correction round, bias aberrations were applied only to modes $PW8-15$ with amplitudes of $\pm\frac{\pi}{4}$. After the MLAO predictions, only the coefficients of $PW8-15$ were updated on the SLM, and the subsequent correction round used this updated SLM pattern. A total of ten correction rounds were performed. The full experimental workflow is illustrated in Figure 9.11.

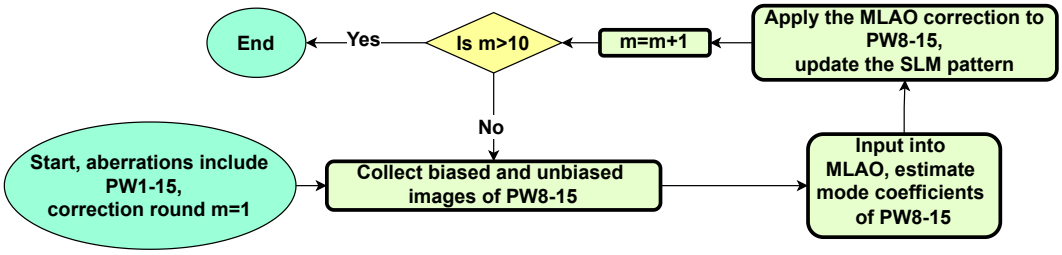
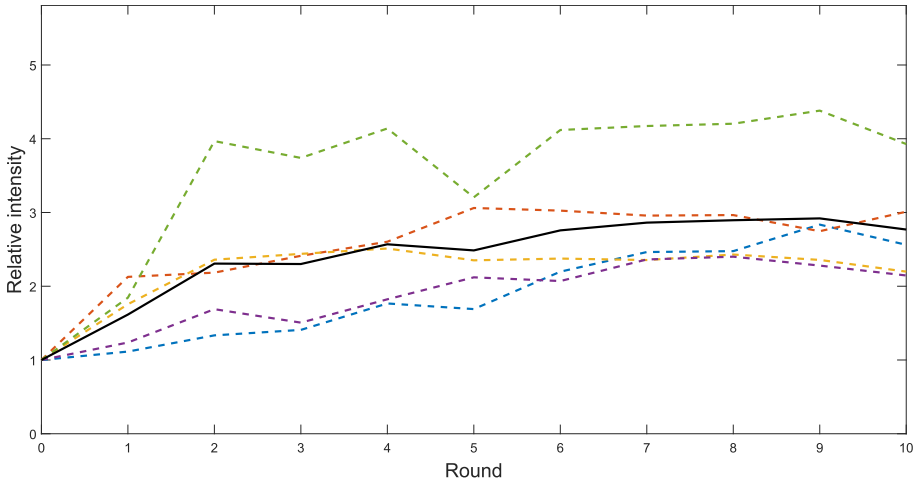
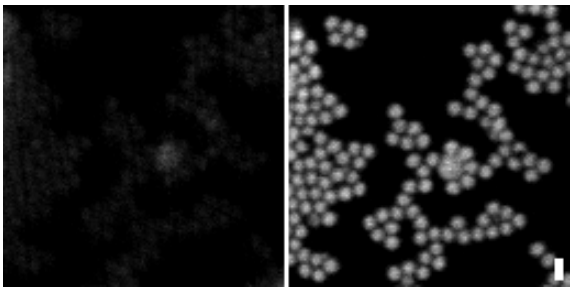


Figure 9.11: Working process of MLAO correction of trained modes in the presence of untrained polar Walsh modes. $PW1-15$ are present, but only $PW8-15$ are corrected over ten rounds.

The resulting relative intensity curves are shown in Figure 9.12. The average relative intensity curve was smooth, and within the first five correction rounds, the average relative intensity reached near-maximal values. Bead image pairs captured before and after correction exhibited significant improvements in both intensity and resolution, even though only a subset of the full polar Walsh mode set was corrected.



(a)



(b)

Figure 9.12: Results of MLAO correcting trained polar Walsh mode set ($PW8-15$) in the presence of untrained polar Walsh mode set ($PW1-7$), under high-SNR and dense specimen structure condition. (a) Relative intensity as a function of correction rounds. Five independent experiments are shown as dotted curves, with the average relative intensity shown in black. (b) Example bead images before (left) and after (right) 10 rounds of aberration correction, the image pair corresponding to one of the experiments above. The scale bar: 2 μm.

From the analysis above, the results suggest that the MLAO is robust and efficient in the presence of an untrained polar Walsh mode subset.

9.6 Comparison of MLAO, CA, and SA under Varying Signal-to-Noise Ratios

In the following sets of experiments, we compared the efficiency and robustness of MLAO with CA and SA under varying SNRs (low and high) with densely populated beads. Only the modes $PW8 - 15$ were added and corrected.

For each imaging scenario, seven experiments were conducted with varying initial aberrations. Note that the FOVs of the seven experiments were not identical. Ten correction rounds were performed for each experiment. The resulting average relative intensities as a function of the number of collected images are shown in Figure 9.13.

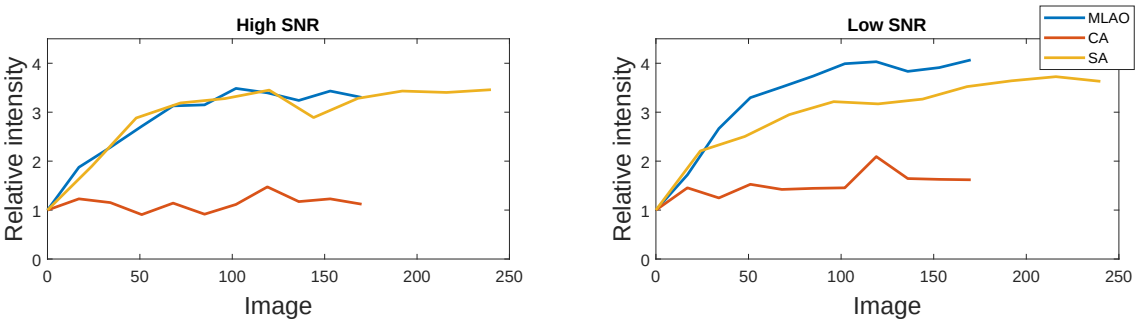


Figure 9.13: Results comparing MLAO with CA and SA under high-/low-SNR and dense specimen structure conditions, correcting $PW8 - 15$. Average relative intensities as a function of the number of collected images are shown. The three curves corresponding to each algorithm are indicated in the top-right legend.

Under high-SNR conditions, MLAO exhibited efficiency comparable to SA and superior to CA, maintaining a relatively stable performance curve. Under low-SNR conditions, MLAO outperformed both SA and CA, demonstrating a faster intensity improvement rate per collected image and achieving a higher final average relative intensity,

while also maintaining performance stability.

Overall, MLAO demonstrates similar or superior working efficiency compared to conventional algorithms, particularly under low SNR conditions. It also exhibits strong robustness across both high and low SNR levels. This makes it especially suitable for biological deep-tissue imaging, which typically involves low SNR environments.

9.7 Summary

We presented experimental results of MLAO when correcting introduced aberrations in a two-photon microscope on a bead specimen. Results suggested that MLAO is robust to different SNR levels and sample structure distributions with high efficiency in image quality improvement. We confirmed that the MLAO trained for certain modes cannot be easily applied to modes that are not used for training, but are robust and efficient in the presence of those untrained modes. When compared with conventional algorithms, MLAO demonstrated robustness across various SNR conditions and exhibited superior image quality improvement efficiency and results, particularly under low-SNR conditions. In conclusion, MLAO provides an innovative, robust, and efficient solution for complex wavefront compensation, outperforming conventional algorithms—particularly under low-SNR conditions relevant to deep-tissue imaging.

Chapter 10

Conclusions and Outlooks

10.1 Conclusions

In this thesis, we discussed the implementation of Walsh modes and neural networks (NN) based on such modes for sensorless adaptive optics aberration correction.

We introduced discrete and binary Walsh modes as alternatives to continuous Zernike modes, which are no longer suitable for deep scattering specimens where the wavefront is highly complex. In practice, when it becomes evident that the complexity of the aberrations exceeds what can be adequately represented by the Zernike modes, Walsh modes should be employed instead to provide a more suitable basis for correction. Based on a single Walsh mode, we proposed a sinusoidal image quality metric function. We further derived the lattice-like metric space for multiple Walsh modes. Exploiting the sinusoidal property, we developed a Walsh mode coefficient estimation algorithm using the single Walsh mode metric function, analogous to phase-shifting interferometry.

However, the single Walsh mode estimation algorithm faces challenges in multidimensional metric spaces, as it often requires excessive image acquisitions. To address this, we adopted a NN to handle the lattice-like metric space. Furthermore, we proposed that, due to the multi-maxima property of this space, the NN search space must be confined to the fundamental unit of the lattice, known as a Voronoi cell, which contains only a single maximum.

We described in detail the construction of the aforementioned NN for the scalar metric space (with the metric value chosen as the Strehl ratio), including the training dataset generation process, NN architecture, and optimization. In practice, however, due to the complexity of aberrated wavefronts, the dimensionality of the Walsh mode set can be very high, making NN training on the full set both complex and time-consuming. To address this limitation, we proposed training a low-dimensional NN on a subset of Walsh modes and then iteratively correcting the remaining modes using the same NN. This approach is justified by the fact that the lattices of these Walsh mode subsets are equivalent. The optimal Walsh mode subset dimension was found to be 8.

Based on this approach, we developed a 7NN + PSI method (the number before NN describes how many Walsh modes the NN can correct in each iteration, and the name after + indicates the extra correcting method it uses). The network was trained on $PW1-7$ and applied to the subsequent seven Walsh modes in each subset, while the lowest-index mode was corrected using the single Walsh mode estimation algorithm, denoted as PSI. The 7NN + PSI method demonstrated better performance than the conventional $2N+1$ algorithm (CA) across various signal-to-noise ratios (SNRs), but its effectiveness was limited when applied to high-dimensional Walsh mode sets. To further

improve NN efficiency by eliminating the additional PSI method, an 8NN was trained directly on a complete set of eight modes ($PW8$–15). This network achieved good performance on $PW8-15$, and when applied to the first polar Walsh mode set, which includes the piston mode, a 'phantom' method was employed. Simulations confirmed that this phantom method is effective and that the 8NN achieves higher efficiency than the 7NN + PSI. Furthermore, more rigorous comparisons were conducted between 8NN and the sequential $3N$ algorithm (SA). The results showed that 8NN performed at a level comparable to SA in simulations, but also demonstrated the potential to surpass SA. Experimental results confirmed this prediction, showing that 8NN can indeed outperform SA.

However, the 8NN is limited by its reliance on an intensity reference and by not fully exploiting the information contained in the images. To address these limitations, the MLAO method was proposed. MLAO effectively utilizes extended images. The pseudo-PSF input data pre-processing method was designed to extract aberration structure information contained in the images. Both the MLAO architecture and training data design—including different input data categories such as intensity-based and pseudo-PSF-based data—aim to fully exploit the information contained in the images. The detailed data generation procedure and MLAO architecture have been described. Furthermore, by analyzing the relationship between the aberration information content of each input data category and the variance of the corresponding trained MLAO parameters, this approach not only informs MLAO architecture modifications to improve performance but also provides deeper insight into the imaging process, guiding further research.

Simulations were conducted to compare the performance of MLAO, CA, and SA under various conditions. When applied to the Walsh mode subset on which MLAO was trained, MLAO demonstrated higher efficiency than SA and robustness to noise. In contrast, SA exhibited limited robustness when confronted with the high-dimensional metric structure. However, MLAO was found to be unsuitable for untrained Walsh mode subsets. Consequently, a new MLAO was trained specifically on that untrained subset. This combined MLAO demonstrated robustness and superior efficiency across most conditions compared with SA. Experimental results demonstrated the efficiency and robustness of MLAO in improving image quality across most SNRs and specimen densities. They also confirmed that MLAO is not suitable for an untrained Walsh mode set but remains efficient and robust when untrained Walsh modes are present, provided only the trained modes are corrected. Finally, comparison with conventional algorithms showed that MLAO demonstrates robustness across a wide range of SNR conditions and exhibits superior image quality improvement efficiency and outcomes than the conventional algorithms, particularly under low-SNR conditions, making it a promising approach for deep-tissue imaging

10.2 Outlooks

Neural networks based on Walsh modes have significant potential for future development, particularly in deep imaging scattering compensation. They may also be applied to other areas, such as atmospheric aberration correction. Walsh modes are promising not only for aberration correction applications but also for their unique mathematical

properties, making them relevant to a wide range of research studies.

Both simulations and experimental demonstrations demonstrated the efficiency and robustness of the established MLAO method based on Walsh modes in aberration correction. The MLAO trained on $PW8-15$ has been thoroughly tested experimentally; however, the newly trained MLAO based on $PW1-7$ has not yet been tested experimentally. Follow-up experiments on this specific MLAO and on the combined MLAO set will be carried out in the future.

Although the MLAO was well designed, refinements to its construction and training process—such as the number of trainable parameters and advanced training strategies—can be implemented. Simulations of such newly trained MLAO will be conducted under harsher conditions, including lower SNRs or more complex specimen structures.

The MLAO we designed shows promise for deep imaging scattering compensation. Its robustness and higher efficiency compared with traditional algorithms make it well-suited for live imaging applications, reducing aberration correction time and specimen exposure, which helps mitigate issues such as photobleaching. Currently, simulations and experimental demonstrations yield encouraging results; however, no experiments on live specimens have yet been conducted. A follow-up project will examine the established MLAOs on live, scattering, and deep specimens in future work.

Walsh modes have great potential not only for aberration correction but also for their promising mathematical properties. A lattice generated from Walsh modes can guide the estimation algorithm and may be applicable to other engineering or mathematical problems. Some aspects remain unexplored, such as the phenomenon whereby the system becomes trapped in the zero regions of the correction space discussed in

Section 8.1, which will be investigated in the future. The applications of these lattices and Voronoi cells will also be explored. Using machine learning methods to study these lattice structures may provide insights into their underlying mathematical problems.

Analyzing the variance of input parameters can not only help improve the neural network construction but also provide insights into the physical imaging process, which is valuable for further research. This approach can also be applied to other studies, for example, to evaluate the information content of different types of data.

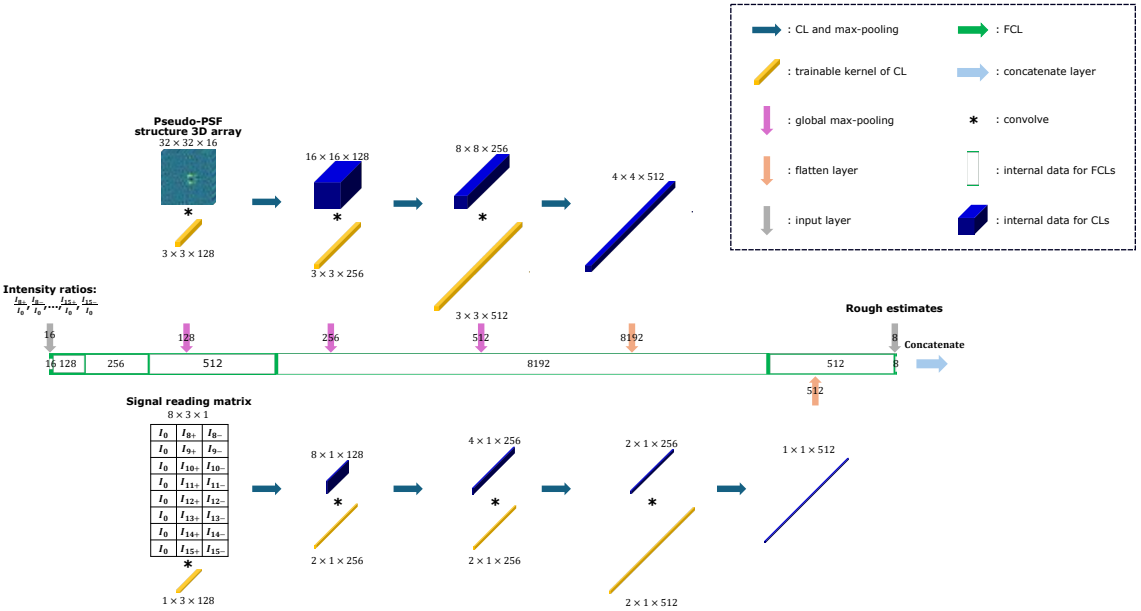
Finally, as the field of machine learning is rapidly evolving, many new cutting-edge methods are being developed, such as the self-supervised learning method. We plan to combine these methods with Walsh modes in future work.

Appendix A

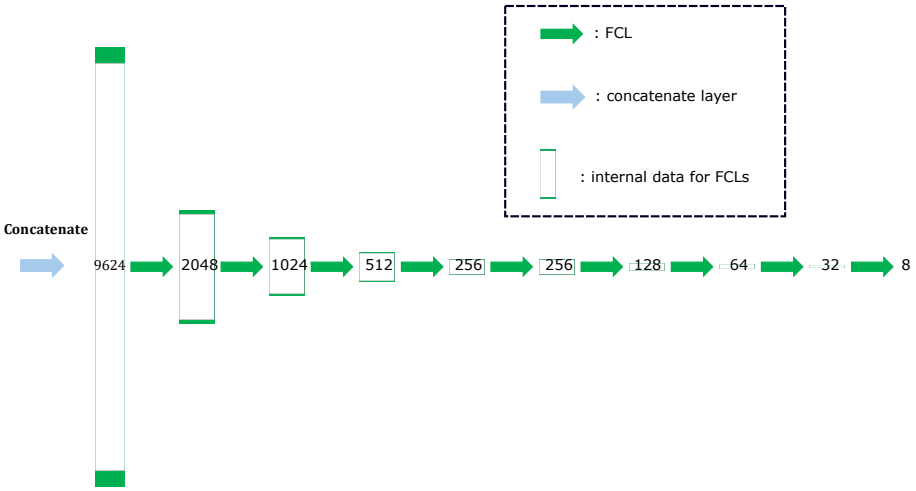
MLAO Architecture

A detailed structure of the MLAO is shown in Figure A.1. The optimizer was Adam. The first three CLs processing the pseudo-PSF 3D array used the HeNormal initializer, ReLU activation, and an L1L2 regularizer, with padding set to 'same'. All three max-pooling layers were 2×2 in size with a 2×2 stride. After each 'CL + max-pooling' module, a global max-pooling layer was applied to extract features for further analysis. This arrangement enabled feature extraction from the pseudo-PSF structure at multiple spatial scales. The outputs of the global max-pooling layers, along with the flattened pseudo-PSF structure convolved by all modules, were then fed into the concatenate layer.

For the signal reading matrix, the first CL used a 1×1 stride, tanh activation, Glorot Uniform initialiser, L1L2 regularizer, and 'valid' padding. In the subsequent three CLs, the padding was changed to 'same'. A max-pooling layer of size 2×1 was applied after the second, third, and fourth CL. The output of this module was flattened and concatenated with the outputs from the pseudo-PSF 3D array module, the intensity ratios, and



(a)



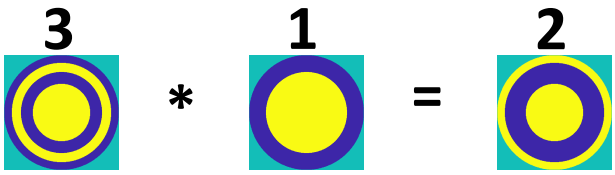
(b)

Figure A.1: Detailed MLAO structure. (a) The first part of the MLAO architecture. The dashed box at the top right denotes the operations, data, and neural network components. (b) The second part of the MLAO architecture. Part (a) is concatenated with part (b) through the concatenate operation, indicated by the light blue arrow. The figures in the dashed box represent operations, data, and neural network components as well.

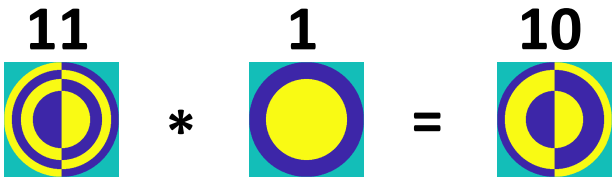
the rough estimates. This concatenated vector was processed by FCLs with HeNormal kernel initializers and ReLU activations. At the output layer, the MLAO generated eight predictions, with the activation set to linear and the initializer set to zeros.

Appendix B

Pairwise Multiplication of Polar Walsh Modes



(a)



(b)

Figure B.1: Examples of pairwise multiplication. The operator $*$ denotes pixel-wise multiplication between two polar Walsh modes, which producing a new mode.

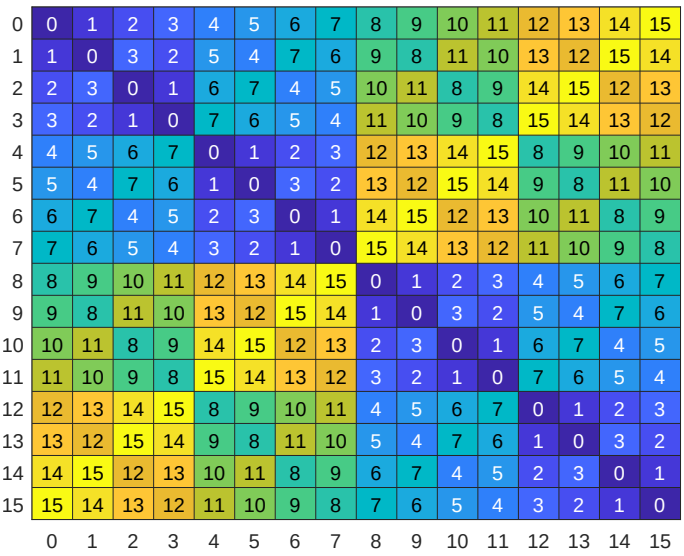


Figure B.2: Results of pairwise multiplication of $PW0-15$. The x- and y-axes denote the indices of the multiplied polar Walsh modes, and the heatmap entries indicate the indices of the resulting modes.

Bibliography

[1] Q. Zhang, Q. Hu, C. Berlage, et al., "Adaptive optics for optical microscopy", Biomedical Optics Express **14**, 1732–1756 (2023).

[2] C. Rodríguez and N. Ji, "Adaptive optical microscopy for neurobiology", Current opinion in neurobiology **50**, 83–91 (2018).

[3] M. J. Booth, "Adaptive optics in microscopy", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **365**, 2829–2843 (2007).

[4] N. Ji, "Adaptive optical fluorescence microscopy", Nature methods **14**, 374–380 (2017).

[5] M. Booth and K. Hampson, "Adaptive optics for microscopy", in *Handbook of adaptive optics* (CRC Press, 2024), pp. 41–58.

[6] M. J. Booth, "Adaptive optical microscopy: the ongoing quest for a perfect image", Light: Science & Applications **3**, e165–e165 (2014).

[7] J. M. Girkin, S. Poland, and A. J. Wright, "Adaptive optics for deeper imaging of biological samples", Current opinion in biotechnology **20**, 106–110 (2009).

[8] D. Débarre, E. J. Botcherby, M. J. Booth, et al., "Adaptive optics for structured illumination microscopy", Optics express **16**, 9290–9305 (2008).

[9] M. Booth, D. Andrade, D. Burke, et al., "Aberrations and adaptive optics in super-resolution microscopy", Microscopy **64**, 251–261 (2015).

[10] D. Débarre, E. J. Botcherby, T. Watanabe, et al., "Image-based adaptive optics for two-photon microscopy", Optics letters **34**, 2495–2497 (2009).

[11] F. Wang, "Wavefront sensing through measurements of binary aberration modes", Applied optics **48**, 2865–2870 (2009).

[12] S. Wang, P. Yang, B. Xu, et al., "Shack-hartmann wavefront sensing based on binary-aberration-mode filtering", Optics Express **23**, 5052–5064 (2015).

[13] F. Wang, "Control of deformable mirror with light-intensity measurements through single-mode fiber", Applied Optics **49**, G60–G66 (2010).

[14] S. Wang, P. Yang, M. Ao, et al., "Wavefront sensing by means of binary intensity modulation", Applied Optics **53**, 8342–8349 (2014).

[15] F. Wang, "High-contrast imaging via modal convergence of deformable mirror", The Astrophysical Journal **751**, 83 (2012).

[16] S. M. Popoff, G. Lerosey, R. Carminati, et al., "Measuring the transmission matrix in optics: an approach to the study and control¡? format?¿ of light propagation in disordered media", Physical review letters **104**, 100601 (2010).

[17] B. Mastiani and I. M. Vellekoop, "Noise-tolerant wavefront shaping in a hadamard basis", Optics express **29**, 17534–17541 (2021).

[18] X. Tao, T. Lam, B. Zhu, et al., "Three-dimensional focusing through scattering media using conjugate adaptive optics with remote focusing (caorf)", Optics express **25**, 10368–10383 (2017).

[19] A. Thendiyammal, G. Osnabrugge, T. Knop, et al., "Model-based wavefront shaping microscopy", Optics letters **45**, 5101–5104 (2020).

[20] B. Mastiani, G. Osnabrugge, and I. M. Vellekoop, "Wavefront shaping for forward scattering", Optics express **30**, 37436–37445 (2022).

[21] B. Rauer, H. B. de Aguiar, L. Bourdieu, et al., "Scattering correcting wavefront shaping for three-photon microscopy", Optics Letters **47**, 6233–6236 (2022).

[22] D. Sinefeld, F. Xia, M. Wang, et al., "Three-photon adaptive optics for mouse brain imaging", Frontiers in neuroscience **16**, 880859 (2022).

[23] V. J. Parot, C. Sing-Long, Y. Adam, et al., "Compressed hadamard microscopy for high-speed optically sectioned neuronal activity recordings", Journal of Physics D: Applied Physics **52**, 144001 (2019).

[24] A. Facomprez, E. Beaurepaire, and D. Débarre, "Accuracy of correction in modal sensorless adaptive optics", Optics express **20**, 2598–2612 (2012).

[25] W. J. Croft, *Under the microscope: a brief history of microscopy*, Vol. 5 (World Scientific, 2006).

[26] C. S. Ball, "The early history of the compound microscope", Bios **37**, 51–60 (1966).

[27] R. Hooke, "Micrographia", in *Literature and science, 1660-1834, part ii vol 5* (Routledge, 2024), pp. 1–14.

[28] H. Gest, "The discovery of microorganisms by robert hooke and antoni van leeuwenhoek, fellows of the royal society", Notes and Records of the Royal Society of London **58**, 187–201 (2004).

[29] A. Badiye, A. Raina, R. Kakad, et al., "Introduction to compound microscope", in *Forensic microscopy* (CRC Press, 2022), pp. 72–82.

[30] A. Holik, "Optical microscopy", Encyclopedia of Materials: Science and Technology, 6458–6463 (2001).

[31] P. W. Hawkes and J. C. Spence, *Springer handbook of microscopy* (Springer Nature, 2019).

[32] M. Locquin and M. Langeron, *Handbook of microscopy* (Butterworth-Heinemann, 2013).

[33] J. Young and F. Roberts, "A flying-spot microscope", Nature **167**, 231–231 (1951).

[34] V. Wilke, "Optical scanning microscopy—the laser scan microscope", Scanning **7**, 88–96 (1985).

[35] M. Abramowitz, K. R. Spring, H. E. Keller, et al., "Basic principles of microscope objectives", BioTechniques **33**, 772–781 (2002).

[36] W. J. Smith, *Modern optical engineering* (McGraw Hill, 2000).

[37] B. Valeur and M. N. Berberan-Santos, "A brief history of fluorescence and phosphorescence before the emergence of quantum theory", Journal of Chemical Education **88**, 731–738 (2011).

[38] G. G. Stokes, "Xxx. on the change of refrangibility of light", Philosophical transactions of the Royal Society of London, 463–562 (1852).

[39] M. Renz, "Fluorescence microscopy—a historical and technical perspective", Cytometry Part A **83**, 767–779 (2013).

[40] N. Rusk, "The fluorescence microscope", Nature Cell Biology **11**, S8–S9 (2009).

[41] G. Y. Wiederschain, "The molecular probes handbook. a guide to fluorescent probes and labeling technologies", Biochemistry (Moscow) **76**, 1276–1277 (2011).

[42] J. K. Sanders and S. E. Jackson, "The discovery and development of the green fluorescent protein, gfp", Chemical Society Reviews **38**, 2821–2822 (2009).

[43] D. C. Prasher, V. K. Eckenrode, W. W. Ward, et al., "Primary structure of the aequorea victoria green-fluorescent protein", Gene **111**, 229–233 (1992).

[44] M. Chalfie, Y. Tu, G. Euskirchen, et al., "Green fluorescent protein as a marker for gene expression", Science **263**, 802–805 (1994).

[45] R. Heim and R. Y. Tsien, "Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer", Current biology **6**, 178–182 (1996).

[46] J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy", Nature methods **2**, 910–919 (2005).

[47] A. D. Elliott, "Confocal microscopy: principles and modern practices", Current protocols in cytometry **92**, e68 (2020).

[48] M. Minsky, "Microscopy apparatus us patent 3013467", USP Office, Ed. US **658** (1961).

[49] M. Minsky, "Memoir on inventing the confocal scanning microscope", (1988).

[50] T. H. Maiman, "Stimulated optical radiation in ruby", nature **187**, 493–494 (1960).

[51] P. Davidovits and M. D. Egger, "Scanning laser microscope", Nature **223**, 831–831 (1969).

[52] J. Pawley, *Handbook of biological confocal microscopy*, Vol. 236 (Springer Science & Business Media, 2006).

[53] G. S. Kino and T. R. Corle, *Confocal scanning optical microscopy and related imaging systems* (Academic Press, 1996).

[54] F. Helmchen and W. Denk, "Deep tissue two-photon microscopy", Nature methods **2**, 932–940 (2005).

[55] C. Sheppard, J. Gannaway, R. Kompfner, et al., "The scanning harmonic optical microscope", IEEE Journal of Quantum electronics **13**, 912–912 (1977).

[56] J. Gannaway and C. Sheppard, "Second-harmonic imaging in the scanning optical microscope", Optical and Quantum Electronics **10**, 435–439 (1978).

[57] C. Sheppard and R. Kompfner, "Resonant scanning optical microscope", Applied optics **17**, 2879–2882 (1978).

[58] A. Diaspro, *Confocal and two-photon microscopy: foundations, applications and advances* (John Wiley & Sons, 2001).

[59] W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy", Science **248**, 73–76 (1990).

[60] M. Pawlicki, H. A. Collins, R. G. Denning, et al., "Two-photon absorption and the design of two-photon dyes", Angewandte Chemie International Edition **48**, 3244–3266 (2009).

[61] W. Kaiser and C. Garrett, "Two-photon excitation in ca f 2: eu 2+", Physical review letters **7**, 229 (1961).

[62] K. Svoboda and S. M. Block, "Biological applications of optical forces", Annual review of biophysics and biomolecular structure **23**, 247–285 (1994).

[63] R. W. Ditchburn, *Light* (Courier Corporation, 2013).

[64] M. Katz, *Introduction to geometrical optics* (World Scientific Publishing Company, 2002).

[65] L. S. Pedrotti, "Basic geometrical optics", Fundamentals of photonics **1**, 73–115 (2008).

[66] J. W. Goodman, *Introduction to fourier optics* (Roberts and Company publishers, 2005).

[67] J. M. Cowley, *Diffraction physics* (Elsevier, 1995).

[68] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).

[69] A. Messiah, *Quantum mechanics* (Courier Corporation, 2014).

[70] E. Hecht, *Optics* (Pearson Education, Inc, 2017).

[71] Q. Hu, "Adaptive optics for corrections of phase and polarisation state aberrations in microscopes", PhD thesis (University of Oxford, 2021).

[72] P. M. Whelan and M. J. Hodgson, *Essential principles of physics.* (London, 1987).

[73] K. R. Castleman and I. T. Young, "Fundamentals of microscopy", in *Microscope image processing* (Elsevier, 2023), pp. 11–25.

[74] *Point spread function*, `https://www.telescope-optics.net/diffractionimage.htm`.

[75] J. M. Beckers, "Adaptive optics for astronomy: principles, performance, and applications", Annual review of astronomy and astrophysics **31**, 13–62 (1993).

[76] C. J. Sheppard and M. Gu, "Aberration compensation in confocal microscopy", Applied optics **30**, 3563–3568 (1991).

[77] J. Pawley, *Handbook of biological confocal microscopy*, Vol. 236 (Springer Science & Business Media, 2006), pp. 239–249.

[78] A. Diaspro, F. Federici, and M. Robello, "Influence of refractive-index mismatch in high-resolution three-dimensional confocal microscopy", Applied optics **41**, 685–690 (2002).

[79] M. J. Booth, M. A. Neil, and T. Wilson, "Aberration correction for confocal imaging in refractive-index-mismatched media", Journal of microscopy **192**, 90–98 (1998).

[80] A. Egner, M. Schrader, and S. W. Hell, "Refractive index mismatch induced intensity and phase variations in fluorescence confocal, multiphoton and 4pi-microscopy", Optics communications **153**, 211–217 (1998).

[81] A. Egner and S. W. Hell, "Aberrations in confocal and multi-photon fluorescence microscopy induced by refractive index mismatch", in *Handbook of biological confocal microscopy* (Springer, 2006), pp. 404–413.

[82] M. J. Booth and T. Wilson, "Refractive-index-mismatch induced aberrations in single-photon and two-photon microscopy and the use of aberration correction", Journal of biomedical optics **6**, 266–272 (2001).

[83] A. Egner and S. Hell, "Equivalence of the huygens–fresnel and debye approach for the calculation of high aperture point-spread functions in the presence of refractive index mismatch", Journal of Microscopy **193**, 244–249 (1999).

[84] M. Booth and T. Wilson, "Strategies for the compensation of specimen-induced spherical aberration in confocal microscopy of skin", Journal of microscopy **200**, 68–74 (2000).

[85] R. K. Tyson and B. W. Frazier, *Principles of adaptive optics* (CRC press, 2022).

[86] K. M. Hampson, R. Turcotte, D. T. Miller, et al., "Adaptive optics for high-resolution imaging", Nature Reviews Methods Primers **1**, 68 (2021).

[87] N. Ji, D. E. Milkie, and E. Betzig, "Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues", Nature methods **7**, 141–147 (2010).

[88] P.-Y. Madec, "Overview of deformable mirror technologies for adaptive optics and astronomy", in Adaptive optics systems iii, Vol. 8447 (International Society for Optics and Photonics, 2012), pp. 844–705.

[89] *Standard deformable mirrors*, (2020) https://bostonmicromachines.com/standard-deformable-mirrors/.

[90] L. J. Hornbeck, "Deformable-mirror spatial light modulators", in Spatial light modulators and applications iii, Vol. 1150 (International Society for Optics and Photonics, 1990), pp. 86–103.

[91] M. J. Booth, *A basic introduction to adaptive optics for microscopy* (Oct. 2019).

[92] U. Efron, *Spatial light modulator technology: materials, devices, and applications*, Vol. 47 (CRC press, 1994).

[93] Y. Yang, A. Forbes, and L. Cao, "A review of liquid crystal spatial light modulators: devices and applications", Opto-Electronic Science **2**, 230026–1 (2023).

[94] P. Salter, M. Booth, et al., "Designing and aligning optical systems incorporating liquid crystal spatial light modulators (slms)", Department of Engineering, University of Oxford, vr1. 0, doi **10**.

[95] K. Philipp, F. Lemke, S. Scholz, et al., "Diffraction-limited axial scanning in thick biological tissue with an aberration-correcting adaptive lens", Scientific reports **9**, 9532 (2019).

[96] P. Pozzi, M. Quintavalla, A. Wong, et al., "Plug-and-play adaptive optics for commercial laser scanning fluorescence microscopes based on an adaptive lens", Optics Letters **45**, 3585–3588 (2020).

[97] P. Rajaeipour, "Deformable phase plate–a new technology for plug-and-play adaptive optics", in Adaptive optics and applications (Optica Publishing Group, 2022), OTh4B–3.

[98] K. Banerjee, P. Rajaeipour, Ç. Ataman, et al., "Optofluidic adaptive optics", Applied optics **57**, 6338–6344 (2018).

[99] P. Rajaeipour, K. Banerjee, A. Dorn, et al., "Cascading optofluidic phase modulators for performance enhancement in refractive adaptive optics", Advanced Photonics **2**, 066005–066005 (2020).

[100] P. Rajaeipour, K. Banerjee, H. Zappe, et al., "Optimization-based real-time open-loop control of an optofluidic refractive phase modulator", Applied optics **58**, 1064–1072 (2019).

[101] K. Banerjee, P. Rajaeipour, H. Zappe, et al., "Refractive opto-fluidic wavefront modulator with electrostatic push-pull actuation", in Adaptive optics and wavefront control for biological systems v, Vol. 10886 (SPIE, 2019), pp. 25–32.

[102] P. Rajaeipour, M. Sauther, K. Banerjee, et al., "Seventh-order wavefront modulation with a gravity-neutral optofluidic deformable phase plate", Journal of Optical Microsystems **1**, 034502–034502 (2021).

[103] B. C. Platt and R. Shack, *History and principles of shack-hartmann wavefront sensing*, 2001.

[104] R. Ragazzoni, E. Diolaiti, and E. Vernet, "A pyramid wavefront sensor with no dynamic modulation", Optics communications **208**, 51–60 (2002).

[105] F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics", Applied Optics **27**, 1223–1225 (1988).

[106] Y. Wu, M. K. Sharma, and A. Veeraraghavan, "Wish: wavefront imaging sensor with high resolution", Light: Science & Applications **8**, 44 (2019).

[107] H. Medecki, E. Tejnil, K. Goldberg, et al., "Phase-shifting point diffraction interferometer", Optics letters **21**, 1526–1528 (1996).

[108] G. P. Andersen, L. Dussan, F. Ghebremichael, et al., "Holographic wavefront sensor", Optical Engineering **48**, 085801–085801 (2009).

[109] L. Seifert, J. Liesener, and H. J. Tiziani, "The adaptive shack–hartmann sensor", Optics Communications **216**, 313–319 (2003).

[110] W. H. Southwell, "Wave-front estimation from wave-front slope measurements", JOSA **70**, 998–1006 (1980).

[111] D. L. Fried, "Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements", JOSA **67**, 370–375 (1977).

[112] J. Herrmann, "Least-squares wave front errors of minimum norm", JOSA **70**, 28–35 (1980).

[113] O. Azucena, J. Crest, S. Kotadia, et al., "Adaptive optics wide-field microscopy using direct wavefront sensing", Optics letters **36**, 825–827 (2011).

[114] P. Vermeulen, E. Muro, T. Pons, et al., "Adaptive optics for fluorescence wide-field microscopy using spectrally independent guide star and markers", Journal of Biomedical Optics **16**, 076019–076019 (2011).

[115] S. A. Rahman and M. J. Booth, "Direct wavefront sensing in adaptive optical microscopy using backscattered light", Applied optics **52**, 5523–5532 (2013).

[116] X. Tao, J. Crest, S. Kotadia, et al., "Live imaging using adaptive optics with fluorescent protein guide-stars", Optics express **20**, 15969–15982 (2012).

[117] N. Verrier and M. Atlan, "Off-axis digital hologram reconstruction: some practical considerations", Applied optics **50**, H136–H146 (2011).

[118] K. Creath, "Phase-shifting speckle interferometry", Applied Optics **24**, 3053–3058 (1985).

[119] K. Creath, "V phase-measurement interferometry techniques", in *Progress in optics*, Vol. 26 (Elsevier, 1988), pp. 349–393.

[120] L. E. Bradsby, A. A. Voitiv, and M. E. Siemens, "Two-step phase-shifting interferometry for phase-resolved imaging from a spatial light modulator", Journal of Optics **27**, 055704 (2025).

[121] A. Jesacher and M. J. Booth, "Sensorless adaptive optics for microscopy", in Mems adaptive optics v, Vol. 7931 (SPIE, 2011), pp. 115–123.

[122] M. Lamb, D. R. Andersen, J.-P. Véran, et al., "Non-common path aberration corrections for current and future ao systems", in Adaptive optics systems iv, Vol. 9148 (International Society for Optics and Photonics, 2014), p. 914857.

[123] Q. Hu, J. Wang, J. Antonello, et al., "A universal framework for microscope sensorless adaptive optics: generalized aberration representations", APL Photonics **5** (2020).

[124] J.-H. Park, W. Sun, and M. Cui, "High-resolution in vivo imaging of mouse brain through the intact skull", Proceedings of the National Academy of Sciences **112**, 9236–9241 (2015).

[125] L. Kong and M. Cui, "In vivo neuroimaging through the highly scattering tissue via iterative multi-photon adaptive compensation technique", Optics express **23**, 6145–6150 (2015).

[126] J. Tang, R. N. Germain, and M. Cui, "Superpenetration optical microscopy by iterative multiphoton adaptive compensation technique", Proceedings of the National Academy of Sciences **109**, 8434–8439 (2012).

[127] N. Ji, T. R. Sato, and E. Betzig, "Characterization and adaptive optical correction of aberrations during in vivo imaging in the mouse cortex", Proceedings of the National Academy of Sciences **109**, 22–27 (2012).

[128] C. Wang and N. Ji, "Pupil-segmentation-based adaptive optical correction of a high-numerical-aperture gradient refractive index lens for two-photon fluorescence endoscopy", Optics letters **37**, 2001–2003 (2012).

[129] D. E. Milkie, E. Betzig, and N. Ji, "Pupil-segmentation-based adaptive optical microscopy with full-pupil illumination", Optics letters **36**, 4206–4208 (2011).

[130] K. M. Hampson, J. Antonello, R. Lane, et al., *Sensorless adaptive optics* (Oct. 2020).

[131] *Zernike polynomials*, (2003) https://wp.optics.arizona.edu/jcwyant/wp-content/uploads/sites/13/2016/08/ZernikePolynomialsForTheWeb.pdf.

[132] F. Zernike, "Diffraction theory of the knife-edge test and its improved form, the phase-contrast method", Monthly Notices of the Royal Astronomical Society **94**, 377–384 (1934).

[133] V. Lakshminarayanan and A. Fleck, "Zernike polynomials: a guide", Journal of Modern Optics **58**, 545–561 (2011).

[134] *Zernike polynomial*, https://mathworld.wolfram.com/ZernikePolynomial.html.

[135] B. H. Shakibaei and R. Paramesran, "Recursive formula to compute zernike radial polynomials", Optics letters **38**, 2487–2489 (2013).

[136] *Strehl ratio — telescope-optics.net*, https://www.telescope-optics.net/Strehl.htm, [Accessed 16-09-2025].

[137] A. Van den Bos, "Aberration and the strehl ratio", Journal of the Optical Society of America A **17**, 356–358 (2000).

[138] L. Murray, "Smart optics: wavefront sensor-less adaptive optics-image correction through sharpness maximisation", NUI Galw (2006).

[139] M. J. Booth, "Wave front sensor-less adaptive optics: a model-based approach using sphere packings", Optics express **14**, 1339–1352 (2006).

[140] M. J. Booth, M. A. Neil, R. Juškaitis, et al., "Adaptive aberration correction in a confocal microscope", Proceedings of the National Academy of Sciences **99**, 5788–5792 (2002).

[141] M. J. Booth, M. A. Neil, and T. Wilson, "New modal wave-front sensor: application to adaptive confocal fluorescence microscopy and two-photon excitation fluorescence microscopy", Journal of the Optical Society of America A **19**, 2112–2120 (2002).

[142] M. Ren, J. Chen, D. Chen, et al., "Aberration-free 3d imaging via dmd-based two-photon microscopy and sensorless adaptive optics", Optics letters **45**, 2656–2659 (2020).

[143] I. Casas, "Networks, neural", (2020).

[144] C. M. Bishop, "Neural networks and their applications", Review of scientific instruments **65**, 1803–1832 (1994).

[145] Y. Jin, Y. Zhang, L. Hu, et al., "Machine learning guided rapid focusing with sensor-less aberration corrections", Optics express **26**, 30162–30171 (2018).

[146] Q. Hu, M. Hailstone, J. Wang, et al., "Universal adaptive optics for microscopy through embedded neural network control", Light: Science & Applications **12**, 270 (2023).

[147] D. Saha, U. Schmidt, Q. Zhang, et al., "Practical sensorless aberration estimation for 3d microscopy with deep learning", Optics express **28**, 29044–29053 (2020).

[148] B. P. Cumming and M. Gu, "Direct determination of aberration functions in microscopy by an artificial neural network", Optics Express **28**, 14511–14521 (2020).

[149] I. Vishniakou and J. D. Seelig, "Wavefront correction for adaptive optics with reflected light and deep neural networks", Optics express **28**, 15459–15471 (2020).

[150] Y. Zheng, J. Chen, C. Wu, et al., "Adaptive optics for structured illumination microscopy based on deep learning", Cytometry Part A **99**, 622–631 (2021).

[151] M. R. Rai, C. Li, H. T. Ghashghaei, et al., "Deep learning-based adaptive optics for light sheet fluorescence microscopy", Biomedical optics express **14**, 2905–2919 (2023).

[152] Y. Guo, L. Zhong, L. Min, et al., "Adaptive optics based on machine learning: a review", Opto-Electronic Advances **5**, 200082–1 (2022).

[153] Y. Zhang, C. Wu, Y. Song, et al., "Machine learning based adaptive optics for doughnut-shaped beam", Optics express **27**, 16871–16881 (2019).

[154] S. Hasegawa and Y. Hayasaki, "Femtosecond laser processing with adaptive optics based on convolutional neural network", Optics and Lasers in engineering **141**, 106563 (2021).

[155] L. Möckl, P. N. Petrov, and W. Moerner, "Accurate phase retrieval of complex 3d point spread functions with deep residual neural networks", Applied physics letters **115** (2019).

[156] E. Bostan, R. Heckel, M. Chen, et al., "Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network", Optica **7**, 559–562 (2020).

[157] I. Muhammad, "Supervised machine learning approaches: a survey", ICTACT Journal on Soft Computing (2015).

[158] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised machine learning: a brief primer", Behavior therapy **51**, 675–687 (2020).

[159] V. Nasteski, "An overview of the supervised machine learning methods", Horizons. b **4**, 51–62 (2017).

[160] S. Dridi, "Supervised learning-a systematic literature review", preprint, Dec (2021).

[161] X. Liu, F. Zhang, Z. Hou, et al., "Self-supervised learning: generative or contrastive", IEEE transactions on knowledge and data engineering **35**, 857–876 (2021).

[162] J. Gui, T. Chen, J. Zhang, et al., "A survey on self-supervised learning: algorithms, applications, and future trends", IEEE Transactions on Pattern Analysis and Machine Intelligence **46**, 9052–9071 (2024).

[163] V. Rani, S. T. Nabi, M. Kumar, et al., "Self-supervised learning: a succinct review", Archives of Computational Methods in Engineering **30**, 2761–2775 (2023).

[164] I. Kang, Q. Zhang, S. X. Yu, et al., "Coordinate-based neural representations for computational adaptive optics in widefield microscopy", Nature Machine Intelligence **6**, 714–725 (2024).

[165] I. N. Papadopoulos, J.-S. Jouhanneau, J. F. Poulet, et al., "Scattering compensation by focus scanning holographic aberration probing (f-sharp)", Nature Photonics **11**, 116–123 (2017).

[166] M. A. May, N. Barré, K. K. Kummer, et al., "Fast holographic scattering compensation for deep tissue biological imaging", Nature Communications **12**, 1–8 (2021).

[167] *Walsh function*, `https://mathworld.wolfram.com/WalshFunction.html`.

[168] K. G. Beauchamp and C. Yuen, "Walsh functions and their applications", IEEE Transactions on Systems, Man, and Cybernetics, 794–795 (1976).

[169] M. De and L. Hazra, "Walsh functions in problems of optical imagery", Optica Acta: International Journal of Optics **24**, 221–234 (1977).

[170] K. G. Beauchamp, "Walsh functions and their applications.", (1975).

[171] H. F. Harmuth, "Applications of walsh functions in communications", IEEE spectrum **6**, 82–91 (1969).

[172] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and synthesis in radio astronomy* (Springer Nature, 2017).

[173] S. Tzafestas, "Walsh functions in signal and systems analysis and design", (1985).

[174] F. Wang, "Utility transforms of optical fields employing deformable mirror", Optics letters **36**, 4383–4385 (2011).

[175] A. Haber and T. Bifano, "Dual-update data-driven control of deformable mirrors using walsh basis functions", Journal of the Optical Society of America A **39**, 459–469 (2022).

[176] C. Stockbridge, Y. Lu, J. Moore, et al., "Focusing through dynamic scattering media", Optics express **20**, 15086–15092 (2012).

[177] D. Sinefeld, H. P. Paudel, D. G. Ouzounov, et al., "Adaptive optics in three-photon fluorescence microscopy", in 2015 conference on lasers and electro-optics (cleo) (IEEE, 2015), pp. 1–2.

[178] Y. M. Wang and C. Yang, *Acoustic-assisted iterative wave form optimization for deep tissue focusing*, US Patent 10,292,589, May 2019.

[179] T. Bifano, Y. Lu, C. Stockbridge, et al., "Mems spatial light modulators for controlled optical transmission through nearly opaque materials", in Mems adaptive optics vi, Vol. 8253 (SPIE, 2012), pp. 149–157.

[180] F. Wang, "Binary phase masking for optical interrogation of matters in turbid media", Optics letters **33**, 2587–2589 (2008).

[181] H. P. Paudel, C. Stockbridge, J. Mertz, et al., "Focusing polychromatic light through strongly scattering media", Optics express **21**, 17299–17308 (2013).

[182] D. Geelen and W. Löffler, "Walsh modes and radial quantum correlations of spatially entangled photons", Optics letters **38**, 4108–4111 (2013).

[183] P. Yang, S. Wang, M. Ao, et al., "A novel wavefront sensing technique for high speed atmospheric measurement based on digital micromirror device", in Op-

tics in atmospheric propagation and adaptive systems xiv, Vol. 8178 (SPIE, 2011), pp. 189–197.

[184] *Walsh system*, http://encyclopediaofmath.org/index.php?title=Walshsystem&oldid=49269.

[185] S. Wolfram, *A new kind of science*, Vol. 5 (Wolfram media Champaign, IL, 2002), p. 1073.

[186] W. Wade, "A walsh system for polar coordinates", Computers & Mathematics with Applications **30**, 221–227 (1995).

[187] Q. Hu, Y. Xiao, J. Cui, et al., "The lattice geometry of walsh-function-based adaptive optics", in Photonics, Vol. 9, 8 (MDPI, 2022), p. 547.

[188] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Vol. 290 (Springer Science & Business Media, 2013).

[189] C. Rycroft, "Voro++: a three-dimensional voronoi cell library in c++", (2009).

[190] *Photon , poisson noise*, (2012) https://people.csail.mit.edu/hasinoff/pubs/hasinoff-photon-2012-preprint.pdf.

[191] K. O'Shea and R. Nash, "An introduction to convolutional neural networks", arXiv preprint arXiv:1511.08458 (2015).

[192] Y. LeCun, L. Bottou, Y. Bengio, et al., "Gradient-based learning applied to document recognition", Proceedings of the IEEE **86**, 2278–2324 (1998).

[193] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", Communications of the ACM **60**, 84–90 (2017).

[194] *Transformation matrices*, (2023) https://chem.libretexts.org/Bookshelves/Physical and Theoretical Chemistry Textbook Maps/Symmetry (Vallance)/01%3A Chapters/1.09%3ATransformationmatrices.

[195] M. J. Booth, "Microscope resolution estimation and normalised coordinates", 10.5281/ZENODO.4302488 (2020).