Radial basis functions for aberration retrieval in adaptive optics

Abhimanyu Gupta

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Radial basis functions for aberration retrieval in adaptive optics

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Abhimanyu Gupta

November 23, 2017

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Abstract

An optical aberration is the departure of optical system performance from that predicted by concepts of paraxial optics. Aberrations are found in most practical imaging systems and their presence deteriorates image quality. The objective of an *adaptive optics* system is the measurement and elimination of such aberrations in incoming wavefront in real time. Wavefront aberration or phase information is not measured directly but has to to be estimated from the intensity distributions on the image plane. Therefore, *aberration retrieval* from intensity data is a special case of the more general problem of *phase retrieval* (PR). Algorithmic phase retrieval based optical wavefront reconstruction offers an attractive means of estimating the aberration from a set of measurements of the point-spread functions (PSF) of some optical systems due to its experimental simplicity.

This thesis deals with the application of the existing PR algorithms to the special case of optical wavefront reconstruction. The representation of the generalized pupil function (GPF) as a linear combination of some convenient basis functions, as opposed to a pixel basis, allows a reduction in computational burden. The complex-valued Zernike polynomials introduced under the Extended Nijboer-Zernike (ENZ) theory constitute the most widely used basis functions. Their inflexibility to represent arbitrary pupil geometry and inability to approximate highly aberrated system are cited as the major limitations in their application to aberration retrieval. On the other hand, Radial Basis Functions (RBF), a standard tool in approximation theory, offer much improved simplicity and geometric flexibility.

The application of Gaussian RBFs to GPF modeling for subsequent aberration retrieval from intensity distributions has been investigated in this thesis. In addition to a real-valued RBF, a complex-valued phase encoded RBF has been proposed followed by an optimal design of RBF hyper-parameters. Finally, the performance of these basis functions is compared with ENZ polynomials in terms of accuracy of aberration retrieval via simulations.

Table of Contents

List of Figures

Acknowledgements

I would like to express my gratitude towards my supervisor prof. M. Verhaegen for giving me an opportunity to work on this project and also for keeping faith in me during the course of the thesis. I would also like to thank Pieter Piscaer for sharing valuable ideas and experiences related to the project. Furthermore, I would like to express my gratitude towards Dean Wilding for helping me with the physics of the system throughout the year.

I am indebted to my parents for their unconditional support. Finally, I would like to thank my friends Divyam, Reoberto, Pépé, Nikolaos, Youssef, Niko, and Salil for their wonderful company. Finally, I would like to thank Nikolaos, Salil and Srinivas for letting me run my simulations on their computers.

Delft, University of Technology
November 23, 2017

Abhimanyu Gupta

Chapter 1

Introduction

The recovery of a signal given the magnitude of its Fourier transform, also known as the *phase retrieval* (PR) problem, is a recurring issue in the field of optics. Since the optical detection devices (e.g. CCD cameras) only measure the photon flux (proportional to the magnitude squared of the field), the phase information is lost. Consequently, access to this information requires additional experimental complexity, like interference with another known field. Alternatively, algorithmic phase retrieval offers a good means of solving the problem by relying on some advanced measurements or prior knowledge to facilitate recovery [1].

This thesis deals with the application of the existing PR algorithms to the special case of optical wavefront reconstruction that arises in the field of adaptive optics.

This chapter provides a brief introduction to the inverse problem of optical wavefront reconstruction and the limitations of the existing approach along with the objective followed by an outline of the thesis report.

1-1 Aberration retrieval

An optical aberration is the departure of optical system performance from that predicted by concepts of paraxial optics [2]. Aberrations are found in most practical imaging systems and their presence deteriorates image quality. Sources of aberrations can either be internal, due to imperfections in or incorrect alignment of the optical components, or it can be external, due to varying refractive index like in the atmosphere (turbulence) or biological samples.

The objective of *adaptive optics* is the measurement and elimination of such aberrations in incoming wavefront, in real time [3]. Wavefront aberration or phase information is not measured directly. Instead, it needs to be estimated from the intensity distributions on a CCD or other area detectors. Therefore, *aberration retrieval* from intensity data is a special case of the more general problem of PR. The application of PR algorithms to aberration retrieval was recently studied in literature [4]. Again, this offers experimental simplicity and does not require any additional optical components.

The procedure of aberration retrieval involves the measurement of the response of an optical system to a *monochromatic point-source* [5]. This impulse response is known as the *point-spread function* (PSF) of the system. The mathematical model of the optical system coalesces all the components into a single plane known as the *pupil plane*. The field at the *image plane* in the focal region, where the light is detected after exiting the optical system, is related to the field in the pupil plane by a diffraction integral.

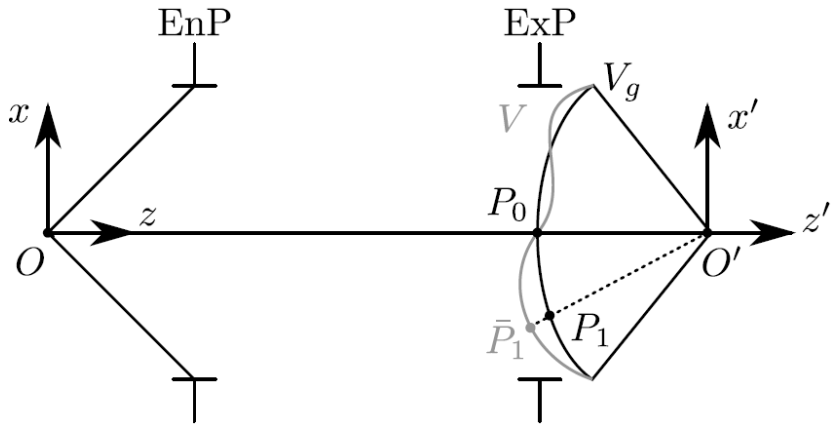


Figure 1-1: Geometrical optics description of the wavefront. Unaberrated light rays from a point-source O enter the optical system through the entrance pupil (EnP), exit at the exit pupil (ExP) and converge in the image plane at O'. An aberrated wave V differs from the Gaussian reference V_g. For the ray depicted by a dashed line, wavefront aberration is given by the optical path length difference $[\bar{P}_1 P_1]$ (*reprinted from* [4]).

In absence of any aberration, the wave follows a reference spherical shape (represented by V_g in Fig. 1-1) that is focused at the image plane. The aberration in the wave are the deviations in the phase and/or amplitude from this reference shape (V in Fig. 1-1). These aberrating effects occurring in the optical system and/or due to the turbulent medium are modeled at the pupil plane as a *generalized pupil function* (GPF) [6]. Therefore, formulation of the aberration retrieval problem requires the parametrization of this GPF.

Expressing the GPF as a linear combination of some convenient basis functions allows the field at the image plane to be represented as a linear combination of transformed basis functions. This is possible due to the linearity property of the diffraction integral. The reconstruction problem is then reduced to retrieval of the coefficients of these basis functions, efficiently calculated using algorithmic PR [4].

Zernike representation of the pupil. Zernike polynomials, the most famous and commonly used set of basis functions for the description of phase distribution in the pupil plane, were introduced as a part of Nijboer-Zernike (NZ) diffraction theory [7]. These form a complete set of orthogonal polynomials defined over a unit disk and have been, since then, extended to represent other pupil shapes [8, 9, 10]. Moreover, these polynomials allowed analytical evaluation of the diffraction integral such that the Zernike coefficients used to describe the phase in the pupil could be translated into an anaytically known contribution to the PSF. However, the theory only permitted phase aberrations upto a few radians. Moreover, the derivation required a uniform amplitude distribution over the pupil [11].

Extended Nijboer-Zernike theory. The aforementioned limitations of the NZ theory were largely mitigated by the introduction of the Extended Nijboer-Zernike (ENZ) theory. Instead of parameterizing the phase aberration, the GPF is represented in a Zernike polynomial basis with only even terms [12], or with complex-valued Zernike polynomial (referred to as ENZ polynomials hereinafter) basis [11]. Both bases require complex-valued coefficients. Again, analytical expression for the PSF has been derived. Like Zernike polynomials, ENZ polynomials offers interesting mathematical properties due to its orthogonality and significant robustness to numerical round-off error.

Although the Zernike modes offer physical interpretation of the phase aberrations present in the wavefront, it is not the case for the ENZ polynomials [11]. These also share the disadvantages of the Zernike polynomials in that they are not flexible to be used with arbitrary aperture shapes and they exhibit smoothing effect that can significantly limit the aberrations that can be modeled [13].

Radial basis functions. Radial basis function, a standard tool in approximation theory, have been efficiently applied in literature to represent the phase aberrations in the wavefront [14] and optical design of freeform surfaces [13]. Their local approximation ability proves to perform better than orthogonal polynomials in representing local deformations in wavefronts and optical freeform surfaces.

Recently, the use of a real-valued Gaussian radial basis function for representing the GPF to evaluate the diffraction integral analytically was studied in the literature [15]. It was demonstrated via simulations that this new scheme was very competitive and provided higher accuracy and speed in comparison to the ENZ theory.

Objective. This thesis focuses on the use of this real-valued Gaussian radial basis function to represent the GPF and subsequent aberration retrieval from intensity distributions in the image plane. In doing so, the optimal design of the hyper-parameters has been studied. The use of a complex-valued radial basis function to represent the GPF, which is essentially complex-valued, has also been proposed in this thesis.

1-2 Outline

The application of radial basis functions to the problem of aberration retrieval from intensity PSF distribution in the image plane is studied.

Chapter 2 deals with the mathematical problem formulation of aberration retrieval introduced in section 1-1.

Chapter 3 is written in the form of a research article. The optimal design of RBF hyper-parameters is proposed. Then the performance of aberration retrieval based on RBF approximation is studied and compared with that based on ENZ theory via simulation. It is noted here that the research article is complete in itself.

Finally, chapter 4 provides the recommendations for future work.

The algorithms used in chapter 3 have been summarized in the Appendices.

Chapter 2

Aberration retrieval in adaptive optics

This chapter deals with the mathematical formulation of the problem of aberration retrieval. The unaberrated imaging model is descibed first. Then the modeling of aberration is given followed by a general problem of aberration retrieval from intensity point spread function. The applications of Nijboer-Zernike theory and its extension to tackle the problem are explained. The application of radial basis function over these existing existing theories is motivated and the design problem is formulated.

2-1 Imaging system

An imaging system usually consists of several lenses and various distances between them. The significant properties of such a system can be completely described by lumping these imaging elements together and specifying only the aggregate properties at the terminals i.e., at the entrance and exit pupil planes as shown in Fig. 2-1. This is what is known as the *generalized model* approach [6]. It is assumed that geometrical optics can be used to describe the passage of light between these terminals.

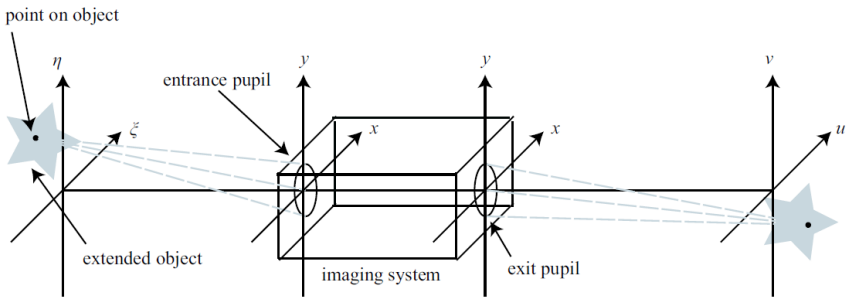


Figure 2-1: Generalized model of imaging system. The imaging system converts a diverging spherical wave at the entrance pupil from a point object into a converging spherical wave at the exit pupil that is finally focused at a point in the image plane (*adapted from* [6]).

A diverging spherical wave, emanating from a point source object, incident on the entrance pupil is converted by the imaging system into a converging spherical wave at the exit pupil, seen at an ideal point in the image plane. Such a system is regarded as being *diffraction-limited* [6].

Point-spread function (PSF) convolution

For a converging wavefront centred on the focal point, the Debye diffraction integral gives the relation between the field in the exit pupil and the field in the focal region [16]. In normalized cylindrical coordinates (ρ, θ) on the exit pupil sphere and (r, ϕ) in the focal region (see Fig. 2-2), this integral takes the form,

$$U(r, \phi, f; \rho, \theta) = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} \exp(if\rho^2) P(\rho, \theta) \times \exp(i2\pi r\rho\cos(\theta - \phi))\rho d\rho d\theta. \tag{2-1}$$

Here, $P(\rho, \theta)$ is known as the pupil function which is unity inside and zero outside the projected aperture. $U(r, \phi, f; \rho, \theta)$ is known as the normalized PSF corresponding to the pupil function. The exit pupil of the optical system is assumed to be a unit disk. f is the defocus parameter ($= 0$ for best focus) normalized with respect to the axial diffraction unit (λ/NA, NA being the image-side numerical aperture of the optical system).

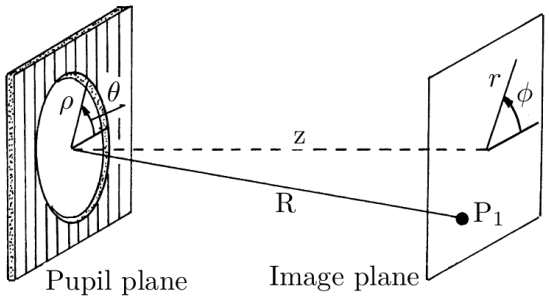


Figure 2-2: Coordinate system for the diffraction integral. The circular pupil has coordinates ρ, θ; the image plane has polar coordinates r, ϕ with z coordinate normal to the pupil plane (*adapted from* [17])

The PSF is measured only in terms of the energy density or the energy flow and is a nonlinear function of the electric field in the focal region,

$$I(r, \phi, f) = |U(r, \phi, f)|^2.$$

Aberration in optical systems

A system is said to have *aberration* if the wavefront leaving the exit pupil departs significantly from ideal spherical shape (reference wave) [5]. Aberrations in image formation can be caused either by inhomogeneities in the medium, for example turbulent atmosphere, or by imperfections in the optical components such as incorrect alignment of lenses. There are two types of deviations from the reference wave. The deviation in phase is referred to as phase aberration,

$\Phi(\rho,\theta)$, and that in amplitude as throughput, $A(\rho,\theta)$, of the optical system. These effects can be modelled using the generalized pupil function (GPF) in a complex phasor form [5],

$$\mathcal{P}(\rho,\theta) = A(\rho,\theta)\exp(i\Phi(\rho,\theta)).$$

Here, $A(\cdot)$ and $\Phi(\cdot)$ are real-valued functions. The GPF is a transmission function that maps the complex amplitude distribution in the object space to the exit pupil sphere in image space and it uniquely characterizes an optical system. The PSF corresponding to GPF can be written similarly as equation (2-1) by replacing $P(\rho,\theta)$ with $\mathcal{P}(\rho,\theta)$. The intensity distribution in the point-spread function strongly depends on the departure of the incident wave from its reference shape.

Purely phase aberrated systems with circular exit pupils are considered under the scope of this thesis. The throughput $A(\rho,\theta)$ is then modelled as an indicator function of the interval $[0,1]$,

$$A(\rho,\theta) = \begin{cases} 1, & 0 \leq \rho \leq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2-2}$$

2-1-1 Inverse problem

The reconstruction of the amplitude and phase of the optical far-field distribution from a particular intensity PSF is known as the *phase retrieval* (PR) problem [5]. Such a procedure is of practical interest as it allows to derive relevant quality data of the optical system, wavefront aberrations in our case, from intensity distributions. The strongly non-linear relationship between the phase departure in the exit pupil of the optical system and the detected intensity in the focal plane leads to an ill-posed problem. The existing PR algorithms have proven to be an attractive way to estimate aberrations for optical systems as they do not require additional optical components [4] (see Appendix A).

The problem formulation requires a convenient and systematic parametrization of the GPF. The easiest choice is a pixel-based parametrization. This serves as the most flexible basis as it can be used with any pupil shape [5]. However, the pixel-basis requires a lot of parameters to be estimated using PR. Other forms of parametrization involve representing the pupil function as a linear superposition of some basis functions [18]. Then, the phase of the GPF can be retrieved by estimating the coefficients of these basis functions which reduces the size of the problem.

$$\hat{\mathcal{P}}(\rho,\theta) = \sum_{n=1}^{N_\beta} \beta_n \psi_n(\rho,\theta) = \boldsymbol{\psi}(\rho,\theta)\boldsymbol{\beta}.$$

Here, the basis functions are written in a matrix format ($\boldsymbol{\psi}(\rho,\theta)$) and the coefficients are collected in a vector $\boldsymbol{\beta}$. The Debye integral in equation (2-1) then gives an analytical expression of the image obtained in the focal region $\hat{U}(r,\phi,f)$. This integral can be evaluated in a discrete format using the FFT algorithm.

Diversity images. To improve the stability of the inversion process, extra information is incorporated in the intensity measurements by, for eg., the addition of known phase aberrations to the system which is a common practice in a more difficult problem of *phase diversity* [19].

One such known aberration is defocus and has been used in this thesis to produce diversity images.

The intensity PSF is measured, using a CCD detector with $N_p = N^2$ pixels, at N_f different positions along the optical axis [4]. The measured pixel values at (r_k, ϕ_k) and defocus f_k are sorted using a single index k, and collected into a vector $\mathbf{I} \in \mathbb{R}^{N_m}$, where $N_m = N_p \cdot N_f$.

The error between the recorded measurements and the predicted PSF for each acquisition can be written as,

$$\epsilon_k = I_k - |\hat{U}(r_k, \phi_k, f_k, \boldsymbol{\beta})|^2.$$

This can be formulated as a PR problem by minimizing the norm of the error,

$$\underset{\boldsymbol{\beta} \in \mathbb{C}^{N_\beta}}{\text{minimize}} \quad \|\boldsymbol{\epsilon}\|. \tag{2-3}$$

2-2 Pupil function modeling

In this section, the existing parametrization of the phase aberration and GPF is first discussed as given by the Nijboer-Zernike theory and then the its extension. The use of radial basis function for local approximations is then motivated to replace these polynomials.

2-2-1 Zernike aberrations

The most famous and common representation of the phase aberration Φ is given by the orthogonal set of basis functions formed by the circle polynomials, Z_n^m, introduced by Zernike [20],

$$\Phi(\rho, \theta) = \sum_{n,m} \alpha_n^m \mathcal{Z}_n^m(\rho, \theta), \tag{2-4}$$

where indices $n \in \mathbb{N}_0$ and $m \in \mathbb{Z}$ respectively denote the radial order and the azimuthal frequency of the Zernike polynomial $\mathcal{Z}_n^m$ such that $n - |m| > 0$ and even. The polynomials are given by the product of a radial polynomial $R_n^{|m|}(\rho)$ and a trigonometric function $\Theta_n^m(\theta)$ with suitable normalization c_n^m,

$$\mathcal{Z}_n^m(\rho, \theta) = c_n^m R_n^{|m|} \Theta_n^m(\theta).$$

where,

$$c_n^m = \begin{cases} \sqrt{n+1} & m = 0 \\ \sqrt{2(n+1)} & m \neq 0 \end{cases}, \quad \Theta_n^m(\theta) = \begin{cases} cos(m\theta) & m \geq 0 \\ -sin(m\theta) & m < 0 \end{cases},$$

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!(\frac{n+m}{2} - s)!(\frac{n-m}{2} - s)!} \rho^{n-2s}.$$

Here, the Zernike polynomials are ordered and normalized according to Noll [21]. This representation of the phase aberration is called *modal representation*, and the basis functions of the decomposition are referred to as *modes*. The first four of these (except the piston) are plotted in Fig. 2-3.

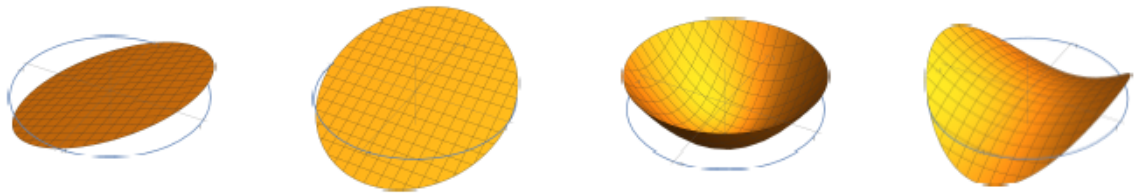


Figure 2-3: First four Zernike polynomials not including the piston (from left to right): $\mathcal{Z}_1^1$, tip; $\mathcal{Z}_1^{-1}$, tilt; $\mathcal{Z}_2^0$, defocus; $\mathcal{Z}_2^2$, astigmatism. These are also some of the most common aberrations found in optical systems (*reprinted from* [3]).

These basis functions have some particularly attractive mathematical properties [3],

- they are orthogonal on a unit disk,
- they have continuous derivatives of all orders,
- $R_n^m(1) = 1$, and thus $\mathcal{Z}_n^m(\rho, \theta)$ becomes unity at the aperture boundary,
- common aberrations observed in optics like tilt, defocus, astigmatism, coma, etc. can be easily expressed through Zernike polynomials,
- they exhibit an inherent smoothing useful for system de-noising.

Another interesting feature of the Zernike circle polynomials is that they allow an analytical evaluation of the Debye integral ((2-1)) [2]. However, this derivation requires the amplitude over the pupil to be uniform and holds only for small aberrations not exceeding a few radians and small defocus [16].

Aberration variance. The Zernike polynomial terms, except the first term, have a zero mean and unit variance. Due to the orthogonality property of the Zernike polynomials, the aberration coefficients α_n^m represent the standard deviation of the corresponding aberration term across the pupil [4]. The root-mean-square (RMS) value of the aberration, in absence of the first term $\mathcal{Z}_0^0$ (piston), can be written in terms of the 2-norm of Zernike coefficients,

$$\mathrm{RMS}(\Phi) = \left(\sum_{n,m} (\alpha_n^m)^2 \right)^{1/2}.$$

Limitations. The major limitation of polynomial representation is that each term extends its influence over the entire pupil [14]. Effort has been made to adapt these polynomials to other pupil shapes, encountered in optics, using Gram-Schmidt orthogonalization [8, 9, 10].

It is very difficult to assess a priori how many terms are necessary to achieve acceptable accuracy. Limiting Zernike analysis to only a few orders may cause incorrect assessment of the severity of the aberration [14]. Although the smoothing effect of the polynomials allows its application to raw data, it significantly limits its ability to capture highly aberrated surfaces [13].

2-2-2 Extended Nijboer-Zernike theory

The limiting requirement of the NZ theory for aberration retrieval is largely mitigated by the Extended Nijboer-Zernike (ENZ) theory as proposed in [12]. Instead of the phase aberration Φ, the GPF is expanded into a Zernike series,

$$\hat{\mathcal{P}}_1(\rho,\theta) = \sum_{n,m} \beta_n^m \mathcal{Z}_n^m(\rho,\theta).$$

Here, the Zernike expansion is used with only cosine (even) terms and the coefficients β are complex-valued but do not have a straight forward physical interpretation as the coefficients α_n^m in equation (2-4). The on-axis intensity of the diffraction image is however, related to the leading coefficient as: $I(0,0) = |\beta_0^0|^2$ [11]. For a special case that $A \approx 1$ and Φ being small,

$$\mathcal{P}(\rho,\theta) \approx 1 + i\Phi(\rho,\theta) = 1 + \sum_{n,m} i\alpha_n^m \mathcal{Z}(\rho,\theta)$$

The authors in [18] derived analytical expressions of the diffraction integral (equation (2-1)) for such a pupil representation. The analytical solution to the minimization in equation (2-3) for the complex coefficients β_n^m is also derived by linearization and solving the PR problem as a linear system of equations. This procedure is not adequate for medium to large aberrations [22] and was improved by considering a predictor-corrector method to overcome the linearization error.

Another basis function comprising both sine and cosine terms in Zernike representation was introduced in [11],

$$\mathcal{N}_n^m(\rho,\theta) = \sqrt{n+1} R_n^{|m|}(\rho)\exp(im\theta).$$

This complex-valued basis function is referred to as ENZ polynomials hereinafter. This representation is physically more suitable as the pupil functions is itself complex-valued. Again, the coefficients β_n^m do not relate directly to the strength of a specific aberration.

To obtain the aberration coefficients α_n^m from the coefficients β_n^m, an additional transformation step, involving phase unwrapping, is required. The retrieval of the coefficients β_n^m and α_n^m using PR algorithms based on convex relaxation and alternating projections is studied in [4] for small aberration correction experiments.

Limitations. Although Zernike polynomials are particularly suitable for the description of phase aberration Φ, it is difficult to derive any physical interpretation from its application to GPF modeling. The ENZ polynomials also share the limitations associated with Zernike polynomial representation to phase aberrations (see Section (2-2-1)).

2-2-3 Gaussian radial basis functions

The RBF interpolation method uses linear combinations of translates of one function $\Psi(\underline{\varrho})$ of real multi-dimensional variable $\underline{\varrho}$. Given a set of centres, $\underline{\varrho}_k$ the RBF interpolate takes the form [23],

$$F(\cdot) = \sum_{k=1}^{N_r} \omega_k \Psi_k(\cdot) = \sum_{k=1}^{N_r} \omega_k \Psi(\left\|\cdot - \underline{\varrho}_k\right\|).$$

Here, ω_k are the weights of the scalar RBF with radial symmetry about its centres $\underline{\varrho}_k$. Many different basis functions $\Psi(\underline{\varrho})$ with global support exist but the scope of this thesis is limited to Gaussian RBF.

$$\Psi_k(\underline{\varrho}) = e^{-\lambda_k \underline{\varrho}^2}$$

The shape parameter $\lambda > 0$ is related to the variance σ^2 of the normal distribution by $\lambda = 1/(2\sigma^2)$. Moreover, it is assumed that the multivariate distribution is radially symmetric i.e., λ_k takes the same value in all directions. The optimal choice of the shape parameter based on a given data is a highly nonlinear problem and therefore, it is assumed that shape parameters for all $k's$ take the same value,

$$\lambda_1 = \cdots = \lambda_{N_r} = \lambda.$$

Unfortunately, severe ill-conditioning may occur in interpolation problems when the shape parameter λ is small corresponding to the flattening of the basis functions [23]. However, this problem is widely studied and it has been shown that limiting interpolants exist and converge to the form of polynomials to obtain a well-conditioned basis [24] (see Appendix B).

Wavefront fitting. The application of radial basis functions for surface description is not new to the optics literature. RBF offers much improved simplicity and geometric flexibility in terms of aperture shapes in exchange of forsaking the orthogonality of the Zernike polynomials. They are simple to implement based upon a summation of a basic function translated across the aperture of the optical element. RBFs provide comparable accuracy to polynomials and spectral convergence can be achieved as opposed to the Zernike polynomials where an optimum number of terms is needed [25]. Another important advantage it offers over the modal methods is the ability to model local deformations and high wavefront slopes at the edge of the pupil [13].

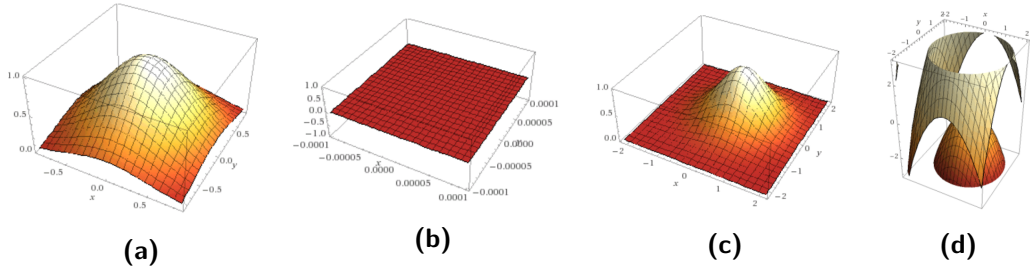
Pupil function modeling. The representation of complex valued pupil function containing both amplitude and phase aberration components using a series of Gaussian RBF has been recently studied [15] to obtain simpler analytical expression for the diffraction integral in equation (2-1). This has been shown to have advantages over the ENZ theory in that, it is not limited to symmetric wavefront errors, exhibits fast convergence for even large values of defocus parameter and is computationally efficient.

The retrieval of the wavefront aberration Φ from intensity data using Gaussian RBF, however, has not been studied in literature. Two different radial basis functions are, therefore, studied as a part of this thesis.

Real-valued RBF

The complex GPF is approximated by a real-valued, radially-symmetric Gaussian RBF on a 2D grid [15],

$$\Psi_k(\rho, \theta; a, b; \lambda) = e^{-\lambda((x-a_k)^2+(y-b_k)^2)}, \quad x = \rho\cos(\theta), \quad y = \rho\sin(\theta).$$



(a) (b) (c) (d)

Figure 2-4: Amplitude and phase plots of 2-4a-2-4b the real-valued RBF and 2-4c-2-4d the complex-valued RBF respectively. The phase profile is zero throughout for the real RBF granting more degree of freedom to the basis. The complex RBF, however, has a Gaussian profile of the phase (upto phase-wrapping).

The amplitude and phase profiles of the RBF are shown in Fig. 2-4a and 2-4b. Using a polar grid for the centres,

$$a_k = \varrho_k \cos(\vartheta_k), \quad b_k = \varrho_k \sin(\vartheta_k),$$

the GPF can be approximated as,

$$\hat{\mathcal{P}}_2(\rho,\theta) = A(\rho,\theta) \sum_{k=1}^{N_r} \gamma_k e^{-\lambda(\rho^2+\varrho_k^2-2\rho\varrho_k\cos(\theta-\vartheta_k))}.$$

Here, $A(\rho,\theta)$ is the indicator function as given in equation (2-2).

Complex-valued RBF

As the GPF is essentially a complex valued function, it is physically more suitable to use complex-valued basis functions for its approximation. This argumentation is based on the same lines with the use of complex-valued Zernike polynomials. The application of complex-valued kernels for image classification problems was recently studied [26]. It was suggested that a complex-valued kernel serves as a regularized model due its sensitivity to phase structure. It was shown that a complex-valued CNN achieved comparable results as its real-valued counterpart and was significantly less vulnerable to over-fitting. The complex Gaussian RBF for approximating the GPF is constructed as a generalization of the real Gaussian RBF by including a phase term,

$$\Psi_k^C(\rho,\theta;a,b;\lambda_1,\lambda_2) = e^{-\lambda_1((x-a_k)^2+(y-b_k)^2)} \cdot e^{-i\pi\lambda_2((x-a_k)^2+(y-b_k)^2)}.$$

Here, the real-valued input features ($\left\|\underline{\varrho}\right\|$, squared distance on a Euclidean space) are phase encoded in $[0,\pi]$ using the transformation $\exp(i\pi\left\|\underline{\varrho}\right\|)$ to obtain the complex valued input features. This is similar to the activation function used in complex-valued neural networks based on real-valued domains [27].

The amplitude and phase profiles of the basis function is shown in Fig. 2-4c and 2-4d. It is noted here that the basis has a Gaussian phase profile as opposed to the flat null-phase of real-valued RBF.

Under polar transformation, the GPF can be approximated as,

$$\begin{aligned}\hat{\mathcal{P}}_3^{(}\rho,\theta) &= A(\rho,\theta)\sum_{k=1}^{N_c}\mu_k e^{-\lambda_1(\rho^2+\varrho_k^2-2\rho\varrho_k\cos(\theta-\vartheta_k))}\cdot e^{-i\pi\lambda_2(\rho^2+\varrho_k^2-2\rho\varrho_k\cos(\theta-\vartheta_k))}\\ &= A(\rho,\theta)\sum_{k=1}^{N_c}\mu_k e^{-(\lambda_1+i\pi\lambda_2)(\rho^2+\varrho_k^2-2\rho q_k\cos(\theta-\vartheta_k))}\end{aligned}$$

It is also noted here that the complex-valued RBF has one extra parameter (λ_2) as compared to its real-valued counterpart.

2-3 Problem formulation

The RBF representation of wavefront offers advantage over orthogonal polynomials such as Zernike modes due to its simplistic implementation, local approximation ability and adaptation to any pupil aperture. This however involves an increased degree of freedom in the design process. Particularly, the optimal choice of shape parameter and the grid for placement of the RBF nodes is an important aspect for the efficient use of this representation. Moreover, it is required to study and rectify the numerical stability issues associated with RBF approximation. To complicate the issue, the wavefront aberration is not known beforehand but has to be estimated from the PSF data observed at the image plane. This restricts the use of the optimization tools available in literature to tune the parameters of the basis functions.

It is desirable to study aberration retrieval using RBF representation of the wavefront and to find avenues to provide a performance better or at least comparable to Zernike representations. As mentioned earlier, the existing research resorts to multiparameter local optimization to design optimal imaging models. These methods are not studied as a part of this thesis but an empirical approach for optimal RBF approximation is designed.

Chapter 3

Radial basis functions for aberration retrieval

It is desired to publish the results of this thesis in the Journal of Optical Society of America A. Therefore, this chapter is written in a research article two-column format.

Aberration retrieval using radial basis functions

A. GUPTA*[1], P. J. PISCAER[1], AND M. VERHAGEN[1]

[1] Delft Centre for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD, The Netherlands
* Corresponding author: a.gupta-9@student.tudelft.nl

Compiled November 23, 2017

The use of radial basis function (RBF) expansion as a means to represent the generalized pupil function for aberration retrieval from intensity point-spread function has been investigated in this research article. The optimal choice of RBF hyper-parameters is derived empirically to achieve an increased accuracy of approximation along with well-conditioned basis. The phase retrieval problem is solved using PhaseLift, an algorithm based on matrix rank minimization and also using a variant of the alternating projections algorithm proposed by Gerchberg-Saxton. The performance of the RBF-based method is compared in terms of accuracy and execution time with that based on the extended Nijboer-Zernike approach. Numerical results based on simulations are presented.

OCIS codes: (100.5070) Phase retrieval; (110.1080) Active or adaptive optics; (010.7350) Wave-front sensing; (220.1000) Aberration compensation; (000.3860) Mathematical methods in physics; (000.4430) Numerical approximation and analysis

1. INTRODUCTION

The inverse problem of aberration retrieval, an important part of any *adaptive optics* system, is a special case of the more general *phase retrieval* (PR) problem. Algorithmic phase retrieval based optical wavefront reconstruction offers an attractive means of estimating the aberration from a set of measurements of the point-spread functions (PSF) of some optical systems [1, 2] due to its experimental simplicity [3]. The representation of the generalized pupil function (GPF) as a linear combination of some convenient basis functions, as opposed to a pixel basis, allows a reduction in computational burden. The complex-valued Zernike polynomials introduced as a consequence of the Extended Nijboer-Zernike (ENZ) theory [4–6] constitute the most widely used basis functions. These basis functions (referred to as ENZ polynomials hereinafter) also allow semi-analytic evaluation of the diffraction integral thereby improving accuracy and efficiency. This was subsequently applied in the field of high resolution optical lithography [7] and the estimation of optical path aberration in very large telescopes [8].

The major limitation of this polynomial representation is that each term extends its influence over the entire pupil and hence is inflexible to be used with arbitrary pupil geometry [9]. Moreover, it is in general difficult to assess the number of terms necessary to achieve acceptable performance a priori. Limiting Zernike analysis to only a few orders may cause incorrect assessment of the severity of the aberration [9]. Although the smoothing effect of the polynomials allows its application to raw data, it significantly limits its ability to capture highly aberrated wavefronts [10]. They are also not suitable for high frequency phase data even if very high-order polynomial terms are used [11].

Recently, GPF approximation based on Gaussian radial basis functions (RBF) was used for semi-analytic evaluation of the diffraction integral [12]. An improvement in terms of complexity, accuracy and execution time was achieved in comparison to the the one based on ENZ theory. This was possible in part due to the parallel evaluation of the integral for a vector of defocus parameters and in part due to a better fit of the complex wavefront achieved using RBFs. RBF offers much improved simplicity and geometric flexibility in terms of aperture shapes in exchange of forsaking the orthogonality of the Zernike polynomials. They provide comparable accuracy to polynomials and spectral convergence can be achieved as opposed to the Zernike polynomials where an optimum number of terms is needed [13]. Another important advantage RBF reconstruction offers over the modal methods is the ability to model local deformations and high wavefront slopes at the edge of the pupil [10].

This article is concerned with the application of Gaussian RBFs to GPF modeling for subsequent aberration retrieval from intensity PSF distribution in the focal region. In addition to a real-valued Gaussian RBF, a complex-valued phase encoded Gaussian function is also proposed. It is asserted that a complex kernel is more suitable for the approximation of a complex-valued GPF and serves as a regularized model. This is based on the argumentation presented in [14] where a complex-valued kernel is used in a convolutional neural network (CNN) framework for real-valued classification. For each of these two RBFs, optimal choice of shape parameters and placement of RBF nodes on a grid are investigated for efficient representation and an empirical design is proposed. An important aspect of RBF-based approximations is to deal with the ill-conditioned system by

using additional means such as Tikhonov regularization [10]. The interplay between the numerical ill-conditioning and the accuracy of the solution to PR problems is also important in practical implementation [1]. This issue is addressed using the RBF-QR algorithm [15].

This RBF representation of the GPF is coupled with PR algorithms based on convex optimization (PhaseLift) [16] and alternating projections (AP) [17] to study aberration retrieval. This approach is similar to the one used in [18] where ENZ polynomials are used instead. The phase aberration data is generated using a Zernike polynomial basis with its coefficients sampled from an assumed distribution. The intensity PSFs corresponding to these phase aberrations are simulated for three values of the defocus parameter. This is done to introduce "phase diversity" in the data to improve the convergence of the AP algorithm [2] which, in general, does not offer any guarantee on uniqueness of the result [19]. Three basis functions, namely real RBF, complex RBF and ENZ polynomials, are compared on the basis of accuracy and efficiency of aberration retrieval from PSF measurements using PhaseLift and AP algorithms.

The structure of the article is as follows. The mathematical formulation of aberration retrieval as a PR problem is presented in Section 2. An overview of the different basis functions used to approximate the GPF is also presented in this section. The PR algorithms used are explained in Section 3. Tuning of RBF parameters for optimal representation of the GPF is explained in Section 4. The simulation results for aberration retrieval are reported in Section 5. A comparative assessment of the performance of different basis functions and the two PR algorithms is performed in Section 6. The conclusions are drawn in Section 7.

2. FORMULATION OF THE PHASE RETRIEVAL PROBLEM

A mathematical formulation of the aberration retrieval problem is briefly presented here. The effects of aberration on an optical system can be modelled using the generalized pupil function (GPF) in a complex phasor form [1] in normalized cylindrical coordinates (ρ,θ) on the exit pupil sphere as

$$\mathcal{P}(\rho,\theta) = A(\rho,\theta)\exp(i\Phi(\rho,\theta)), \tag{1}$$

where $A(\cdot)$ and $\Phi(\cdot)$ are real-valued functions and denote the deviation in amplitude and phase respectively. Under the assumption of purely phase aberrated systems with circular exit pupils, the throughput $A(\rho,\theta)$ is modelled as an indicator function of the interval $[0,1]$ such that its value is unity inside the pupil and zero otherwise. The field in the focal region is related to that in the exit pupil by the Debye diffraction integral as

$$\begin{aligned} U(r,\phi,f;\rho,\theta) =& \frac{1}{\pi}\int_0^1\int_0^{2\pi}\exp(if\rho^2)\mathcal{P}(\rho,\theta) \\ &\times\exp(i2\pi r\rho cos(\theta-\phi))\rho d\rho d\theta, \end{aligned} \tag{2}$$

where (r,ϕ) are the cylindrical coordinates in the focal region and f is the defocus parameter, all normalized with respect to the axial diffraction unit (λ/NA), NA being the image-side numerical aperture of the optical system. Here, $U(r,\phi,f)$ is the complex point-spread function corresponding the GPF. It is measured only in terms of the energy density or the energy flow and is a nonlinear function of the electric field in the focal region,

$$I(r,\phi,f) = |U(r,\phi,f)|^2. \tag{3}$$

Formulation of the phase retrieval problem requires a convenient and systematic parametrization of the GPF, the easiest being pixellation of the pupil. This serves as the most flexible basis as it can be used with any pupil shape [1]. However, the pixel-basis requires a large number of parameters to be identified using PR. As pointed out in [4], parametrizations based on approximating the GPF as a linear superposition of some basis functions reduce the size of the problem. The phase of the GPF can then be retrieved by estimating the coefficients of the basis.

A. Extended Nijboer-Zernike theory

The representation of phase aberration Φ in terms of Zernike polynomials [20] was generalized to represent the GPF under the Extended Nijboer-Zernike (ENZ) theory [6, 21–23]. The GPF is approximated as a truncated series of ENZ polynomials [6],

$$\hat{\mathcal{P}}_1(\rho,\theta) = \sum_{n,m}\beta_n^m\mathcal{N}_n^m(\rho,\theta). \tag{4}$$

Here, n and m denote respectively the radial order and azimuthal frequency of the ENZ polynomial $\mathcal{N}_n^m(\rho,\theta)$ (see Appendix A). The polynomials are ordered according to Noll [24] and the coefficients can then be collected into a single vector $\boldsymbol{\beta}\in\mathbb{C}^{N_\beta}$ where $N_\beta = (n_\mathrm{M}+1)(n_\mathrm{M}+2)/2$, n_M being the maximum radial order considered. The coefficients β_n^m are complex-valued but do not share the physical interpretation pertinent to the coefficients of real-valued Zernike polynomials used to represent Φ [6].

B. Radial basis functions

The pupil function can be approximated by a linear combination of Gaussian radial basis functions. The approximation theory using RBF has been studied widely in literature [25]. Two different strictly positive Gaussian RBFs with local support have been studied in this article. The coefficients are assumed to be complex-valued.

B.1. Real valued RBF

The complex GPF is approximated by a real-valued, radially-symmetric Gaussian RBF [12],

$$\hat{\mathcal{P}}_2(\rho,\theta) = A(\rho,\theta)\sum_{k=1}^{N_\mathrm{r}}\gamma_k\Psi(\rho,\theta;\varrho_k,\vartheta_k;\lambda) \tag{5}$$

$$\Psi_k(\rho,\theta;\lambda) = e^{-\lambda^2(\rho^2+\varrho_k^2-2\rho\varrho_k\cos(\theta-\vartheta_k))}, \tag{6}$$

where (ϱ_k,ϑ_k) are the polar coordinates of the RBF nodes on a polar grid and $A(\rho,\theta)$ is the same as in Eq. (1). Also, $\lambda>0$ is the shape parameter inversely proportional to the square of variance of the normal distribution. The amplitude and phase profiles of the RBF are shown in Fig. 1a and Fig. 1b respectively.

B.2. Complex valued RBF

The complex Gaussian RBF (CRBF) for approximating the GPF is constructed as a generalization of the real Gaussian RBF by including a phase term,

$$\hat{\mathcal{P}}_3(\rho,\theta) = A(\rho,\theta)\sum_{k=1}^{N_\mathrm{c}}\gamma_k^\mathrm{C}\Psi^\mathrm{C}(\rho,\theta;\varrho_k,\vartheta_k;\bar{\lambda}) \tag{7}$$

$$\Psi_k^\mathrm{C}(\rho,\theta;\bar{\lambda}) = e^{-(\bar{\lambda}+i\pi)(\rho^2+\varrho_k^2-2\rho\varrho_k\cos(\theta-\vartheta_k))}. \tag{8}$$

The real-valued input features ($||\cdot||$, squared distance on a Euclidean space) are phase encoded in $[0,\pi]$ using the transformation $\exp(i\pi||\cdot||)$ to obtain the complex valued input features. This is similar to the activation function used in complex-valued neural networks based on real-valued domain [26]. The amplitude and phase profiles of the basis are shown in Fig. 1c and Fig. 1d respectively.

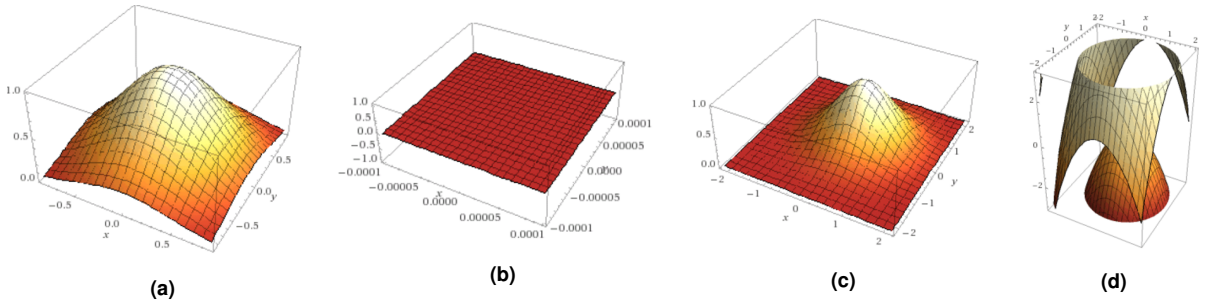


Fig. 1. Amplitude and phase plots of 1a-1b the real-valued RBF and 1c-1d the complex-valued RBF respectively. The phase profile is zero throughout for the real RBF granting more degree of freedom to the basis. The complex RBF, however, has a Gaussian profile of the phase (upto phase-wrapping).

C. Inverse problem

Once a choice of basis is made, the PR formulation can be done. The intensity PSF can be measured using a CCD detector with $N_\mathrm{p} = N^2$ pixels at N_f different positions along the optical axis [18]. The measured pixel values at (r_j, ϕ_j) and defocus f_{dj} are collected into a vector $\mathbf{I} \in \mathbb{R}^{N_\mathrm{m}}$ where $N_\mathrm{m} = N_\mathrm{p} \cdot N_\mathrm{f}$ sorted using a single index j. Acquiring several images along the optical axis improves stability of the inversion process by incorporating extra information in the intensity measurements [1]. Here this phase diversity is added to the system by introducing some known defocus aberration, f_d, $d = 1 \ldots N_\mathrm{f}$. The error between the recorded measurements and the predicted PSF for each acquisition can be written as,

$$\epsilon_j = I_j - |\hat{U}(r_j, \phi_j, f_j, \boldsymbol{\gamma})|^2. \tag{9}$$

The predicted PSF is a linear combination of transformed basis functions weighted by the same coefficients as for the GPF owing to the linearity property of the diffraction integral in Eq. (2). This transformation is performed using two-dimensional FFT denoted by $\mathcal{F}_d\{\cdot\}$ for a defocus f_d,

$$\hat{U}(r, \phi, f_d, \boldsymbol{\gamma}) = \mathcal{F}_d\{\hat{P}_2\} = \sum_{k=1}^{N_\mathrm{r}} \gamma_k \mathcal{F}_d\{\Psi_k(\rho, \theta; \lambda)\}. \tag{10}$$

The norm of the error in Eq. (9) can be minimized for the solution of the PR problem,

$$\underset{\boldsymbol{\gamma} \in \mathbb{C}^{N_r}}{\text{minimize}} \quad ||\boldsymbol{\epsilon}||. \tag{11}$$

The first application of ENZ to aberration retrieval considered only cosine (even) terms of the Zernike polynomials with complex coefficients [4, 5]. The solution of the PR problem in Eq. (11) was derived by linearizing Eq. (9), solved as a system of linear equations, and was adequate for only small aberrations. A predictor-corrector method was later proposed for medium to large aberrations to rectify the error due to linearization [6]. Recently, ENZ polynomials were coupled with algorithms based on convex optimization and alternating projection to solve the problem without any explicit linearization [18]. These algorithms are also used in this article to solve the PR problem.

3. ALGORITHMIC PHASE RETRIEVAL

This section briefly explains two phase retrieval algorithms widely used for PR, one based on convex relaxation of Eq. (9) and another based on non-linear iterative minimization. Comprehensive overview and comparisons can be found in [3].

A. Convex relaxation: PhaseLift

A smooth cost function in Eq. (11) can be written as

$$\underset{\boldsymbol{\gamma} \in \mathbb{C}^{N_r}}{\text{minimize}} \left(\sum_{j=1}^{N_\mathrm{m}} \left(I_j - |\langle \boldsymbol{\gamma}, \mathbf{a}_j \rangle|^2 \right)^2 \right)^{1/2}, \tag{12}$$

where the notation $\langle \boldsymbol{\gamma}, \mathbf{a}_j \rangle = \mathbf{a}_j^H \boldsymbol{\gamma}$ denotes an inner product and complex vectors $\mathbf{a}_j \in \mathbb{C}^{N_\mathrm{r}}$ are the conjugate transposed, $\cdot^H$, rows of the Fourier transformed basis,

$$\mathbf{a}_j = \begin{bmatrix} \vdots \\ \mathcal{F}_d\{\Psi(\rho_j, \theta_j; \lambda)\}^H \\ \vdots \end{bmatrix}. \tag{13}$$

Then the set of quadratic equations in Eq. (12) can be written as linear equations in a higher dimension by considering the linear transformation $|\langle \boldsymbol{\gamma}, \mathbf{a}_j \rangle|^2 = \mathrm{tr}(A_j B)$ such that $A_j = a_j a_j^H$ and $B = \boldsymbol{\gamma}\boldsymbol{\gamma}^H$ and $\mathrm{tr}(\cdot)$ denotes trace of a matrix. Here, the second constraint is such that the matrix B is required to have rank 1. The minimization problem in Eq. (12) can be recast as a matrix recovery problem [27],

$$\begin{aligned} \text{minimize} \quad & ||\boldsymbol{\epsilon}|| + \mu_\mathrm{r} \mathrm{rank}(B) \\ \text{subject to} \quad & \epsilon_j = I_j - \mathrm{tr}(A_j B) \quad j = 1, \ldots, N_\mathrm{m} \\ & B \succeq 0. \end{aligned} \tag{14}$$

As the rank minimization problem is known to be NP-hard [28], a convex relaxation is considered. This can be written as a trace-minimization problem,

$$\begin{aligned} \text{minimize} \quad & ||\boldsymbol{\epsilon}|| + \mu_\mathrm{r} \mathrm{tr}(B) \\ \text{subject to} \quad & \epsilon_j = I_j - \mathrm{tr}(A_j B) \quad j = 1, \ldots, N_\mathrm{m} \\ & B \succeq 0. \end{aligned} \tag{15}$$

This complex semi-definite program can be solved using convex optimization packages like CVX [29, 30]. Then the solution is factorized as $B = \boldsymbol{\gamma}\boldsymbol{\gamma}^H$ to solve the PR problem.

Uniqueness is guaranteed by oversampling and stability can be achieved if the rows of the basis matrix A have a random Gaussian distribution even in presence of noise [16]. The algorithm also works well on more structured, non-random phase retrieval problems [31]. It is also noted here that this good precision comes at a high computational cost as a higher dimensional problem has to be solved. However, the solution space has a

very low dimension when polynomial or RBF approximation is considered as compared to that of a pixel basis. This algorithm is denoted as RBFPL, CRBFPL and ENZPL for RBF, CRBF and ENZ bases respectively.

B. Alternating projections

Alternating projections constitute the most popular class of phase-retrieval algorithms [32, 33]. A variant of the conventional AP algorithm is described here [17]. The following non-smooth cost function is minimized,

$$\underset{\boldsymbol{\gamma}\in\mathbb{C}^{N_r}}{\text{minimize}} \sum_{j=1}^{N_m} \left(\sqrt{I_j} - |\langle \boldsymbol{\gamma}, \mathbf{a}_j\rangle|\right)^2. \tag{16}$$

The method relies on a two step-scheme: (i) initialization, (ii) iterative minimization of the cost in Eq. (16) using a "local search" alternating projections algorithm. A good initialization procedure for random Gaussian distributed $\mathbf{a}_j$ is described in [17] and has been summarized in Table. Defining a basis matrix $M \in \mathbb{C}^{N_m\times N_r}$ and a measurement vector $\mathbf{b} \in \mathbb{R}^{N_m}$ as

$$M = \begin{bmatrix} \vdots \\ a_j^H \\ \vdots \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \vdots \\ \sqrt{I_j} \\ \vdots \end{bmatrix}. \tag{17}$$

Then the associated phase retrieval problem can be written as

$$\text{reconstruct} \quad \boldsymbol{\gamma} \quad \text{from} \quad \mathbf{b} = |M\boldsymbol{\gamma}|. \tag{18}$$

Assuming that the basis matrix M is independent from $\boldsymbol{\gamma}$ and is injective (true for $N_m \gg N_r$ [28]). Then it is sufficient to recover $z = M\boldsymbol{\gamma}$ in the intersection of following two sets,

1. $z \in \{z' \in \mathbb{C}^{N_m}, |z'| = \mathbf{b}\}$;
2. $z \in \text{Range}(M)$.

The solution to the PR problem is then obtained from z via least squares. The algorithm is summarized in Algorithm 1.

Although the use of non-convex algorithm offers an improvement over convex-relaxation algorithms in terms of minimal computational costs, they are based on gradient descent and enjoy weak form of convexity only in the neighbourhood of the solution [17]. Therefore, it is required to have a good initialization and/or more measurements. This algorithm is denoted as RBFAP, CRBFAP and ENZAP for RBF, CRBF and ENZ bases respectively.

4. OPTIMAL RBF APPROXIMATION OF THE PUPIL FUNCTION

In this section, an optimal design of RBF parameters is described and some empirical rules are postulated under the assumption that the wavefront data is available. The wavefront data is generated randomly using Zernike polynomials with varying Zernike order N_α (see Section 5A) and rms$(\cdot)$ values (see Appendix B).

There are basically three parameters that need to be determined for Gaussian approximation namely, shape parameter λ_k, node position (ϱ_k, ϑ_k) and the weights γ_k. A practical way to obtain an efficient representation is to choose for each index k, values of (ϱ_k, ϑ_k) and λ_k. This is, however, a highly non-linear problem usually solved by cross-validation [12]. Therefore, the shape parameters are assumed to be the same for all indices and

Algorithm 1. Alternating projections algorithm

Input: Observations: $\mathbf{b} = \sqrt{I} \in \mathbb{R}^{N_m}$,
Sampling basis: $M \in \mathbb{C}^{N_m\times N_r}$.

procedure INITIALIZATION($\mathbf{b}$,$\mathbf{a}_j$) ▷ $\mathbf{b}_j = |\mathbf{a}^H\boldsymbol{\gamma}|$.

Set

$$\zeta^2 = N_r \frac{\sum_i \mathbf{b}_j}{\sum_i ||\mathbf{a}_j||}.$$

Set $\boldsymbol{\gamma}_0$ to be leading eigenvector of

$$\mathbf{B} = \frac{1}{N_m}\sum_{j=1}^{N_m} \mathbf{b}_j \mathbf{a}_j \mathbf{a}_j^H.$$

return initial guess normalized as $||\boldsymbol{\gamma}_0|| = \zeta$.

procedure AP($\mathbf{b}, M, \boldsymbol{\gamma}_0$)

Initialize $z_0 = M\boldsymbol{\gamma}_0$, $t = 0$.

while $t <$ maxIter **do** ▷ Stop: maximum iterations.

$z'_t \leftarrow \mathbf{b} \odot \text{phase}(z_t)$ ▷ Projection onto set 1.

$z_{t+1} \leftarrow (MM^\dagger) z'_t$ ▷ Projection onto set 2.

$t \leftarrow t + 1$

return $\boldsymbol{\gamma} = M^\dagger z_t$. ▷ $\cdot^\dagger$: pseudo-inverse.

Output: Coefficient vector $\boldsymbol{\gamma}$

different distribution of the nodes are investigated. The accuracy of the Gaussian function and the stability of corresponding linear system depends on the number of data points and on the shape parameter λ.

A. Stability and numerical conditioning

As mentioned earlier, numerical conditioning of RBF basis matrix is an important issue in RBF approximation. This ill-conditioning increases exponentially as the separation between the RBF nodes is decreased for a fixed λ or as the value of λ is decreased for a fixed separation between the nodes [25]. A standard criterion for measuring the numerical stability of an approximation method is its condition number [25]. Condition number near unity indicates a well-conditioned matrix. The inverse calculations of the basis matrix become more sensitive to small errors for high condition numbers.

Sever ill-conditioning occurs in the flat basis function limit $(\lambda \to 0)$. In this limit, it has been proved that limiting interpolants exist and converge to the form of polynomials [15]. This algorithm dubbed as RBF-QR is used in this literature to yield a well-conditioned basis. However, this algorithm works well only for small values of λ but the conditioning worsens for larger values, in which case the regular basis matrix works better (Fig. 2). Therefore, a hybrid algorithm is desirable.

B. Node distribution

The following regular and quasi-random node distributions have been investigated,

- rectangular grid with equally spaced points in $[-1.1, 1.1]$.
- Halton points based grid generated from quasi-random number sequence [34]. According to literature, an improvement in approximation error is expected with anisotropic placement of nodes when representation of possible sharp features or asymmetric local variations of the surface is desired.

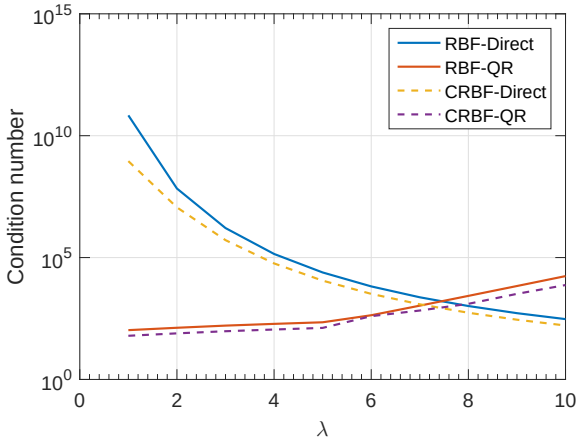


Fig. 2. Matrix condition number for RBF versus the shape parameter for $N_r = 64$ for a Fibonacci grid (as in 1c). The cross-over shape parameter λ_{cross}, for which the RBF-QR and RBF-Direct give the same matrix conditioning, can be used to design a hybrid algorithm.

- Fibonacci grid has been proved to be a competitive and robust choice when the shape parameter is optimally chosen for isotropic RBF representations [10]. A planar Fibonacci grid can be defined using spiral representation in the polar coordinates. For the k^{th} point,

$$\varrho_k = \varrho_0\sqrt{k-1/2}, \quad \vartheta_k = 2\pi k/\varphi, \tag{19}$$

  where ϱ_0 is an arbitrary scale factor and $\varphi = (1+\sqrt{5})/2$ is the Golden ratio. The points on the grid tend to be evenly distributed over the disk.

The grids are defined to have an area slightly larger than the unit disk ($[-1.1, 1.1]$) as it is an efficient method to deal with Runge phenomenon observed in interpolation problems [35]. These node distributions are shown in Fig. 3.

C. Optimal shape parameter

The choice of shape parameter is significant as it affects numerical stability, accuracy of fit and speed of convergence. The practical design of the shape parameter is data dependent, in that it depends on the variance of wavefront aberration and its spatial frequency content as well [13]. As the data of GPF is not available beforehand but has to be estimated using PR techniques, it is desired to find a systematic empirical approach of shape parameter selection.

Two designs are compared here. Firstly, in context of data independent tuning based on mathematical literature [25],

$$\lambda_{adhoc} = 1/(0.815\delta), \quad \delta = (1/N_r)\sum_{k=1}^{N_r}\delta_k. \tag{20}$$

Here, N_r is the total number of nodes and δ_k is the nearest neighbour distance between nodes. This will be referred to as ad-hoc approach hereinafter. Secondly, an algorithm based on *leave-one-out cross validation* (LOOCV) is used [36] to estimate optimal values and the relation in Eq. (20) is updated as

$$\lambda_{opt} = 1/(\zeta\delta). \tag{21}$$

The numerical tuning results are presented in Table 1. The optimal values obtained for real RBF with Fibonacci nodes using the two tuning approaches are shown in Fig. 4a. The variation of the optimal shape parameter with number of terms N_α is shown in Fig. 4b.

Table 1. Conditioning and tuning parameters for radial basis functions for $N_r = 69$. **N** is the type od node distribution. Subscripts: adhoc Eq. (20); opt optimal value calculated using Rippa's LOOCV algorithm; cross: Fig. 2. The adhoc relation is updated as in Eq. (21). R: Rectangular, F: Fibonacci and H: Halton.

Ψ	**N**	λ_{adhoc}	λ_{opt}	λ_{cross}	**cond**(Ψ)	ζ
RBF	R	4.86	3.23±0.11	8.00	$4.0\cdot10^2$	0.580
	F	5.47	3.21±0.08	7.40	$1.5\cdot10^2$	1.333
	H	6.92	4.55±0.10	7.80	$2\cdot10^2$	0.803
CRBF	R	4.86	3.53±0.21	8.10	$3\cdot10^2$	0.786
	F	5.47	3.26±0.04	7.40	$1\cdot10^2$	0.757
	H	6.92	4.57±0.05	7.90	$1.5\cdot10^2$	0.753

5. SIMULATION

In this section, a comparative assessment of performance of the three basis functions, described in Section 2, is carried out to reconstruct static aberrations from intensity PSFs. Firstly, the wavefront aberration data is obtained in terms of Zernike polynomials with the coefficients drawn from an assumed distribution. Then the corresponding PSFs are computed. Lastly, these aberrations are retrieved using the two PR algorithms. The last step is repeated using each of the three different basis functions.

A. Preparation of aberration data

The phase aberration data is simulated using the Zernike representation

$$\Phi(\rho,\theta) = \sum_{n,m}\alpha_n^m\mathcal{Z}_n^m(\rho,\theta), \tag{22}$$

where the Zernike polynomial $\mathcal{Z}_n^m$ is defined in Appendix A. It is assumed that the leading coefficient $\alpha_0^0 = 0$. This piston coefficient does not affect the image quality [18]. The total number of Zernike terms considered is denoted by $N_\alpha = (n_\alpha+1)(n_\alpha+2)/2 - 1$, where n_α is the maximum radial order considered. The coefficient distribution is similar to the one derived experimentally in [18] and is reported in Fig. 5a.

The wavefront data is generated over a circular pupil on a 256×256 Nyquist-sampled grid (Fig. 5b). Twenty sets of coefficient vectors $\boldsymbol{\alpha}$ with $N_\alpha = 104$ terms (corresponding to Zernike polynomials up to the 13^{th} radial order) are sampled from the distribution in Fig 5a as explained in Appendix C. For a single set of coefficients $\boldsymbol{\alpha}$, the phase aberration Φ is generated for each increment in the radial order from $n_\alpha = 5-13$. The final data set, thus, consists of 160 phase aberration maps (8 for each of the 20 coefficient vectors) which is used for further analysis.

A scalar number is associated with each phase map that quantifies the variance of the aberration and is denoted as rms(Φ) (see Appendix B). This scalar quantity is intended to represent the spatial frequency contained in a given Zernike order and by design, increases with increasing radial order for a single set of coefficients $\boldsymbol{\alpha}$.

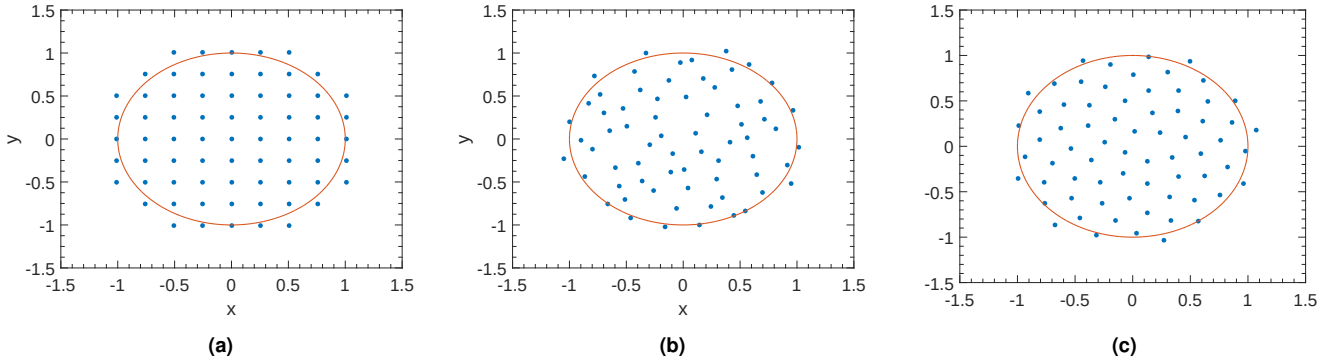


Fig. 3. Node distributions on a 2D grid for a unit disk pupil aperture: 3a Rectangular (R), 3b Halton (H) and 3c Fibonacci (F).

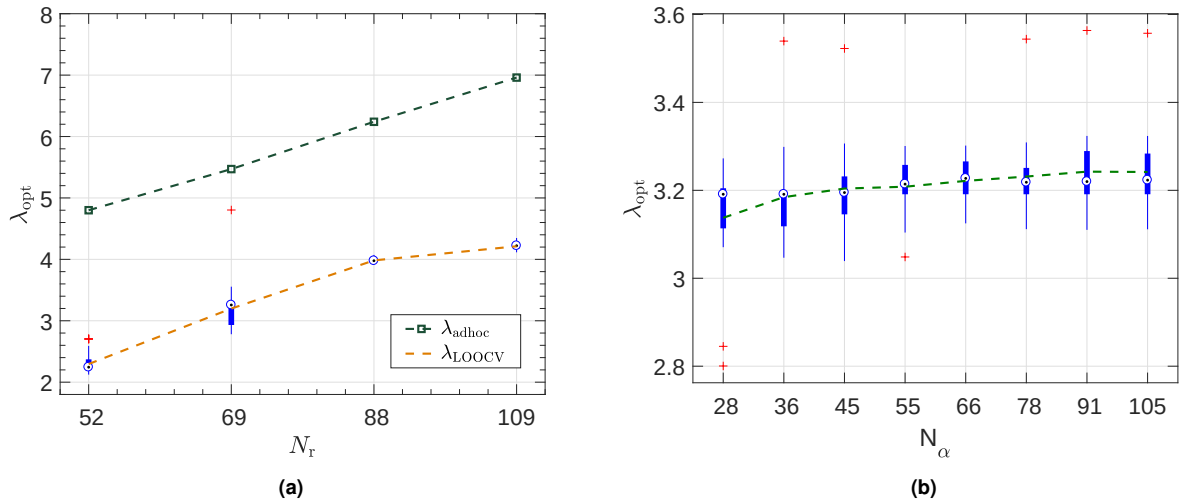


Fig. 4. Tuning of RBF shape parameter. Fig. 4a Optimal shape parameter for real RBF with Fibonacci nodes with varying number of nodes N_{r}. At every value of N_{r}, a boxplot of the data-dependent optimal shape parameter is plotted. A line is then drawn through the mean values of these distributions. The ad-hoc values of the shape parameter are also plotted here (Eq. (20)). Fig. 4b Optimal shape parameter for RBF with Fibonacci nodes versus the number of Zernike polynomial terms used to generate the phase data. The whiskers of the boxplots denote the range of the variables with circumpunct as the median. A line is drawn via the means of optimal shape parameters.

B. Error in approximating the generalized pupil function

The errors in approximating the GPF using the ENZ polynomials and RBF are compared in this section. A purely phase aberrated pupil function is considered for this analysis i.e.,

$$\mathcal{P}(\rho,\theta) = \exp\left(i\sum_{n,m}\alpha_n^m\,\mathcal{Z}_n^m(\rho,\theta)\right). \tag{23}$$

The coefficient vector α_n^m is generated as explained in the previous section. This pupil function is approximated using the three basis functions discussed in Section 2 and is denoted as $\hat{\mathcal{P}}(\rho,\theta)$. For a purely phase aberrated GPF, an approximation of the phase aberration Φ is obtained as

$$\hat{\Phi}(\rho,\theta) = \arg(\hat{\mathcal{P}}(\rho,\theta)). \tag{24}$$

The approximation error, $E = \Phi - \hat{\Phi}$, is quantified using a scalar number, $\text{rms}(E)$, as explained in Appendix B.

Firstly, it is observed in Fig. 6a that the error in approximation decreases monotonously with increase in number of radial orders of ENZ basis. The number of RBF centres used for the approximation are assumed to be the same with $N_{\text{r}} = N_{\text{c}} = 69$. The ENZ polynomials can thus be truncated to the nearest radial order, i.e. $n_{\text{M}} = 13$ with $N_\beta = 66$ terms. The λ_{opt} for RBF is taken to be constant equal to the mean values in Table 1. The $\text{rms}(E)$ is plotted against the number of Zernike polynomial terms used to generate the data, N_α, in Fig 6b. By design, the severity of aberration increases with the number of Zernike terms. Consequently, it is observed that the least square error in approximation of the GPF increases with the increasing order of aberration. Also, this error for the RBF Fibonacci (F) case is always below that of the ENZ approximation indicating a better fit. The rms error for the RBF based on Halton (H) and rectangular (R) nodes show intermediate performance diverging from the Fibonacci case with increase in order of aberration. The performance of the complex-valued RBF was found to be same as its real valued counterpart. Therefore, it has not been reported in Fig. 6b. However, an improvement in matrix conditioning can be observed in Fig. 2 and Table 1.

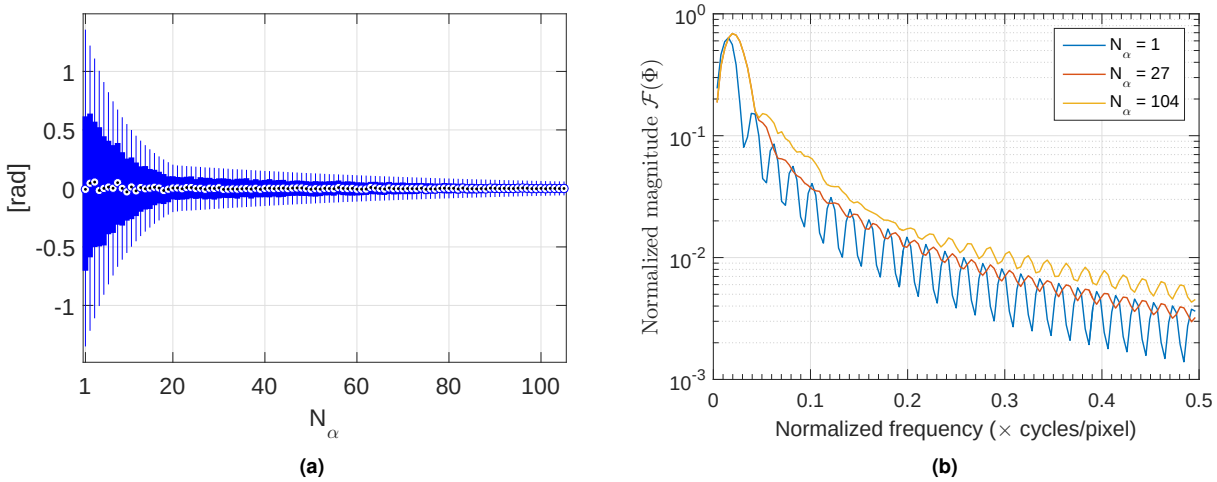


Fig. 5. Wavefront aberration data. Fig. 5a Boxplot of 500 realizations of randomly generated Zernike aberration coefficients α_n^m with an exponential decay in the values with order. On each box, central mark indicates the median, the top and bottom edge of the blue box indicate first and third quartiles, respectively. The whiskers (thin blue lines) extend to the most extreme data points not outliers. 5b Normalized frequency spectrum of the phase screen Φ (Eq. (22)) for different number of coefficients from the distribution in Fig. Fig. 4a plotted against normalized spatial frequency in cycles/pixel. Most of the energy of the signal is confined to 2π rad/4pixel signifying Nyquist limit.

C. Aberration correction simulation

The next step in the assessment of radial basis functions against ENZ polynomials involves aberration correction simulations that estimate the phase aberration Φ from intensity PSF using PR algorithms. The intensity PSF data is generated for the pupil function model, Eq. (23), using Eq. (10) at defocus positions $f_1 = 0.0$, $f_2 = -2.77$, and $f_3 = 3.46$. This data is generated as 256 pixel $\times$ 256 pixels images.

Again, the approximation error is quantified using a single scalar quantity, rms(E), in terms of the reconstruction of the coefficient vector, $\boldsymbol{\alpha}$, used to generate the data. The procedure is summarized in Algorithm 2. The phase unwrapping [37] in step 10 is applicable only when $\arg(\hat{\mathcal{P}}(\rho,\theta))$ extends beyond the interval $[-\pi,\pi]$. The rms error values are normalized with the rms of the phase aberration (Step 12) for comparison across different wavefront datasets.

The aberration correction simulations are carried out for each PR algorithm and each basis function under study. In case of the AP algorithm (see Section 3B), $N_\mathrm{r} = N_\mathrm{c} = 69$ terms are taken for the RBF approximations and the ENZ polynomials are similarly truncated at $N_\beta = 66$. The value of λ_opt is kept constant as the mean values reported in Table 1. A total number of measurements $N_\mathrm{m} = 51040$ are used to estimate the parameters. The stopping criterion is set at 400 iterations or earlier if the successive estimate update is less than $1 \cdot 10^{-7}$. The simulations based on AP algorithm are carried out on an Intel Core i7 quad-core processor (2.50 GHz) with 4GB RAM.

The PhaseLift algorithm (see Section 3A) was found to be slow and therefore, the data is instead sampled on smaller 32 pixel $\times$ 32 pixels images. The number of basis functions is set at $N_\mathrm{r} = N_\mathrm{c} = N_\beta = 45$ and only the Fibonacci case is studied for RBF. The value of $\lambda_\mathrm{opt} = 2.74$ is kept constant and is calculated using the relation in Eq. (21) for $\delta = 0.274$. A total number of measurements $N_\mathrm{m} = 2220$ are used to estimate the parameters. The regularization parameter is fixed at $\mu_\mathrm{r} = 1$ and the convex optimization is solved using the CVX toolbox using the SDPT3 solver (infeasible path-following algorithm).

Algorithm 2. Aberration retrieval

Input: Aberration coefficients $\boldsymbol{\alpha} \in \mathbb{R}^{N_\alpha}$, Fig. 5a,
Basis matrix and its Fourier transform: $\Psi, \mathcal{F}_d\{\Psi\} \in \mathbb{C}^{N_\mathrm{m} \times N_\mathrm{r}}$.

2: **procedure** SIMULATE ABERRATION($\boldsymbol{\alpha}, N_\alpha$)
Generate $\Phi(\rho,\theta)$ using Eq. (22).
4: Obtain GPF $\mathcal{P}(\rho,\theta)$ using Eq. (23).
Obtain PSF $\mathbf{I} \in \mathbb{R}^{N_\mathrm{m}}$, Eq. (2),(3), at three defocus f_d.
6: **procedure** RETRIEVE ABBERATION($\mathbf{I}, \Psi, \mathcal{F}_d\{\Psi\}$)
Estimate $\boldsymbol{\gamma} \leftarrow PR(\mathbf{I}, \Psi, \mathcal{F}_d\{\Psi\})$ ▷ PL/AP (Section 3)
8: $\hat{\mathcal{P}} \leftarrow \sum_{k=1}^{N_\mathrm{r}} \gamma_k \Psi_k$ ▷ Estimated GPF
$\bar{\Phi} \leftarrow \arg(\hat{\mathcal{P}})$ ▷ Phase of the GPF
10: $\hat{\Phi} \leftarrow \text{phase\_unwrap}(\bar{\Phi})$ ▷ Phase unwrappping
$\hat{\boldsymbol{\alpha}} \leftarrow \boldsymbol{\mathcal{Z}} \backslash \hat{\Phi}$ ▷ Least squares solution
12: **return** $\text{RMSE}(\Phi - \hat{\Phi}) \leftarrow \frac{||\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}||}{||\boldsymbol{\alpha}||}$. ▷ Normalized error

Output:

The threshold value is same as that of AP algorithm but the maximum number of iterations is kept at the default value. The simulations based on PhaseLift algorithm are carried out on an Intel Xeon CPU E5-1620 v3 processor (3.5 GHz) with 16GB RAM.

It is noted here that the assumptions underlying the PhaseLift algorithm that guarantee uniqueness and stability of the solution to the PR problem, as mentioned in 3A, could not be verified in the cases presented here. However, a consistent aberration retrieval performance is achieved in simulations. Unfortunately, the AP algorithm only offers a weak form of convexity in the neighbourhood of the solution [17] and the entire measurement needs to be utilized to achieve convergence in simulations.

The results of the aberration retrieval simulations are summarized in Fig 7. It was found that complex valued RBF did not offer any significant advantage over its real-valued counterparts in terms of accuracy. Therefore, the results are not shown in the figure. For the case of PhaseLift, a least squares fit using 136 ENZ polynomials is also plotted in Fig. 7b.

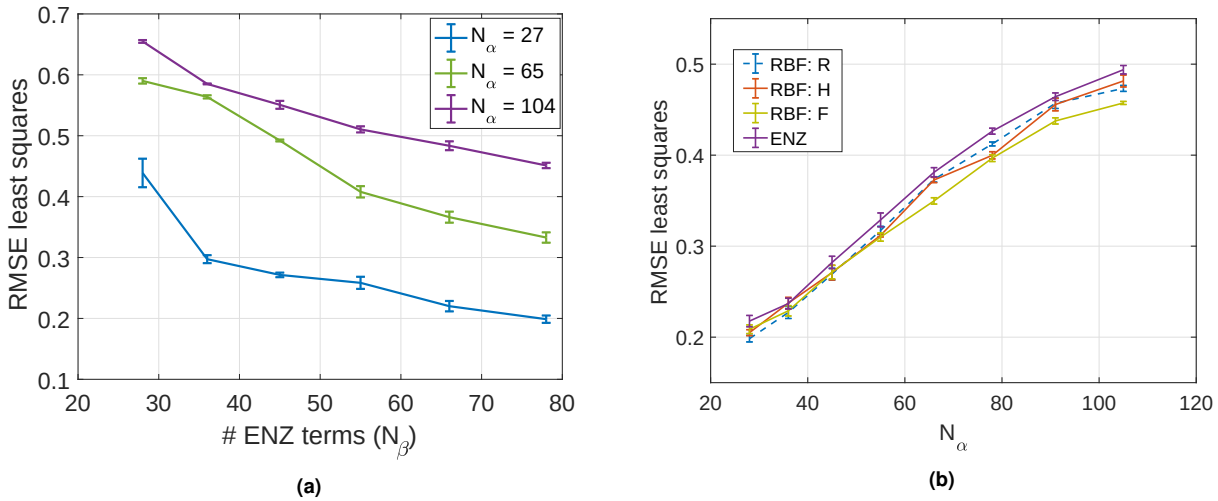


Fig. 6. Least squares approximation of GPF. Fig. 6a RMS least squares error (Algorithm 2 step 12) for ENZ polynomial reconstruction of Zernike aberrations. The error is calculated as in Algorithm 2 step 12. It decreases with increase in number of ENZ terms N_β. The error plots are shown for three different number of Zernike polynomials (N_α) used for generating the data. The error bars denote one standard deviation in the values. Fig. 6b RMS least squares error (Algorithm 2 step 12) for ENZ polynomial ($N_\beta = 66$) and RBF based ($N_r = 69$) reconstruction of Zernike aberrations. The error bars denote one standard deviation in the values. RBF grids are denoted as R: rectangular, H: Halton and F: Fibonacci.

The computational time spent by the alternating projections algorithm and PhaseLift algorithm for different basis functions are found to be the same for different basis functions at mean values of 25 sec and 650 sec respectively. These could not be compared because the simulations for AP and PhaseLift were performed on systems with different configurations.

6. DISCUSSION

The shape parameter of the RBF provides an extra degree of freedom that allows the capture of small details in the wavefront data with an easy implementation. It was observed that the RBF-QR algorithm achieved a better conditioning of the RBF system. This is important because the standard basis is prone to small errors in the region of the optimal shape parameter. However, this algorithm is generally more suited to circular pupils [15]. Other algorithms exist but have more complex implementation and are therefore not investigated. These can be found in [38] with a detailed comparison. It is noted here that RBF-QR algorithm does not provide any significant improvement in the approximation of the data but only improves the conditioning of the system of equations for further calculations.

The guideline proposed for the optimal tuning of λ based on the ad-hoc technique (Section 4C) was found to be very useful when a different number of basis functions is used. However, λ_{opt} also depends on the frequency content of the aberration as is evident from Fig. 4b. The optimal λ increases with the number of terms N_α in the Zernike aberration (Fig. 4b) which, by design, also means an increase in frequency content. As the generated Zernike aberrations are not able to capture a wide spectrum of frequency content (Fig. 5b), the variation in λ_{opt} is very less. Therefore, a mean value was used for aberration correction experiments. A practical value based on bandwidth $\mathbf{f}_\text{b}$ of the aberration is given as $\lambda = (\pi \mathbf{f}_\text{b})^2$ [13]. This relation can be combined with the ad-hoc relation for aberrations of varying frequency content to find a generalized tuning relation for the shape parameter. A similar method was adopted in [11] for estimating the Zernike order based on "targeted" frequency content.

The RBF based on Fibonacci nodes were found to be the most efficient in estimating the GPF. This can be attributed to the even distribution of the nodes over the pupil aperture. The RBF based on Halton and rectangular nodes were found to exhibit similar performances. The simulations for approximation of the GPF showed that the RBF have a better accuracy in comparison to ENZ polynomials. A tuning based on frequency content of the aberration can possibly result in better performance.

The AP aberration correction simulations were found to be in agreement with the GPF approximation results. Moreover, it was observed that the accuracy of RBF based approximation over that based on ENZ polynomials increased (upto 10% for the Fibonacci case) with increase in severity of the aberration even with constant shape parameters. The standard deviation in the percentage retrieval values was low for all the cases because of the way the aberration data is collected. The standard deviation for RBFAP was found to be even smaller than that for ENZAP.

The PhaseLift aberration correction simulations provide an optimal fit that is not obtained using the AP algorithm. Although the percentage retrieval values differ from the AP case due to different sampling of the aberration data, the trend remains the same. The RBF approximation provides a considerably better accuracy for higher order aberrations. Also, the standard deviation in the obtained values were very less.

It is noted that the aberration data collected might not relate to that found in practice. However, Zernike polynomials have been used widely to fit wavefront data in the field of adaptive optics and provide a good basis. Lastly, as ENZ polynomials are also based on Zernike polynomials, the data is believed to be biased towards it. Even then, a considerable improvement in accuracy is obtained with RBF.

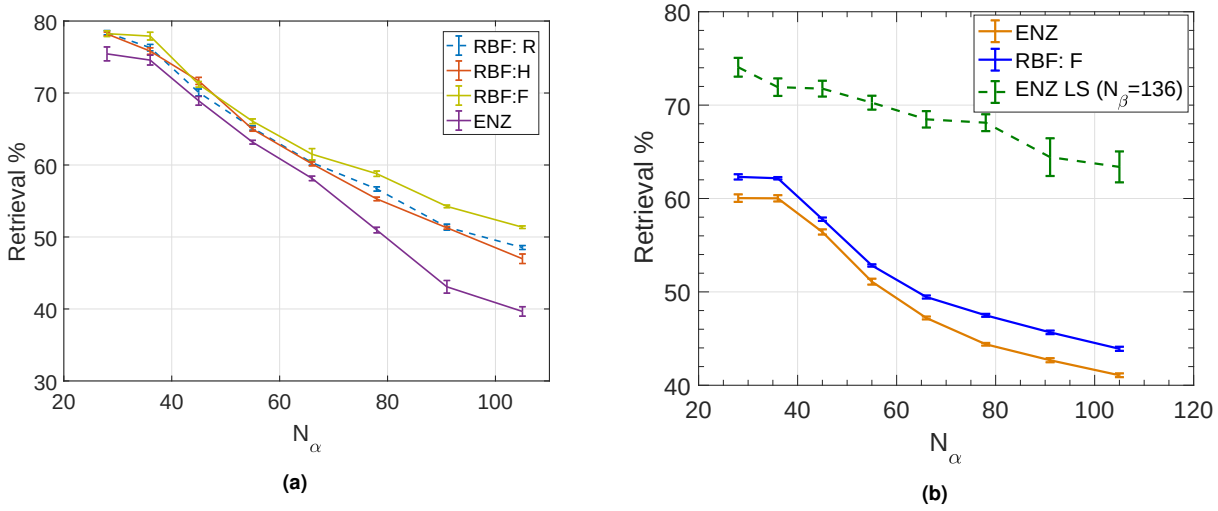


Fig. 7. Summary of 160 aberration correction simulations. Fig. 7a reports the Alternating projections results with an error plot of aberration retrieval performance (Algorithm 2 step 12) with a line drawn through the mean over the different cases and error bar denote one standard deviation in the values. Fig. 7b reports the PhaseLift results for the ENZ and RBF (Fibonacci) cases. A least squares fit using 136 ENZ polynomials is also shown here that provides a reference maximum fit for under-sampled data.

7. CONCLUSION

The inverse problem of aberration retrieval from intensity point spread function is a non-convex optimization problem. It has already been proved that the use of basis functions to represent complex-valued generalized pupil function can result in an improvement in approximation and computational time. The application of radial basis functions to model the GPF has been explored in this article. As the GPF data is not available but has to be estimated by solving an optimization problem, empirical relations for the optimal choice of the shape parameter and grid choice have been derived. The RBF-QR algorithm has been used to improve the numerical conditioning of the RBF-approximation. The complex-valued RBF resulted in a basis that is better conditioned than its real-valued counterpart with similar accuracy.

The RBF based approach proposed here has been compared with extended Nijboer-Zernike based complex-valued polynomials to retrieve the phase aberration. The result of the comparison show that use of RBF is competitive in terms of accuracy especially for higher order aberrations. It has been found that Fibonacci nodes are more efficient as compared to other grid types due to their symmetric distribution over the pupil.

As variation in the frequency content of the aberration data is low, the optimal shape parameter was chosen to be constant for a particular set of nodes. However, a slight monotonous increase in its value was observed with increasing frequency content of the aberration. This observation indicates that the proposed framework to tune the shape parameter can be extended to include target bandwidth of the aberration.

8. APPENDIX

A. REAL AND COMPLEX-VALUED ZERNIKE POLYNOMIALS

The phase aberration Φ can be analyzed by the orthogonal set of basis functions formed by the circle polynomials, Z_n^m, introduced by Zernike [20],

$$\Phi(\rho,\theta) = \sum_{n,m} \alpha_n^m \mathcal{Z}_n^m(\rho,\theta), \tag{25}$$

where indices $n \in \mathbb{N}_0$ and $m \in \mathbb{Z}$ respectively denote the radial order and the azimuthal frequency of the Zernike polynomial $\mathcal{Z}_n^m$ such that $n - |m| > 0$ and even. The polynomials are given by the product of a radial polynomial $R_n^{|m|}(\rho)$ and a trigonometric function $\Theta_n^m(\theta)$ with suitable normalization c_n^m,

$$\mathcal{Z}_n^m(\rho,\theta) = c_n^m R_n^{|m|} \Theta_n^m(\theta). \tag{26}$$

where,

$$c_n^m = \begin{cases} \sqrt{n+1} & m = 0 \\ \sqrt{2(n+1)} & m \neq 0 \end{cases}, \quad \Theta_n^m(\theta) = \begin{cases} cos(m\theta) & m \geq 0 \\ -sin(m\theta) & m < 0 \end{cases},$$
$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} \rho^{n-2s}. \tag{27}$$

The expression of the radial function $R_n^m(\rho)$ can be found in [24, 39]. The GPF can be analyzed using a truncated series of ENZ polynomials [6],

$$\mathcal{N}_n^m(\rho,\theta) = \sqrt{n+1} R_n^{|m|}(\rho) \exp(im\theta). \tag{28}$$

Again, the coefficients can be collected in a single vector using Noll's indexing. The normalization used here is as given in [18].

B. ROOT-MEAN-SQUARE VALUE OF ZERNIKE PHASE ABERRATIONS

Owing to the orthogonality of the Zernike polynomial expansion, the standard deviation of separate aberration terms across the pupil can be represented by the corresponding aberration coefficient, α_n^m. The mean square value of the function $\Phi(\rho,\theta)$ is defined as,

$$E_2[\Phi] = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} \Phi(\rho,\theta)^2 \rho d\rho d\theta. \tag{29}$$

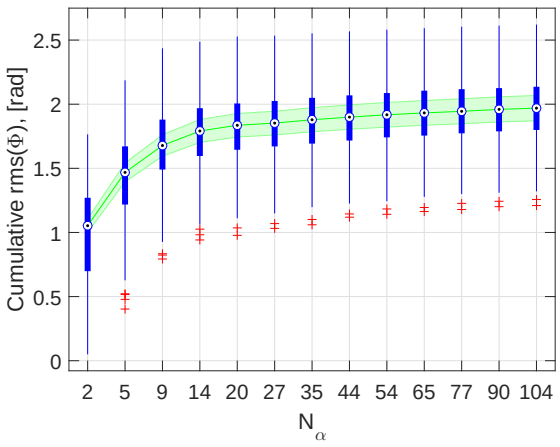


Fig. 8. The cumulative variance of the aberration versus the number of Zernike polynomial terms considered for the distribution of Zernike coefficients α_n^m in Fig. 5a. The green shaded region denotes 5% of the median values.

The aberration variance is defined as,

$$\text{rms}(\Phi) = (E_2[\Phi] - E_1[\Phi])^{1/2}. \quad \textbf{(30)}$$

The mean value of the aberration, $E_1[\Phi]$, is given by the norm of the first Zernike aberration coefficient α_0^0. This piston term has been assumed to be zero in this article as it does not affect the image quality. Therefore, for the choice of normalization c_n^m in Eq. (27), $\text{rms}(\Phi) = \|\boldsymbol{\alpha}\|$ [40]. The error in reconstruction can also be expressed in terms of the norm of the coefficient vector as $\text{rms}(\Phi - \tilde{\Phi}) = \|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|$.

C. SAMPLING ZERNIKE ABERRATION COEFFICIENTS

The Zernike coefficients are sampled from the distribution shown in Fig. 5a. The variability in the variance of the aberration, $\text{rms}(\Phi) = \|\boldsymbol{\alpha}\|$, was found to be very high for a particular number of polynomial terms, N_α, considered. This is expected due to the randomness in the data. To reduce this randomness, 20 realizations of $\boldsymbol{\alpha}$ were selected such that the cumulative $\text{rms}(\Phi)$ was within 5% of the median as shown in Fig. 8.

REFERENCES

1. D. R. Luke, J. V. Burke, and R. G. Lyon, "Optical wavefront reconstruction: Theory and numerical methods," SIAM review **44**, 169–224 (2002).
2. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Optical Engineering **21**, 829–832 (1982).
3. Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," IEEE signal processing magazine **32**, 87–109 (2015).
4. P. Dirksen, J. Braat, A. J. Janssen, and C. Juffermans, "Aberration retrieval using the extended nijboer-zernike approach," Journal of Micro/Nanolithography, MEMS, and MOEMS **2**, 61–68 (2003).
5. C. Van der Avoort*, J. Braat, P. Dirksen, and A. Janssen, "Aberration retrieval from the intensity point-spread function in the focal region using the extended nijboer–zernike approach," Journal of Modern Optics **52**, 1695–1728 (2005).
6. S. Van Haver, "The extended nijboer-zernike diffraction theory and its applications," Ph.D. thesis, Delft University of Technology (2010).
7. P. Dirksen, J. Braat, and A. J. Janssen, "Estimating resist parameters in optical lithography using the extended nijboer-zernike theory," Journal of Micro/Nanolithography, MEMS, and MOEMS **5**, 013005–013005 (2006).
8. P. Riaud, D. Mawet, and A. Magette, "Nijboer-zernike phase retrieval for high contrast imaging-principle, on-sky demonstration with naco, and perspectives in vector vortex coronagraphy," Astronomy & Astrophysics **545**, A150 (2012).
9. A. Martínez-Finkelshtein, A. M. Delgado, G. M. Castro, A. Zarzo, and J. L. Alió, "Comparative analysis of some modal reconstruction methods of the shape of the cornea from corneal elevation data," Investigative ophthalmology & visual science **50**, 5639–5645 (2009).
10. M. Maksimovic, "Optical design and tolerancing of freeform surfaces using anisotropic radial basis functions," Optical Engineering **55**, 071203–071203 (2016).
11. B. H. Dean and C. W. Bowers, "Diversity selection for phase-diverse phase retrieval," JOSA A **20**, 1490–1504 (2003).
12. A. Martinez-Finkelshtein, D. Ramos-Lopez, and D. Iskander, "Computation of 2d fourier transforms and diffraction integrals using gaussian radial basis functions," Applied and Computational Harmonic Analysis **43**, 424–448 (2017).
13. M. Montoya-Hernández, M. Servín, D. Malacara-Hernández, and G. Paez, "Wavefront fitting using gaussian functions," Optics Communications **163**, 259–269 (1999).
14. N. Guberman, "On complex valued convolutional neural networks," arXiv preprint arXiv:1602.09046 (2016).
15. B. Fornberg, E. Larsson, and N. Flyer, "Stable computations with gaussian radial basis functions," SIAM Journal on Scientific Computing **33**, 869–892 (2011).
16. J. Emmanuel, "Candes, thomas strohmer, and vladislav voroninski," Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. CoRR, abs/1109.4499 (2011).
17. I. Waldspurger, "Phase retrieval with random gaussian sensing vectors by alternating projections," arXiv preprint arXiv:1609.03088 (2016).
18. J. Antonello and M. Verhaegen, "Modal-based phase retrieval for adaptive optics," JOSA A **32**, 1160–1170 (2015).
19. H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization," JOSA A **19**, 1334–1345 (2002).
20. Z. von F, "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode," Physica **1**, 689–704 (1934).
21. A. J. Janssen, "Extended nijboer–zernike approach for the computation of optical point-spread functions," JOSA A **19**, 849–857 (2002).
22. J. Braat, P. Dirksen, and A. J. Janssen, "Assessment of an extended nijboer–zernike approach for the computation of optical point-spread functions," JOSA A **19**, 858–870 (2002).
23. J. J. Braat, S. van Haver, A. J. Janssen, and P. Dirksen, "Assessment of optical systems by means of point-spread functions," Progress in Optics **51**, 349–468 (2008).
24. R. J. Noll, "Zernike polynomials and atmospheric turbulence," JOsA **66**, 207–211 (1976).
25. G. E. Fasshauer, *Meshfree Approximation Methods with Matlab:(With CD-ROM)*, vol. 6 (World Scientific Publishing Co Inc, 2007).
26. M. F. Amin and K. Murase, "Single-layered complex-valued neural network for real-valued classification problems," Neurocomputing **72**, 945–955 (2009).
27. E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," Communications on Pure and Applied Mathematics **66**, 1241–1274 (2013).
28. E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," SIAM review **57**, 225–251 (2015).
29. M. Grant, S. Boyd, and Y. Ye, "Cvx toolbox," (2009).
30. M. X. Goemans and D. Williamson, "Approximation algorithms for max-3-cut and other problems via complex semidefinite programming," in "Proceedings of the thirty-third annual ACM symposium on Theory of computing," (ACM, 2001), pp. 443–452.

31. I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," Mathematical Programming **149**, 47–81 (2015).
32. R. W. Gerchberg, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," Optik **35**, 237–246 (1972).
33. J. R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," Optics letters **3**, 27–29 (1978).
34. L. Kocis and W. J. Whiten, "Computational investigations of low-discrepancy sequences," ACM Transactions on Mathematical Software (TOMS) **23**, 266–294 (1997).
35. J. P. Boyd, "Six strategies for defeating the runge phenomenon in gaussian radial basis functions on a finite interval," Computers & Mathematics with Applications **60**, 3108–3122 (2010).
36. M. Mongillo, "Choosing basis functions and shape parameters for radial basis function methods," SIAM Undergraduate Research Online **4**, 190–209 (2011).
37. B. M. Hanser, M. G. Gustafsson, D. A. Agard, and J. W. Sedat, "Phase retrieval for high-numerical-aperture optical systems," Optics letters **28**, 801–803 (2003).
38. S. A. Sarra and S. Cogar, "An examination of evaluation algorithms for the rbf method," Engineering Analysis with Boundary Elements **75**, 36–45 (2017).
39. M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).
40. V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," Applied optics **33**, 8121–8124 (1994).

Chapter 4

Recommendations for future work

A general framework for the design of radial basis functions (RBF) has been proposed in this thesis. It was mentioned that a large variation in frequency content in the phase aberration could not be simulated using Zernike polynomials. This limited the tuning of shape parameter to only a specific case of low to mid spatial frequency content. Therefore it is required to generalize the empirically obtained tuning relation to a wider class of aberrations. A method to achieve this objective has also been proposed in the Discussion section of Chapter 3.

It is again emphasized that the Zernike aberrations produced data that was possibly biased towards approximation using extended Nijboer-Zernike (ENZ) polynomials. Therefore, it is required to test the proposed basis functions on a more general class of aberrations found in the domain of adaptive optics. One possibility is to generate atmospheric turbulence data based on Kolmogorov phase screens [28].

The most cited advantage of RBF approximation is its relative robustness of the computation with respect to the underlying geometry of the pupil. The simulations presented in this thesis are all based on unit circular disk pupils. Therefore, this local character of RBF could not be brought to light. The pupil shape usually varies across application domain. For example, for the case of Hubble telescope, the pupil has a hexagonal annular geometry [5]. Moreover, phase jumps across separate panels makes it difficult to carry out Gram-Schmidt orthogonalization of the Zernike polynomials. On the other hand, RBF can handle such situations with ease.

It was also noted in the report that the RBF approximation worked better for higher order aberrations for the same number of terms as Zernike polynomials. The limitation of Zernike polynomials to capture aberrations of high frequency content has demotivated its use Extreme Adaptive Optics systems (ExAO) with application to high contrast imaging [29, 30, 31]. This creates interesting avenues for the use of RBF in aberration retrieval for adaptive optics systems that can capture high spatial frequency with a reduced computational burden.

Finally, the use of semi-analytic formulation of the fast Fourier transform based on RBF as described in [15] can be extended to formulate the aberration retrieval problem as done using ENZ theory in [4]. This can result in a better approximation and a decrease in computational time.

Appendix A

Phase retrieval

This chapter sheds light on some of the existing algorithmic phase retrieval techniques. This reconstruction of a complex vector from its linearly transformed magnitude is a non-convex optimization problem.

The typical arrangement of the phase-retrieval problem is discussed as follows [32].

For a discrete signal $\mathbf{x} \in \mathbb{C}^n$, the measurements are the squared modulus of the inner product between the signal and some measurement vectors a_μ,

$$b_\mu^2 = |\langle \mathbf{x}, a_\mu \rangle|^2, \quad \mu = 1, 2, \ldots, M.$$

Here M are the number of measurements. The phase information inside the modulus squared notation is lost. Therefore, it is intended to recover x from the data matrix $\mathbf{b} = [b_1, b_2, \ldots b_M]$.

$$\langle \mathbf{x}, a_\mu \rangle = \sigma_\mu b_\mu.$$

For a real-valued signal $\mathbf{x}$, σ_μ's are just signs having 2^M combinations. For the complex-valued case, σ_μ's are also complex valued with each on a unit circle making the optimization problem more complicated [32].

There are several methods which impose conditions on M and a_μ to guarantee efficient numerical recovery while minimizing cost function of the form,

$$\text{minimize} \quad \mathcal{L}(\mathbf{x}) = \sum_{\mu=1}^{M} l(b_\mu, |\langle \mathbf{x}, a_\mu \rangle|).$$

These algorithmic phase retrieval techniques are explained briefly in the following subsections. Comprehensive overview and comparisons can be found in [1].

Note: It is important to observe here that for any solution $\mathbf{x}$ to the above problem, $\mathbf{x}e^{i\phi}$, $\forall \phi \in [0, 2\pi)$ is also a solution. Therefore, the signal can be reconstructed only upto a global phase factor $e^{i\phi}$. This is usually not of any concern in imaging problems.

A-1 Semidifinite Programming

The following cost function is considered,

$$\text{minimize} \quad \mathcal{L}_1(\mathbf{x}) = \frac{1}{2M}\sum_{\mu=1}^{M}(b_\mu^2 - |\langle \mathbf{x}, a_\mu\rangle|^2)^2. \tag{A-1}$$

This method involves rewriting the set of quadratic equations (A-1) as linear equations in a higher dimension. By considering a linear transformation $\mathcal{A}$ that maps the Hermitian matrices into real-valued vectors [1].

$$\begin{aligned}\mathcal{H}^{n\times n} &\to \mathbb{R}^P \\ \boldsymbol{X} &\mapsto \{a_\mu^H \boldsymbol{X} a_\mu\}_{1\le\mu\le P} \\ \mathbf{b} &= \mathcal{A}(\mathbf{x}\mathbf{x}^H)\end{aligned}$$

Uniqueness is guaranteed by oversampling and stability can be achieved if the measurements are random (i.e., a_μ are random vectors). Unfortunately, this good precision comes at a high computational cost as a higher dimensional problem has to be solved.

PhaseLift. The constraint $\boldsymbol{X} = \mathbf{x}\mathbf{x}^H$ is equivalent to requiring the matrix $\boldsymbol{X}$ to have rank 1. Also, this matrix is positive semi-definite. Then this can be cast as matrix recovery problem.

$$\begin{aligned}\text{minimize} \quad & \text{rank}(\boldsymbol{X}) \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{X}) = \mathbf{b} \\ & \boldsymbol{X} \succeq 0.\end{aligned}$$

With convex relaxation, this can be written as a trace-minimization problem,

$$\begin{aligned}\text{minimize} \quad & \text{Tr}(\boldsymbol{X}) \\ \text{subject to} \quad & \mathcal{A}(\boldsymbol{X}) = \mathbf{b} \\ & \boldsymbol{X} \succeq 0.\end{aligned}$$

Then the solution is factorized as $\mathbf{x}\mathbf{x}^H$ to obtain the solution of the phase retrieval problem up to a scalar multiplier $c \in \mathbb{C}$ sucth that $|c| = 1$.

Limitations. In [32], it is shown that *PhaseLift* requires $M \sim \mathcal{O}(n\log n)$ random Gaussian measurements to yield true vector $\mathbf{x}$ with large probability. The algorithm also works well on more structured, non-random phase retrieval problems [33].

Unfortunately, the precision that PhaseLift offers comes at the cost of a high computational cost making it impractical for online implementation but allows efficient solving as the constraints are singleton.

A-2 Alternating Projections

These constitute the most popular class of phase-retrieval methods. The method pioneered by Gerchberg and Saxton [34], (referred to as GS algorithm hereinafter), recovers the complex image from magnitude measurements at two different planes–the real (imaging) plane and Fourier (diffraction) plane. The algorithm is iterative and involves projections, imposing real and Fourier plane constraints. However, recovery to the true solution is not guaranteed due to non-convexity of the constraints and it can get stuck in local minimum. Modifications have been proposed by Fienup [35] to overcome stagnation problems. The Fienup algorithm, however, requires several initializations to increase performance. A comparison of these iterative algorithms with convex-optimization approaches can be found in [1].

A variant of the GS algorithm is described here [36]. The following non-smooth cost function is minimized,

$$\text{minimize} \quad \mathcal{L}_2(\mathbf{x}) = \sum_{\mu=1}^{M} (b_\mu - |\langle \mathbf{x}, a_\mu \rangle|)^2.$$

. The method relies on the following two-step procedure,

- an initialization scheme.
- Iterative minimization of cost function $\mathcal{L}_2(\mathbf{x})$ using a "local search" alternating projections algorithm.

Initialization The first step has important practical consequences. The initialization depends on the probability distribution of the measurement vectors a_μ's. When these vectors are independent and Gaussian, a good initialization procedure is described in [36], and has been summarized in Table. Then, it has been shown that the algorithm should converge, at a linear rate, with high probability, for $M \sim \mathcal{O}(n)$.

Assuming that the measurement vectors are known, a measurement matrix $A \in \mathbb{C}^{M \times n}$ can be defined as,

$$A = \begin{pmatrix} a_1^H \\ \vdots \\ a_M^H \end{pmatrix}.$$

The associated *phase retrieval* problem is,

$$\text{reconstruct} \quad \mathbf{x} \quad \text{from} \quad \mathbf{b} = |A\mathbf{x}|.$$

It is assumed that the measurement matrix is independent from $\mathbf{x}$ and is injective (true for $M > 4n$ for generic measurement vectors [37]). Then, it is sufficient to recover $z = A\mathbf{x}$ in the intersection of the following two sets,

- $z \in \{z' \in \mathbb{C}^m, |z'| = b\}$;
- $z \in \text{Range}(A)$.

Limitations. Although the use of non-convex algorithm offers an improvement over convex-relaxation algorithms in terms of minimal computational costs, they are based on gradient descent and enjoy weak form of convexity only in the neighbourhood of the solution. Therefore, it is required to have a good initialization and/or much more number of measurements in relation to the dimension of the signal to be reconstructed. The last requirement increases the probability by decreasing the stagnation points in the optimization.

Moreover, the algorithm is prone to several sources of instability, most prominently the possibility of division by zero (for eg., due to the numerical instability of measurement vectors). Therefore, it is important to take machine precision into consideration in context of implementation.

Appendix B

RBF-QR algorithm

Severe ill-conditioning occurs especially in the flat basis function limit ($\epsilon \to 0$). An alternative basis has been proposed in this limit that spans the same space and is based upon the polynomial expansion of Gaussian functions. In this limit, it has been proved that limiting interpolants exist and converge to the form of polynomials [24, 38, 39].

The algorithm is explained for the 2-dimensional case. Here, the RBF centred at node points $\underline{\varrho}_k = (r_k, \theta_k)$ are evaluated at points $\underline{\varrho} = (r, \theta)$.

The standard radial basis function, $\phi_k(\underline{\varrho})$, is expanded on a suitable expansion function $t_l(\underline{\varrho})$,

$$\phi_k(\underline{\varrho}) = \sum_{l=1}^{\infty} c_l(\underline{\varrho}_k) d_l t_l(\underline{\varrho}), \quad k = 1, \ldots, N.$$

This infinite expansion can be truncated at a finite value $M \geq N$, determined based on machine precision and shape parameter value, and written in matrix form as

$$\Phi(\underline{\varrho}) \approx C \cdot D \cdot T(\underline{\varrho}), \tag{B-1}$$

The matrix $C \in \mathbb{R}^{N \times M}$ contains the elements $c_{kl} = c_l(\underline{\varrho}_k)$. In this thesis, the expansion functions are Chebyshev polynomials for the radial component, with a Gaussian weighing function along that dimension and trigonometric functions for the angular component. The expressions for the elements of the coefficient matrix C and expansion function $T(\underline{\varrho})$ can be found in [24]. The diagonal scaling matrix D, with its elements as d_l, contains the shape parameter

$$d_l = \frac{2\lambda^l}{l!}.$$

The elements d_l decay rapidly with l for small shape parameters. The goal of this factorization is to confine the ill-conditioning, due to scaling of RBF, to the matrix D which can be safely inverted.

Left multiplication with any non-singular matrix in equation (B-1) results in a new basis which spans the same space as the original basis. To achieve a well-conditioned basis, the

matrix C is split using QR factorization,

$$\begin{aligned}\Phi(\underline{\varrho}) &= (Q \cdot R)D \cdot T(\underline{\varrho}) = Q \cdot \begin{bmatrix} R_1 & R_2 \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \cdot T(\underline{\varrho}) \\ &= Q \cdot \begin{bmatrix} R_1 D_1 & R_2 D_2 \end{bmatrix},\end{aligned}$$

where $Q \in \mathbb{R}^{N \times N}$ is an orthogonal matrix and $R \in \mathbb{R}^{N \times M}$ is upper triangular. The matrices R_1 and D_1 are $N \times N$. The expansion functions $T(\underline{\varrho})$ is chosen to be better conditioned and insensitive to the value of λ. The new basis is chosen as

$$\Psi(\underline{\varrho}) = D_1^{-1} R_1^{-1} Q^H \Phi(\underline{\varrho}) = \begin{bmatrix} I & D_1^{-1} R^{-1} R_2 D_2 \end{bmatrix} \cdot T(\underline{\varrho}) = \begin{bmatrix} I & \tilde{R} \end{bmatrix} \cdot T(\underline{\varrho}).$$

It is asserted that the unstable effect of the leading powers of λ are contained in D_1 and D_2, but the resulting effect on $\tilde{R}$ is harmless.

The approximation problem is solved as a least square problem for coefficients $\mathbf{c}$,

$$A\mathbf{c} = f.$$

The matrix A is computed in the new basis as

$$A = \begin{bmatrix} \Psi(\underline{\varrho}_1) & \dots & \Psi(\underline{\varrho}_M) \end{bmatrix}^T \begin{bmatrix} I \\ \tilde{R}^T \end{bmatrix} = T_1^T + T_2^T \tilde{R}^T,$$

where $T_1(\underline{\varrho})$ contains the first N expansion functions and $T_2(\underline{\varrho})$ contains the remaining expansion functions evaluated at the same location.

The RBF-QR method explained here may suffer from numerical overflow as the expansion coefficients, d_l start to diverge quickly depending upon the value of λ. Therefore, this method is not suitable for larger values of λ and the standard basis can be used instead. Also, this method is generally more suitable for circular domains but other methods exist and either require multi precision computing toolbox for MATLAB or a more complicated variant of RBF-QR explained here [39]. These methods are not studied as a part of this thesis.

Appendix C

Rippa's LOOCV algorithm

The goal of a standard interpolation problem is to find a continuous function Ψ_f that interpolates the given function values $f(\underline{\varrho}_i)$ on some scattered data sites $X = \{\underline{\varrho}_1, \ldots, \underline{\varrho}_N\}$ such that,

$$\Psi_f(\underline{\varrho}_i) = f(\underline{\varrho}_i), \quad i = 1, \ldots, N.$$

An RBF interpolant is of the form,

$$\Psi_f(\underline{\varrho}) = \sum_{j=1}^{N} c_j \phi(\left\| \underline{\varrho} - \underline{\varrho}_j \right\|).$$

Here, the basis ϕ is assumed to be strictly positive definite and the least squares solution for the coefficients $\mathbf{c} = [c_1, \ldots, c_N]^T$ is given as,

$$\mathbf{c} = \mathbf{A}^{\dagger}\mathbf{f},$$

where $A_{ij} = \phi(\left\| \underline{\varrho}_i - \underline{\varrho}_j \right\|)$ and $\mathbf{f} = [f(\underline{\varrho}_1, \ldots, f(\underline{\varrho}_N)]^T$.

A variant of cross validation method, *leave-one-out cross validation* (LOOCV)algorithm, a cost function based on errors for a sequence of partial fits to the data is minimized to estimate the optimal value of the shape parameter λ. This partial fit error is computed at a single data point split from the rest of the data ($N-1$ points), on which the approximation of the interpolant is based. This is repeated for each one of the N data points. The cost function is provided by some norm of the so found error vector. The predicted optimal is usually close to the actual optimum.

For a vector of data sites with the point $\underline{\varrho}^{[k]}$ removed defined as,

$$\underline{\varrho}^{[k]} = [\underline{\varrho}_1, \ldots, \underline{\varrho}_{k-1}, \underline{\varrho}_{k+1}, \ldots, \underline{\varrho}_N]^T,$$

the partial RBF interpolant to the data $f^{[k]}$ is defined as,

$$\Psi_f^{[k]}(\underline{\varrho}) = \sum_{j=1}^{N-1} c_j^{[k]} \phi(\left\| \underline{\varrho} - \underline{\varrho}_j^{[k]} \right\|).$$

Algorithm 1 Rippa's LOOCV algorithm

Input: Data sites $\underline{\varrho}$, function values $f(\underline{\varrho})$. Fix λ

procedure LOOCV ALGORITHM(ϕ,f,$\underline{\varrho}$)

2: **while** $k = 1, \ldots, N$ **do** ▷ Iterate for each data site

$\underline{\varrho}^{[k]} = [\underline{\varrho}_1, \ldots, \underline{\varrho}_{k-1}, \underline{\varrho}_{k+1}, \ldots, \underline{\varrho}_N]^T$,

4: $\Psi_f^{[k]}(\underline{\varrho}) = \sum_{j=1}^{N-1} c_j^{[k]} \phi(\left\| \underline{\varrho} - \underline{\varrho}_j^{[k]} \right\|)$ ▷ Partial fit.

$e_k = \left| f(\underline{\varrho}_k) - \Psi_f^{[k]}(\underline{\varrho}_k) \right|$ ▷ error estimate at k^{th} data site.

6: $e_k = \frac{c_k}{\mathbf{A}_{kk}^{-1}}$ ▷ Rippa's step.

return cost vector $\mathbf{e} = [e_1, \ldots, e_N]^T$.

Output: Optimal shape parameter λ by minimizing $\|\mathbf{e}\|$.

Then the algorithm is summarized in Algorithm 1.

Rippa showed that the step 5 can be simplified to step 6 where c_k is the k^{th} expansion coefficient of the interpolant Ψ_f based on full data set, and $\mathbf{A}_{kk}^{-1}$ is the k^{th} diagonal element of the inverse of the corresponding interpolation matrix. It is noted here that step 7 can be computed in a single step. The MATLAB function `fminbnd` is used to find the minimum of the cost function in step 7.

Bibliography

[1] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

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

[3] M. Verhaegen, G. Vdovin, and O. Soloviev, "Control for high resolution imaging." 2016.

[4] J. Antonello and M. Verhaegen, "Modal-based phase retrieval for adaptive optics," *JOSA A*, vol. 32, no. 6, pp. 1160–1170, 2015.

[5] D. R. Luke, J. V. Burke, and R. G. Lyon, "Optical wavefront reconstruction: Theory and numerical methods," *SIAM review*, vol. 44, no. 2, pp. 169–224, 2002.

[6] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[7] B. R. A. Nijboer, *The diffraction theory of aberrations*. PhD thesis, Wolters, 1942.

[8] G.-m. Dai and V. N. Mahajan, "Orthonormal polynomials in wavefront analysis: error analysis," *Applied optics*, vol. 47, no. 19, pp. 3433–3445, 2008.

[9] W. Swantner and W. W. Chow, "Gram–schmidt orthonormalization of zernike polynomials for general aperture shapes," *Applied optics*, vol. 33, no. 10, pp. 1832–1837, 1994.

[10] R. Upton and B. Ellerbroek, "Gram–schmidt orthogonalization of the zernike polynomials on apertures of arbitrary shape," *Optics letters*, vol. 29, no. 24, pp. 2840–2842, 2004.

[11] S. Van Haver, *The Extended Nijboer-Zernike diffraction theory and its applications*. PhD thesis, Delft University of Technology, 2010.

[12] A. J. Janssen, "Extended nijboer–zernike approach for the computation of optical point-spread functions," *JOSA A*, vol. 19, no. 5, pp. 849–857, 2002.

[13] M. Maksimovic, "Optical design and tolerancing of freeform surfaces using anisotropic radial basis functions," *Optical Engineering*, vol. 55, no. 7, pp. 071203–071203, 2016.

[14] A. Martínez-Finkelshtein, A. M. Delgado, G. M. Castro, A. Zarzo, and J. L. Alió, "Comparative analysis of some modal reconstruction methods of the shape of the cornea from corneal elevation data," *Investigative ophthalmology & visual science*, vol. 50, no. 12, pp. 5639–5645, 2009.

[15] A. Martinez-Finkelshtein, D. Ramos-Lopez, and D. Iskander, "Computation of 2d fourier transforms and diffraction integrals using gaussian radial basis functions," *Applied and Computational Harmonic Analysis*, vol. 43, no. 3, pp. 424–448, 2017.

[16] J. J. Braat, S. van Haver, A. J. Janssen, and P. Dirksen, "Assessment of optical systems by means of point-spread functions," *Progress in Optics*, vol. 51, pp. 349–468, 2008.

[17] R. K. Tyson, *Principles of adaptive optics*. CRC press, 2015.

[18] P. Dirksen, J. Braat, A. J. Janssen, and C. Juffermans, "Aberration retrieval using the extended nijboer-zernike approach," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 2, no. 1, pp. 61–68, 2003.

[19] R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Optical Engineering*, vol. 21, no. 5, pp. 829–832, 1982.

[20] Z. von F, "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode," *Physica*, vol. 1, no. 7-12, pp. 689–704, 1934.

[21] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *JOsA*, vol. 66, no. 3, pp. 207–211, 1976.

[22] C. Van der Avoort*, J. Braat, P. Dirksen, and A. Janssen, "Aberration retrieval from the intensity point-spread function in the focal region using the extended nijboer–zernike approach," *Journal of Modern Optics*, vol. 52, no. 12, pp. 1695–1728, 2005.

[23] G. E. Fasshauer, *Meshfree Approximation Methods with Matlab:(With CD-ROM)*, vol. 6. World Scientific Publishing Co Inc, 2007.

[24] B. Fornberg, E. Larsson, and N. Flyer, "Stable computations with gaussian radial basis functions," *SIAM Journal on Scientific Computing*, vol. 33, no. 2, pp. 869–892, 2011.

[25] M. Montoya-Hernández, M. ServıÌĄn, D. Malacara-Hernández, and G. Paez, "Wavefront fitting using gaussian functions," *Optics Communications*, vol. 163, no. 4, pp. 259–269, 1999.

[26] N. Guberman, "On complex valued convolutional neural networks," *arXiv preprint arXiv:1602.09046*, 2016.

[27] M. F. Amin and K. Murase, "Single-layered complex-valued neural network for real-valued classification problems," *Neurocomputing*, vol. 72, no. 4, pp. 945–955, 2009.

[28] C. M. Harding, R. A. Johnston, and R. G. Lane, "Fast simulation of a kolmogorov phase screen," *Applied optics*, vol. 38, no. 11, pp. 2161–2170, 1999.

[29] B. Paul, L. Mugnier, J.-F. Sauvage, K. Dohlen, and M. Ferrari, "High-order myopic coronagraphic phase diversity (coffee) for wave-front control in high-contrast imaging systems.," *Optics express*, vol. 21, no. 26, pp. 31751–31768, 2013.

[30] M. Ygouf, L. M. Mugnier, D. Mouillet, T. Fusco, and J.-L. Beuzit, "Simultaneous exoplanet detection and instrument aberration retrieval in multispectral coronagraphic imaging," *Astronomy & Astrophysics*, vol. 551, p. A138, 2013.

[31] T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper, *et al.*, "High-order adaptive optics requirements for direct detection of extrasolar planets: Application to the sphere instrument," *Optics Express*, vol. 14, no. 17, pp. 7515–7534, 2006.

[32] J. Emmanuel, "Candes, thomas strohmer, and vladislav voroninski," *Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. CoRR, abs/1109.4499*, 2011.

[33] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[34] R. W. Gerchberg, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[35] J. R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.

[36] I. Waldspurger, "Phase retrieval with random gaussian sensing vectors by alternating projections," *arXiv preprint arXiv:1609.03088*, 2016.

[37] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM review*, vol. 57, no. 2, pp. 225–251, 2015.

[38] G. E. Fasshauer and M. J. McCourt, "Stable evaluation of gaussian radial basis function interpolants," *SIAM Journal on Scientific Computing*, vol. 34, no. 2, pp. A737–A762, 2012.

[39] S. A. Sarra and S. Cogar, "An examination of evaluation algorithms for the rbf method," *Engineering Analysis with Boundary Elements*, vol. 75, pp. 36–45, 2017.