Department of Precision and Microsystems Engineering

Measurement Uncertainty Analysis of a Twyman-Green Interferometer for Lens Testing

Lyuben Davidov

Report no : 2023.044
Coach : Dr. ing. F. Zijp
Professor : Dr. N. Bhattacharya
Specialisation : Micro & Nano Engineering
Type of report : Master Thesis
Date : 10.07.2023

TUDelft Delft University of Technology

Challenge the future

Measurement Uncertainty Analysis of a Twyman-Green Interferometer for Lens Testing

MSc. Mechanical Engineering
Thesis

Lyuben Davidov
July 2023

Faculty of Mechanical, Maritime, and Materials Engineering
Department of Precision and Microsystems Engineering

Supervisors:

Dr. N. Bhattacharya

Dr. ing. F. Zijp

Prof. Dr. ir. W.M.J.M. Coene

Dr. ir. S. Konijnenberg

Abstract

There is a growing demand for precision and high quality optical objectives and lenses due to their numerous advantages and uses. The market for high precision objective lenses has tremendous potential for development and is a sector that is expanding quickly. [1]

Measuring the aberrations of optical systems is an essential step in the fabrication of high precision optical components. However, when working at the cutting-edge of technology, it is increasingly difficult to provide trustworthy measurements as the used metrology instrument has to be of comparable or higher precision. This poses a major problem especially when working with high numerical aperture (NA) optics.

In this thesis, we will analyze and quantify the measurement uncertainty of a Twyman-Green interferometer used for lens testing of high-NA microscope objectives.

To quantify the measurement uncertainty of the interferometer, various sources of uncertainty that affect the accuracy and precision of the measurements are considered. These include environmental and instrumentation factors such as incorrect phase-stepping, laser instability, camera noise, stray light, photon shot noise, effects of mid-spatial frequencies originating in the optical reference, as well as computational shortcomings such as: incorrect phase unwrapping, polynomial fit errors, incorrect pupil scaling and edge detection.

By carefully analyzing these individual sources of uncertainty and their impact, we determine the overall measurement uncertainty of the interferometer and provide an assessment of its accuracy through Monte Carlo simulations, where the introduced uncertainties are obtained from real measurement data. The uncertainty analysis procedure described in this paper is a useful tool that can also be applied to different types of interferometers by taking proper considerations into account.

Contents

1   Introduction

Optical systems are an indispensable part of the modern world and find their application in numerous aspects of our life. From cameras, contact lenses, and lighting systems, to chip manufacturing, medical equipment, and gravitational wave detection. As optical systems consist of various optical components of all shapes and sizes, such as lenses, mirrors, prisms, etc., it is important to be able to measure whether an optical component or an optical system meets certain design criteria in terms of its performance. For example, controlling whether an optical flat has a sufficiently low surface error, or whether an aspherical lens has the correct aspherical profile and correctly focuses the incoming light within tolerances.

Twyman-Green interferometers are a popular choice for lens testing. A Twyman-Green interferometric setup can be used to measure the wavefront aberrations from which one can draw a conclusion on the quality of the imaging system in question. With the advent of digital interferometry, it was possible to increase the resolution of interferometric measurements by means of recording a series of digital interferograms with varying spatial phase and retrieve it pointwise. There are several different steps involved, including – phase retrieval, phase unwrapping, and wavefront analysis. During each step, various algorithms are used depending on the situation. However, all algorithms are prone to numerical errors due to their discrete nature which brings errors in the final results. Moreover, measurement uncertainty may be introduced by external sources, such as nonlinearities in the phase shift, instability of the light source, detector nonlinearities, electronic noise, vibrations. To achieve meaningful measurement results, it is important to quantify the measurement uncertainty of the metrology setup.

In this thesis, we first investigate the existing literature on interferometric measurements using phase-stepping, phase unwrapping, and Zernike polynomial fitting techniques, as well as describe common errors arising from the instrumentation and environment. Then a methodology is proposed for quantifying errors using different procedures. In the conducted experiments a high-NA microscope objective is tested against a reference spherical mirror of slightly higher numerical aperture. The total measurement uncertainty of the interferometer is then evaluated using Monte Carlo simulations, where the results show a consistent measurement uncertainty of about $\pm 3 m\lambda$ per Zernike coefficient.

2   Problem Statement

The purpose of this project is to develop and evaluate a lens-testing procedure using a Twyman-Green Phase Stepping Interferometry (PSI) setup for the aberration metrology of diffraction-limited high-NA infinite conjugate lenses. The procedure is able to detect and measure the intrinsic aberrations of the test lens/objective, correct for the aberrations originating from the repherence spherical mirror and other systematic errors, and has a set of uncertainty values per calculated Zernike coefficient.

It is necessary to evaluate a Twyman-Green interferometric setup that can test the wavefront

aberrations of a lens or objective, for short referred to as an "Optical Imaging System" (OIS). The OIS is to be viewed as a black box. The information that is present is as indicated by the supplier and is therefore limited. That includes the focal length, the numerical aperture, and the location of the entry and exit pupil planes. Information regarding the individual lenses in an objective or their respective coatings is a company secret and therefore not available.

The lenses under test have a specified "field size", in the order of $50\mu m$ in diameter, meaning that a beam that enters the lens at an angle to the optical axis (i.e. a beam with a field offset) introduces aberrations that must fall within certain spec limits. The metrology procedure should be able to measure the aberrations over the full field size.

Depending on the chosen phase retrieval technique, phase unwrapping algorithm, and polynomial fitting for wavefront analysis, we can either increase or decrease both the accuracy and precision of our measurement results for the wavefront aberrations. The goal is to obtain meaningful results for the measurements, select and evaluate different algorithms in terms of their performance, and determine an error budget for the measurements.

3   Test Setup

A Twyman-Green interferometer, together with a moving flat reference mirror is used to create phase-shifted interferograms that are imaged on the CMOS sensor of a camera. The type of light source is a monochromatic continuous-wavelength (CW) laser. Several different wavelengths are used for testing as the microscope objective is designed to operate over a wide wavelength spectrum. An overview of the used CW lasers can be found in Table 1.

Table 1: Light sources.

λ [nm]	Power: [W]	Class:	Type:	Diameter:
405	0.004	3B	Single channel benchtop laser diode unit	2 inch
406	0.004	3B	Multi-diode fiber coupled unit	2 inch
488	0.018	3B	Multi-diode fiber coupled unit	2 inch
520	0.008	3B	Multi-diode fiber coupled unit	2 inch
543	0.005	3R	Green HeNe	1 inch
650	0.017	3B	Multi-diode fiber coupled unit	2 inch
705	0.010	3B	Multi-diode fiber coupled unit	2 inch
850	0.011	3B	Multi-diode fiber coupled unit	2 inch
1064	0.025	3B	Multi-diode fiber coupled unit	2 inch
1310	0.060	3B	Multi-diode fiber coupled unit	2 inch
1550	0.01	3B	Multi-diode fiber coupled unit	2 inch

The source delivers light through a polarization-maintaining glass fiber, which is then vertically polarized using a polarizer, and expanded to the "Diameter", as indicated in Table 1, and collimated. A shearing plate is used to confirm that the beam is indeed collimated and the collimator lens is positioned accordingly. The layouts is depicted in Fig.(1).
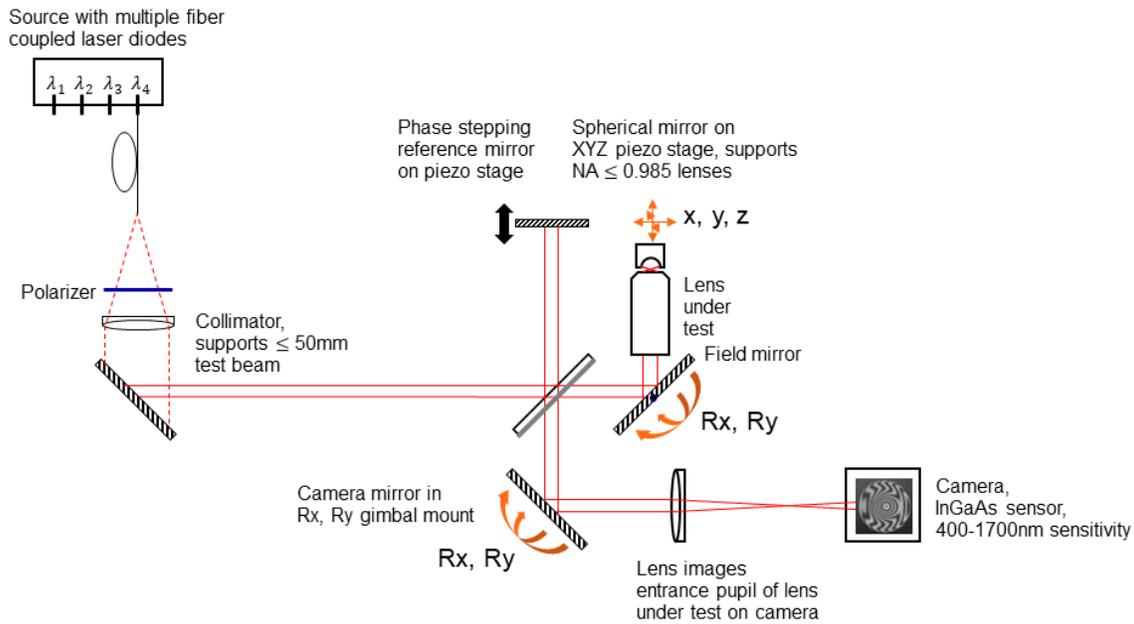

Figure 1: Twyman-Green setup for testing an infinite conjugate lens

The laser beam is expanded and a planar mirror is used in the reference arm. We can choose between multiple wavelengths as indicated in Table 1, however, the position of the collimator lens needs to be readjusted in order to prevent chromatic aberrations. A retroreflector and an additional Fizeau interferometer, not shown above, are used to assist in the alignment of the various optical component as depicted in the setup. The incoming laser light is vertically polarized out of the plane of Fig.(1) in order to achieve an approximately 50/50 split ratio in the beam splitter.

The resulting two beams are propagated through the test and reference arms and back, and a controlled phase shift is introduced by moving the planar reference mirror using a piezo actuator. The optical path distance in the test and reference arm is equal. The beams are propagated back, and recombined, as will be explained in section 4.1, and subsequently, an imaging lens, located in front of the camera, images the test lens' exit pupil on the CMOS sensor of the camera. The image of the exit pupil is magnified by a factor of 2.

There are two ways to propagate back the beam in the test arm – by using a spherical high-NA mirror, as depicted in Fig.(1,2), and by using a planar mirror in the focal plane of the lens under test.

In section 2, it was mentioned that the aberrations over the "field size" of the lens need to be measured, for which purpose the beam needs to enter the lens at an angle to the optical axis. A spherical mirror with a high NA allows to back-propagate the beam. The principle is shown in Fig.(2).
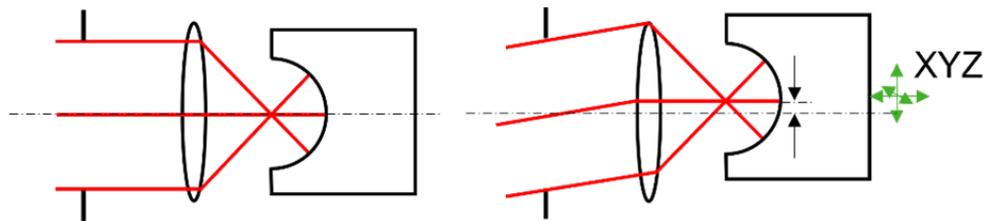
Figure 2: Spherical mirror with an on-axis and skewed incoming beam for a telecentric lens.

Both the lens under test and the mirror are positioned using their own X, Y, Z-stages with μm precision. The stages can be controlled by either μm hand spindles or stepper motors. The hand spindles have the advantage of faster positioning when we change between lenses, while the stepper motors can be controlled by the computer and are used to automate measurements. An example of the measurement arm for the setup with a spherical mirror can be seen in Fig.(3).

Figure 3: Measurement arm of the interferometer with the mounted spherical mirror

The spherical mirror is used for both on-axis and off-axis measurements. A planar mirror, for example, can be used for off-axis measurements as outlined on the right of Fig.(2) but the measurement will only consist of even aberrations as the odd aberrations will change sign upon reflection and cancel out on their way back throught the lens/objective. However, the spherical mirror also introduces aberrations, which if not accounte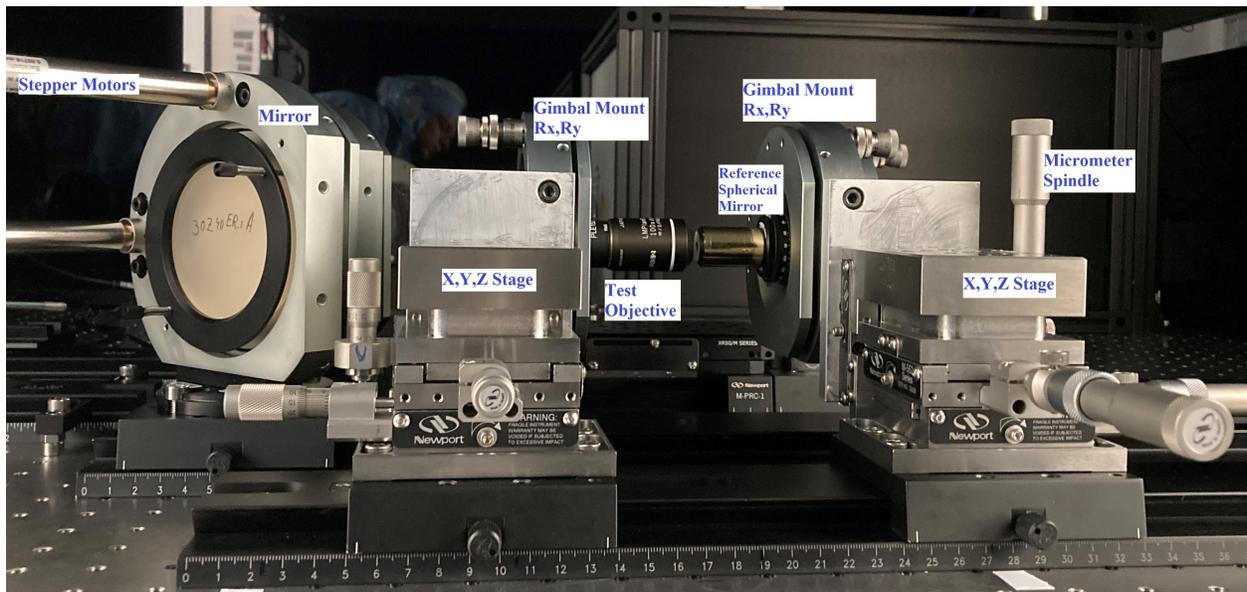d for, will combine with the aberrations present in the lens under test, that we want to measure. Therefore, it is necessary to properly quantify its full aberration profile including both the odd and even aberrations introduced by it.

Due to the geometry and existing polishing techniques, a planar surface is easier to produce and is of better quality than a spherical surface. This can be observed in the interferogram shown in Fig.(4), where you can see concentric grooves. They are especially visible in the areas marked by the 3 small circles. The spherical mirror used in our project has been machined using "Single-Point Diamond Turning". The concentric grooves originate from the machining process and testify that there are indeed imperfections present in the mirror.

A flat mirror, on the other hand, is produced using polishing and grinding, where two flat surfaces are ground against one another. This produces better results as there are factors that can introduce defects during the turning process, such as "beating", "chatter", "chipping" or "dulling" of the cutting tool. For these reasons, a spherical mirror of high quality (low surface roughness) would be very expensive and labor involving to produce. On the other hand, a planar mirror of comparable quality would cost significantly less and can be mass-produced.
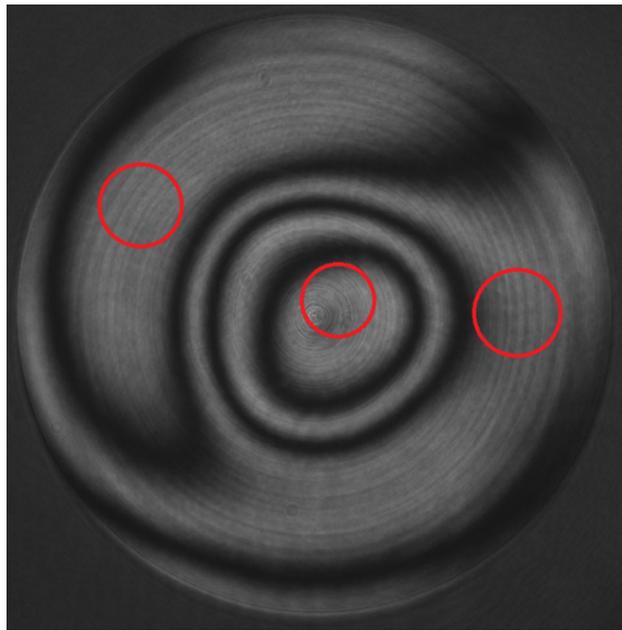

Figure 4: An interferogram of a low-NA test lens, created with the spherical mirror.

We will proceed to highlight the approach for determining the aberration profile of the spherical

mirror. The measured aberrations are represented in terms of Zernike polynomials, therefore individual polynomials may be angularly even or odd. This means that we can change the sign of odd aberrations by rotating the spherical mirror by 180°. An example of the effect of rotation on 3rd order coma along the x-axis (Z_{-1}^3) can be seen in Fig.(5, 6). The sum of the two functions is 0, hence why rotation is a suitable way to find the coefficients of the odd aberrations in our spherical mirror.
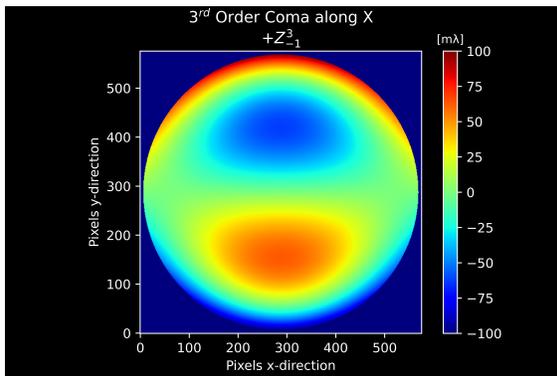

Figure 5: $100m\lambda$ of Z_{-1}^3
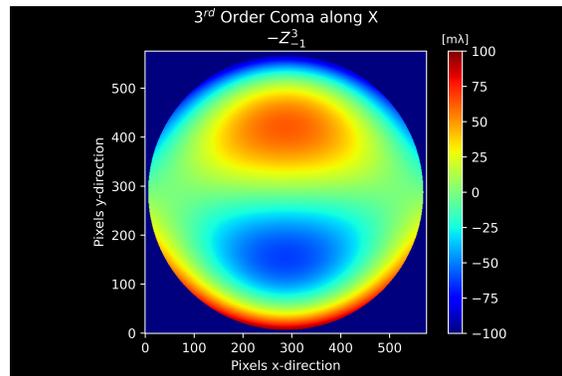
Figure 6: $100m\lambda$ of Z_{-1}^3, 180° rotated

The even aberrations of the spherical mirror are determined using an additional on-axis measurement with a planar mirror. To determine the full aberration profile of the spherical mirror 3 measurements are required. The first one using a spherical mirror at 0°, the second one using a spherical mirror rotated by 180°, and a third measurement using a planar mirror located at the focal point of the lens/objective. The full procedure is explained in detail in section (6.2.1) and the aberration profile of the spherical mirror is shown.

4 Theory: Lens Testing Interferometry

There are three main operations involved in the measurement process – phase retrieval through phase stepping, phase unwrapping, and a polynomial fitting . It is first necessary to understand how they are used together to conduct a measurement.

First, the two beams, in the measurement arm and reference arm are brought together in order to interfere and create an interferogram. The interferogram is converted into an electric signal using the CMOS sensor of the camera and sent to the computer. The sensor measures the irradiance over the area of each pixel and stores it as a data point in an array.

The interferogram is a measurement of the irradiance for a set of data points. From this measurement, we would like to extract information regarding the phase, and therefore reconstruct the wavefront. This is done using "Phase Shifting Interferometry" (PSI) algorithms.

Due to the nature of the PSI algorithms, the calculated phase is between $-\pi$ and $+\pi$. This comes from the fact that the phase is calculated using an arctangent function which gives discontinuities at the edges of the specified range. A wavefront, however, is a smooth geometrical surface and the discontinuities must be removed by "unwrapping" the phase. This is a crucial step as an incorrectly unwrapped wavefront yields incorrect aberration results.

Once the phase has been calculated it can be converted to surface errors or optical path length differences to obtain the wavefront using Eq.(10). The wavefront can then be represented as a weighted sum of individual polynomials that are orthogonal over the shape of the pupil. Different sets of polynomials are orthogonal over different shapes and the correct set must be chosen based on the pupil shape of our system. The high-NA lenses and objectives that take part in the experimental validation of this thesis are all designed to be circularly symmetric, hence their pupils are circular as per the nominal design. However, "as-built" lenses are no longer perfectly symmetric which can lead to errors and must be considered. For systems with circular and approximately circular symmetry, the common set of used polynomials is "Zernike polynomials" as they are orthogonal over the unit circle. The amount of a certain aberration that is present in the measured wavefront, as obtained through phase-unwrapping is determined by the coefficient for each individual polynomial in the set.

This raises the question what errors can each operation introduce and how do they affect the results altogether. In view thereof, assessment of the measurement uncertainty of an interferometer is a difficult task due to the numerous processes involved. It may be advantageous to separate the problem into two main categories – computational and physical aspects. Computational aspects refer to the many different choices for "Phase Stepping Algorithms", "Phase Unwrapping Algorithms" and "Wavefront Fitting Techniques" available to us, including their corresponding strengths and weaknesses. We look into different ways to quantify their suitability and capability to compensate for errors and undesired external influences, as well as the numerical errors they may introduce. A flowchart showing a general overview of the process is shown in Fig.(7).

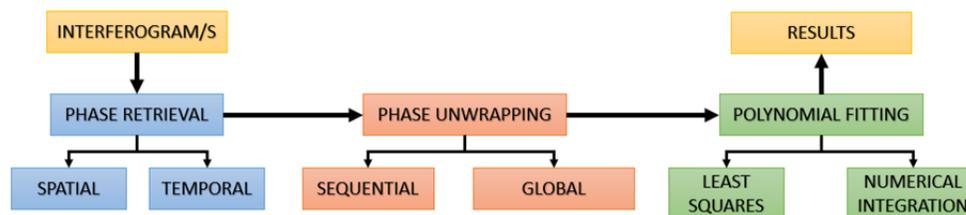Figure 7: The computational process of measuring the wavefront aberrations, including the major algorithm subtypes.

Physical aspects, on the other hand, refer to effects on the measurements originating in the test setup, such as phase-shifter errors, camera quantization, shot noise, electronic noise, and external vibrations. Different approaches as found in the literature will be considered, in order to limit their effect.

4.1 Two-Beam Interference

To understand the working principle of the setup we will briefly explain what interference is and how two-beam interferometry works. The interference phenomenon was first discovered by Isaac Newton, where he observed localized concentric fringes, called "Newton rings" around the point of contact of a plano-convex lens with a large radius of curvature, pressing against a flat glass plate. Later on, Thomas Young proceeded to explain the phenomenon in terms of wave optics. In the book "Optics" by E. Hecht the following short description is given: "Optical interference corresponds to the interaction of two or more light waves yielding a resultant irradiance that deviates from the sum of the component irradiances..." [2]. It follows that interference cannot be explained by geometrical optics and is instead caused by the wave nature of light. We can think of light as a wave motion possessing crests and valleys. When two correlated and coherent waves meet, if they are in phase – that is the position of their crests overlap and respectively the position of their valleys as well, then we can observe constructive interference where the two waves strengthen each other. On the contrary, if the waves are out of phase – that is the position of the crests of the first wave overlap with the position of the valleys of the second wave, then we observe destructive interference where the waves cancel one another. This can be seen in Fig.(9).
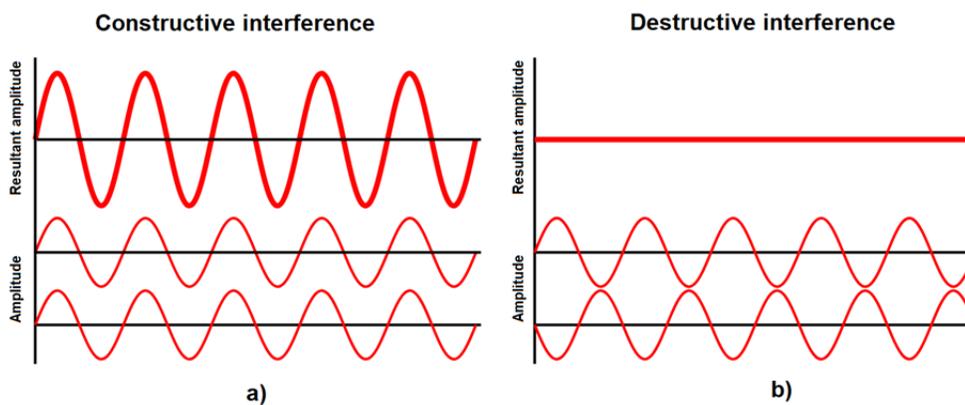
Figure 8: Examples of two-beam interference. [source: wikipedia]

Light is a vector phenomenon, and it obeys the principle of superposition, therefore it is possible to combine two or more waves and express the combined wave as a vector sum of its components. We will restrict the following derivation to two waves. The interference pattern from the interaction of two light waves shown on a screen, in the sense of the relative irradiance of the screen, is the intensity profile of the combined wave, that results from the superposition of the electric fields of the two individual waves. We will denote the intensity profile of the first time-harmonic wave as $U_A(x, y, z, t)$ and of the second one as $U_B(x, y, z, t)$.

$$U_A(x, y, z, t) = \vec{A}_0 e^{i\vec{k}\cdot\vec{r}} e^{i(wt + \phi_A)}$$
$$U_B(x, y, z, t) = \vec{B}_0 e^{i\vec{k}\cdot\vec{r}} e^{i(wt + \phi_B)}$$

(1)

Where $\vec{A}_0$, $\vec{B}_0$ denote the amplitude and polarization direction of the two waves, $\vec{k}$ denotes the wave vector, $\vec{r} = \begin{bmatrix} \hat{i}, & \hat{j}, & \hat{k} \end{bmatrix}^T$ denotes the unit vector, w denotes the angular frequency of the oscillations in time, ϕ_A, ϕ_B denote the phases of the two waves. We can simplify the two expressions by considering the case where the two waves propagate in the same direction. We can therefore omit the $\vec{k} \cdot \vec{r}$ dependency and rewrite the expressions and give the resulting wave as:

$$
\begin{aligned}
U_A(x, y, z, t) &= \vec{A}_0 e^{-i(wt+\phi_A)} \\
U_B(x, y, z, t) &= \vec{B}_0 e^{-i(wt+\phi_B)}
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
U_C(x, y, z, t) &= U_A(x, y, z, t) + U_B(x, y, z, t) = \vec{A}_0 e^{-i(wt+\phi_A)} + \vec{B}_0 e^{-i(wt+\phi_B)} \\
U_C^*(x, y, z, t) &= U_A^*(x, y, z, t) + U_B^*(x, y, z, t) = \vec{A}_0^* e^{+i(wt+\phi_A)} + \vec{B}_0^* e^{+i(wt+\phi_B)}
\end{aligned}
\tag{3}
$$

Where the $*$ denotes complex conjugate. The irradiance that we observe on a screen is the time average of the magnitude of the field.

$$
I = \| U_C^2(t) \|_T = < U_C U_C^* >_T
\tag{4}
$$

Substituting Eq.(3) into Eq.(4) and assuming $\vec{A}_0$ and $\vec{B}_0$ are parallel gives the following expression:

$$
I = A_0^2 + B_0^2 + 2A_0 B_0 cos(\phi_A - \phi_B)
\tag{5}
$$

Where the modulation term in Eq.(5) proves that the resultant irradiance is not the summation of the individual irradiances of the two waves: $I_A = A_0^2$, $I_B = B_0^2$. In section 4.2 we will show how the irradiance from the interference of two waves given by Eq.(5) can be used to our advantage to extract information about the phase of the combined wave, indicated by $\phi_A - \phi_B$.

4.2 Phase Stepping Algorithms

Phase stepping algorithms (PSAs) are a type of phase-shifting interferometry algorithm where the phase is varied in steps. In general, all phase shifting algorithms are used to acquire the raw phase map from the recorded interferograms – irradiance measurements. The recorded irradiance can be represented as a sum of the average irradiance and the irradiance modulation as seen in Eq.(6).

$$
I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi_t(x, y) - \phi_r(x, y) + \delta(t)]
\tag{6}
$$

Here $I(x, y, t)$ is the irradiance as measured by the CMOS sensor at a certain location – pixel, with coordinates x and y. $I'(x, y)$ represents the average irradiance over the set of recorded interferograms at a given location, also known as the DC component, and $I''(x, y)$ – the irradiance modulation amplitude [3]. The irradiance is modulated by the cosine term, which depends on the phase in the test arm of the interferometer $\phi_t(x, y)$ and the one in the reference arm $\phi_r(x, y)$. We introduce a phase change which is denoted by $\delta(t)$, by varying the OPD in the reference arm of the interferometer in a controlled fashion. This phase change may either be linear – "Integrating Bucket Technique", or in steps – "Phase Stepping".

As the wavefronts from both arms are recombined by the beam splitter and imaged on the camera using the imaging lens, located in front of it, it is more useful to work with the phase difference between the two wavefronts in the test and reference arm, rather than their separate values – ϕ_t, ϕ_r. Therefore, we can introduce a single phase term by substituting: $\phi(x, y) = \phi_t(x, y) - \phi_r(x, y)$.

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos\left[\phi(x, y) + \delta(t)\right] \tag{7}$$

The goal here is to calculate $\phi(x, y)$ which is the "raw phase map". There are two main types of PSAs depending on how the phase $\phi(x, y)$ is detected – "synchronous" and "asynchronous". The most common way to obtain the raw phase, as per literature, is by recording four interferograms where the phase step, $\delta(t)$, varies by $\frac{\pi}{2}$. This algorithm was first described by Wyant and is known as the "4-step" or "4-sample algorithm" [4]. Algorithms where the phase step is known, such as the "4-step" one, are of the "synchronous" type, while algorithms where we have no information about the phase step are called "asynchronous". A version of the same algorithm that relies on the least amount of recorded interferograms was also developed by Wyant and is respectively called the "3-sample algorithm" [5]. Both are very comparable in terms of their inability to compensate for errors and simplicity of the expression, however, the 4-step one will be shown below. First, let us look at the intensity profiles of the recorded interferograms. They are as follows:

$$
\begin{aligned}
I_0(x, y, t) &= I'(x, y) + I''(x, y) \cos\left[\phi(x, y)\right] \\
I_1(x, y, t) &= I'(x, y) + I''(x, y) \cos\left[\phi(x, y) + \frac{\pi}{2}\right] \\
I_2(x, y, t) &= I'(x, y) + I''(x, y) \cos\left[\phi(x, y) + \pi\right] \\
I_3(x, y, t) &= I'(x, y) + I''(x, y) \cos\left[\phi(x, y) + \frac{3\pi}{2}\right]
\end{aligned}
\tag{8}
$$

This is a linear system of four equations and three unknowns - $I'(x, y), I''(x, y), \phi(x, y)$ which can easily be solved. Note that it is possible to use three measurements instead of four, however, four measurements result in a simpler formula for the phase. Using simple trigonometric identities an expression for $\phi(x, y)$ in terms of the measured intensities is obtained:

15

$$\phi(x, y) = \arctan\left[\frac{I_3 - I_1}{I_0 - I_2}\right] \tag{9}$$

The calculated phase map is linked to the wavefront in a simple fashion. A wavefront is a surface over which all points of a wave field have the same phase. An aberrated wavefront may be thought of as the deviation from a perfectly spherical reference wavefront. A flat wavefront can be thought of as a spherical wavefront with radius at infinity. The deviation at a certain point (x,y) on that reference sphere is simply a height error. The height error results in an optical path difference (OPD) which is linked to the wavefront phase according to the following expression:

$$OPD(x, y) = \phi(x, y)\frac{\lambda}{4\pi} \tag{10}$$

While the OPD may be considered as the surface height errors, the wavefront phase $\phi(x, y)$ may be thought of as the wavefront error. In Eq.(10) we see the factor $\frac{\lambda}{4\pi}$, which is the scaling factor used in double pass interferometers such as the Twyman-Green. For single pass interferometers the scaling factor is $\frac{\lambda}{2\pi}$. Therefore the major interest falls on the correct calculation of the phase.

In the case of a perfect imaging system, a deduction from the Malus-Dupin theorem states that "if the rays of a pencil intersect in a single point after refraction through the optical system, then the wavefronts will be spherical" [6]. Therefore, under ideal circumstances, the wavefront in the exit pupil will be perfectly spherical, hence, the phase will be constant over the sphere's surface and would not vary with (x, y). It follows that the points on the surface experience the same optical path distance. In this case, the OPD is zero. However, in the presence of monochromatic aberrations for example, the imperfections of the optical system would manifest as points on the refence sphere having different phase and thus experiencing different optical path distances in the system. The geometric surface over which all points have the same phase would no longer be spherical. The departure, height error, at a certain location from the shape of an ideal sphere is the OPD.

The above-mentioned algorithms are simple, yet not very robust, especially when it comes to variations in the phase step $\delta(t)$. This has led to other, more robust algorithms being developed, such as the Carré, Hariharan algorithm, and others. Throughout the early stages of PSI, the major limitation was the computation power and thus the amount of steps or iterations in the algorithm. However, modern computers have unparalleled capabilities that have made it possible to synthesize more robust and exotic PSI algorithms, albeit at a higher computational price. Algorithms can be designed to limit the effect of certain errors or external influences. Because of the large variety of available algorithms, it is important to categorize them based on their properties.

4.2.1 Fourier Representation of Phase Stepping Algorithms

The Fourier representation of PSAs is a theory developed by several authors, which analyzes the error-compensating abilities of phase stepping algorithms. It is important to note that the authors, who developed the theory have mostly focused their research on Fizeau interferometers (Hibino, De Groot). In a Fizeau interferometer, errors caused by harmonics are common and can lead to additional errors caused from the coupling of the harmonics with different error sources. This was a driving force for the design of new and better algorithms. In a Twyman-Green interferometer, on the other hand, harmonics generally do not present a problem, especially when considering that the lenses/objectives under test usually have anti-reflective coatings applied to them. Nevertheless, the theory that was developed can be particularly useful in the selection of a phase stepping algorithm because it encompasses the other error sources as well. Moreover, the effect of harmonics in a Twyman-Green interferometer hasn't been studied extensively and it is worth comparing the results for the calculated phase, obtained by different PSAs also in terms of their harmonic attenuating abilities.

Before we can describe the technique to analyze PSAs it is first necessary to derive a more general expression for the intensity profile. The formula of Eq.(7) is not a full representation of the intensity profile $I(x, y)$ but an approximation that is very often found in textbooks. The simplified expression shows the intensity distribution for the fundamental frequency of the modulation and therefore the effects of harmonics cannot be taken into account when applying it. The ability of an algorithm to compensate for the error caused by harmonics is a driving force for the design of newer and better algorithms and for that reason it will benefit us to look at a more general expression shown below.

$$I_r(x, y) = I'(x, y) + \sum_{n=1}^{\infty} I_n''(x, y) \cos [\phi_n(x, y) + n\delta_r] \tag{11}$$

Here $I_r(x, y)$ represents the intensity profile for the rth interferogram, $I'(x, y)$ is the DC component as shown in Eq.(6), $I_n''(x, y)$ is the intensity modulation amplitude for a given phase of the nth order harmonics - $\phi_n(x, y)$, n is an integer indicating the harmonic's order and δ_r is the phase step (i.e. 0, $\frac{\pi}{2}$, π, $\frac{3\pi}{2}$) for the rth recorded interferogram.

When comparing Eq.(7) & (11), it is evident that Eq.(7) is the expression as given in Eq.(11) for $\phi_1(x, y)$, which is the fundamental order, or the phase corresponding to the fundamental frequency. The reason why Eq.(7) is commonly used in the literature is that $\phi_1(x, y)$ describes the wavefront shape and is therefore the object of interest, while $\phi_{2,3,4,...}(x, y)$ are higher order harmonics that need to be attenuated, filtered out.

In Eq.(7, 11) the phase step δ_r is shown to take discrete values (i.e. 0, $\frac{\pi}{2}$, π, $\frac{3\pi}{2}$), however, we are actually modulating the signal $I(x, y)$ with a periodic in time function, with a period of T_s. As such an even more general expression can be written as:

$$I_r(x, y) = I'(x, y) + \sum_{n=1}^{\infty} I_n''(x, y) \cos [\phi_n(x, y) + n\delta(t)] \tag{12}$$

The phase step $\delta(t)$ is a periodic-in-time function that can be rewritten as $\delta(t) = 2\pi v_s t$, where v_s is the fundamental frequency of the intensity modulation, equal to $v_s = \frac{1}{T_s}$. For the following analysis we will only consider the time dependance. The formula can now be rewritten as:

$$I(t) = \sum_{n=0}^{\infty} a_n \cos\left[2\pi n v_s t + \phi_n\right] \tag{13}$$

Where a_n is a weight factor identical to the intensity modulation amplitude $I_n''(x, y)$ for the nth harmonic. In the paper of K. Freischlad and C. Koliopoulos [10] they propose an approach for evaluating the phase ϕ_n that can be represented using a heterodyne process. This evaluation is done using Fourier theory and allows for the "Fourier Representation of PSAs".

The idea behind is that the phase ϕ_n can be extracted from the convolutions of the time-periodic signal $I(t)$ with two sampling functions $f_1(t)$ and $f_2(t)$. Additionally, the whole phase stepping algorithm may be visualized in the frequency domain, which is the algorithm assessment tool developed by Larkin and Oreb [11]. The sampling signals $f_1(t)$ and $f_2(t)$ may be thought of as sampling weights for a set of samples – the measured interferograms. The two signals in time-domain are given as:

$$f_1(t) = \sum_{n=0}^{N-1} \alpha_n \delta(t - t_n)$$
$$f_2(t) = \sum_{n=0}^{N-1} \beta_n \delta(t - t_n) \tag{14}$$

Here, $\delta(t)$ is the Dirac-delta function and α_n, β_n are real coefficients and t_n are sample positions. The sample positions are given as:

$$t_n = \frac{nT_s}{N-1} - \frac{T_s}{2} \tag{15}$$

Where N is the number of steps used in the algorithm, T_s is the fundamental period as described earlier and is usually equal to 2π, and $n \in [0; N-1]$ is an integer.

If we look at the formula for the 5-step Hariharan algorithm as an example, we can see that the expression for the phase is given as:

$$\phi(x, y) = \arctan\left[\frac{2(I_1 - I_3)}{2I_2 - I_0 - I_4}\right] = \arctan\left[\frac{I_1 - I_3}{I_2 - \frac{1}{2}I_0 - \frac{1}{2}I_4}\right] \tag{16}$$

The coefficients α_n, β_n correspond to the normalized sampling weights, therefore for this particular case:

$$\alpha_0 = 0 \quad \alpha_1 = 1 \quad \alpha_2 = 0 \quad \alpha_3 = -1 \quad \alpha_4 = 0$$
$$\beta_0 = -\frac{1}{2} \quad \beta_1 = 0 \quad \beta_2 = 1 \quad \beta_3 = 0 \quad \beta_4 = -\frac{1}{2}$$

Therefore, a general expression for the phase as provided by Kim [7] can be written as:

$$\phi(x,y) = \arctan \frac{\sum_{n=0}^{N-1} \alpha_n I_n}{\sum_{n=0}^{N-1} \beta_n I_n} = \arctan \frac{\int_{-\infty}^{\infty} F_1(\nu) J(\nu) d\nu}{\int_{-\infty}^{\infty} F_2(\nu) J(\nu) d\nu} \tag{17}$$

Where $F_1(\nu)$, $F_2(\nu)$ are the Fourier transforms of $f_1(t)$ and $f_2(t)$, and $J(\nu)$ is the Fourier transform of $I(t)$ as described in Eq.(13). The Fourier transforms of $f_1(t)$ and $f_2(t)$ are given as:

$$F_1(\nu) = \sum_{n=0}^{M} \alpha_n e^{-i\delta_n \nu} = -i \sum_{n=0}^{M} \alpha_n \sin(2\pi \nu t_n)$$
$$F_2(\nu) = \sum_{n=0}^{M} \beta_n e^{-i\delta_n \nu} = \sum_{n=0}^{M} \beta_n \cos(2\pi \nu t_n) \tag{18}$$

The expression on the left represents the pure Fourier transform of the sampling functions, while the expression on the right is the simplified expression. The reason for the simplification is that the sampling function $f_1(t)$ is real and odd, while $f_2(t)$ is real and even. It follows from this that $F_1(\nu)$ is imaginary and odd, while $F_2(\nu)$ is real and even. Here δ_n is the phase step over the summation range M, where M is given as:

$$M = \frac{N-1}{2}, \ for \ N \ odd$$
$$M = \frac{N-2}{2}, \ for \ N \ even \tag{19}$$

The reason for this is that a heterodyne process requires for the time-periodic signal $I(t)$ to be mixed with two sinusoidal signals that are 90° out of phase – the sampling functions. Therefore the properties of $F_1(\nu)$, $F_2(\nu)$ stem from the properties of the sampling functions in time-domain. For example, for the five-step algorithm described by Eq.(16), using the expressions given by Eq.(14,15), we can write the sampling functions as:

$$f_1(t) = \delta(t + \frac{\pi}{2}) - \delta(t - \frac{\pi}{2})$$
$$f_2(t) = -\frac{1}{2}\delta(t + \pi) + \delta(t) - \frac{1}{2}\delta(t - \pi) \tag{20}$$

The sampling period T_s was taken as $T_s = 2\pi$. The coefficients α_n, β_n can either be determined by looking directly at the expression for the PSA (i.e. Eq.(9, 16)), or by applying a least-squares-fit as described by Greivenkamp and Morgan [12], [13], to N samples over a period $T_s = 2\pi$

19

and acquiring the expressions given in Table 2. Beware that the expressions vary depending on whether the algorithm has an even or odd number of steps N.

Table 2: Sampling weights for N-step PSAs with phase steps over a 2π range

N is even	N is odd
$\alpha_n = \sin(\frac{2\pi n}{N}), N \in [0; N-1]$	$\alpha_n = \sin(\frac{2\pi n}{N-1}), N \in [0; N-1]$
$\beta_n = -\cos(\frac{2\pi n}{N}), N \in [0; N-1]$	$\beta_n = -\cos(\frac{2\pi n}{N-1}), N \in [1; N-2], \beta_0 = \beta_{N-1} = -\frac{1}{2}$

The sampling weights α_n, β_n, number of steps N, and modulation period T_s is the required information to construct $F_1(v), F_2(v)$. By plotting the two functions versus the fundamental frequency v_s, we obtain a powerful tool for the assessment of PSAs. We can plot the frequency spectrum of the sampling functions $f_1(t)$ and $f_2(t)$ for various algorithms and extract useful information regarding the behavior of the said algorithm. Examples of the behavior of a 6-step algorithm and a 19-step algorithm are provided in Fig.(9).
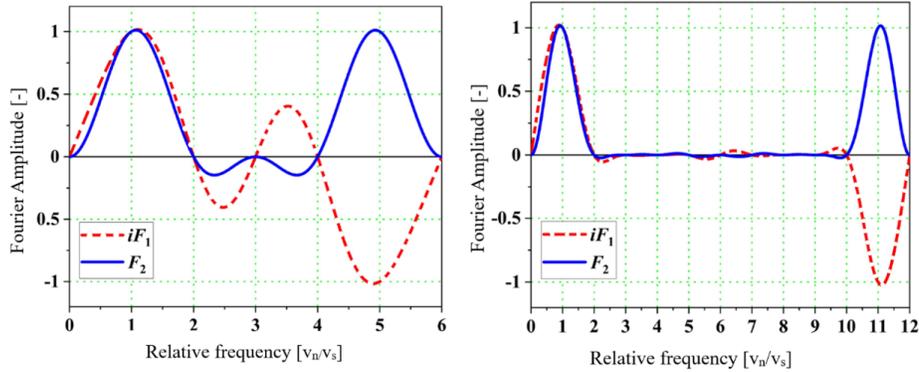

Figure 9: left – 6 step algorithm as proposed by [11], right 19-step algorithm as proposed by [14], [15], source: [7].

The horizontal axis represents the frequency in terms of the fundamental frequency v_s. Both $F_1(v)$, $F_2(v)$ are periodic functions with a period of $\frac{N-1}{T_s}$ for N odd, and a period of $\frac{2(N-1)}{T_s}$ for N even.

These plots carry a lot of useful information for the error-compensation abilities of the algorithm at hand. For example, for the algorithm on the left we can see that the Fourier amplitudes at the 2nd, 3rd, and 4th harmonic are effectively 0. This means that the algorithm can filter out these harmonics. The 19-sample algorithm on the right performs much better in that regard, where the harmonics up-to and including the 10th harmonic are filtered. The height of the sidelobes also plays a very important role as the gradient of the Fourier amplitude is related to errors such as coupling. An overview of some common relationships can be found in Table 3.

Table 3: Conditions for the suppression of certain error types (PSE = Phase Shift Error)

Conditions	At $\frac{v}{v_s} = 1$	At Higher Harmonics $\frac{v}{v_s}$
$iF_1(v_s) = F_2(v_s)$	Absolute Requirement	N/A
$|F_1(v)| = |F_2(v)|$	N/A	Linear PSEs
$iF_1(nv_s) = F_2(nv_s) = 0,\ n \neq 1$	N/A	n-th harmonic
$\frac{\partial iF_1(v)}{\partial v} = \frac{\partial F_2(v)}{\partial v}$	Linear PSEs	Coupling error
$\frac{\partial^2 iF_1(v)}{\partial v^2} = \frac{\partial^2 F_2(v)}{\partial v^2}$	Nonlinear PSEs	Nonlinear PSEs
$iF_1(v_s \pm \varepsilon) = F_2(v_s \pm \varepsilon) = 1$	DC error	N/A

Table 3, however, is not complete and consists of loose information extracted from several papers - [7], [10], [11]. More detailed relationships between the frequency spectrum and the error-compensating ability of algorithms would be a useful addition to the analysis of PSAs.

4.2.2 Error-Compensating Capabilities of Phase Stepping Algorithms

A convenient way to categorize PSAs is by arranging them based on the errors the algorithms can compensate for as in the paper of S. Kim et al. [7]. However, it is first necessary to explain the different errors that are present and can affect a measurement. There are five error sources a PSA can compensate for – harmonics, linear phase-shift error, nonlinear phase-shift error, coupling, and DC error.

In a perfect interferometer, mirrors have perfect reflectivity and optical transmission components also transmit light perfectly. In the real world this is often not the case and "ghost" beams that are multiply reflected can give rise to spurious fringes like those seen in a Fizeau interferometer [8].

Nonlinear phase-shift errors, on the other hand, are usually caused by external vibrations from the environment, while linear errors originate from miscalibrations, which is identical to very low frequency vibration [9]. As seen in the simple 4-step algorithm, the expression for the phase $\phi(x, y)$ relies on the fact that the introduced phase step, $\delta(t)$, indeed varies by $\frac{\pi}{2}$. External vibrations may change the OPD and introduce a small error in the phase step between adjacent measurements of the intensity (recorded interferograms), meaning that the phase step between I_0 and I_1, for example, may become $\frac{\pi}{2} \pm \epsilon$, where ϵ is the error.

Harmonics and the phase-shift errors, when both present induce a new error, called "coupling", which can be a reason for concern. However, more robust algorithms that involve more steps compensate good for harmonics, which prevents the occurrence of coupling.

The DC error is simple to comprehend as it relates to variations in the DC component, $I\prime(x, y)$, in the recorded interferograms. These variations may be caused by different sources including instability of the light source used, in our case – a diode laser. Depending on how the algorithms compensate for these errors, 7 main categories can be distinguished, which are shown in Table 4.

Table 4: Categorization of groups according to error compensation ability [7]

Group	Compensation: Harmonics	Lin. PS Error	Nonlin. PS Error	Coupling Error	DC Error
I	A	N/A	N/A	N/A	N/A
II	A	A	N/A	N/A	N/A
III	A	A	A	N/A	N/A
IV	A	A	N/A	A	N/A
V	A	A	A	A	N/A
VI	A	A	A	N/A	A
VII	A	A	A	A	A

Table 5: Categorization of phase-shifting algorithms [7]

Group	PSA	Design Method
I	Synchronous Detection	Fourier Description
	Wyant 4-sample	Trigonometric Function
	Wyant 3-sample	Trigonometric Function
II	Hariharan 5-sample	Averaging Theory
	Larkin-Oreb N+1	Fourier Description
	Zhu 9-sample	Linear Equation
III	Schmit and Creath 6-sample	Averaging Theory
	De Groot 7-sample	Data-Sampling Window
	Fang 11-sample	Linear Equation
IV	Hibino 7-sample	Linear Equation
	Surrel 2N-1	Characteristic Polynomial
	Hibino 19-sample	Fourier Description
	Hanayama 2N-1	Characteristic Polynomial
	Estrada 9-sample	Fourier Description
	Jeon 11-sample	Data-Sampling Window
V	Hibino 9-sample	Linear Equation
	Zhang 8-sample	Averaging Theory
	De Groot 13-sample	Data-Sampling Window
	Wu 10-sample	Averaging Theory
	Shi 13-sample	Data-Sampling Window
	Kim 3N-2	Characteristic Polynomial
	Kumagai 13-sample	Data-Sampling Window
	Yu 13-sample	Data-Sampling Window
	Choque 9-sample	Fourier Description
	Padilla C(N-1)-1	Fourier Description
VI	Kim 9-sample	Linear Equation
	Choque 8-sample	Fourier Description
VII	Kim 13-sample	Linear Equation
	Bae 19-sample	Linear Equation
	Kim 15-sample	Linear Equation
	Kim 4N-1	Characteristic Polynomial

Table 4 and Table 5 provide a detailed overview of many algorithms and their error-compensating properties, however, it will be very useful to have a method that can be used to analyze the behavior of PSAs. In the literature such methods have been designed with the purpose to assist in their design but as the purpose of this project is not the creation of a new algorithm but rather the error analysis thereof, the methods will be used to help us select an appropriate algorithm which compensates for errors as much as possible. One such method has been described by K. Freischlad and C. Koliopoulos [10] and later used by K. Larkin and B. Oreb [11]. It stems from the Fourier representation of PSAs.

4.3 Phase Unwrapping

The goal of optical metrology is to convert the conducted measurements, in our case interferograms, into useful information – displacements, strains, refractive indices, etc., often through the process of phase retrieval. This is an inverse problem that is ill-posed in many cases, one of which is radar imaging. Here, the height-errors of the aberrated wavefront represent the surface topography of the terrain and a large amount of noise, fringes and discontinuities are present in each recorded interferogram. However, in the case of phase stepping interferometry, where multiple interferograms with varying phase step are recorded in a controlled environment, the problem of phase retrieval becomes well-posed for 3 recorded interferograms, and over-defined in the case of more recorded interferograms with varying phase step.

This was discussed in 4.2, where it was shown by Eq.(7) that there are three unknowns $I^{'}(x, y)$, $I^{''}(x, y)$, $\phi(x, y)$ and each recorded interferogram represents a linear equation, where for a system of linear equations the minimum number of equations required to solve for 3 unknowns is 3 equations. Hence, more than 3 recorded interferograms have the effect of overdetermining the phase retrieval problem. This is a good thing as ill-posed problems require elaborate phase unwrapping algorithms and techniques to extract the phase map from the interferogram.

For that purpose, the major scientific advances in the field of phase unwrapping algorithms have been focused on designing algorithms that for example deal with the common problems in "Synthetic Aperture Radar" (SAR) imaging, rather than interferometry for optical testing. In our case, the classical phase unwrapping algorithms are more than sufficient and more complex ones would be unnecessary as they don't show any advantages in terms of accuracy, speed, or ease-of-use [20]. Nevertheless, phase unwrapping is an important step as incorrectly unwrapped phase map would represent an incorrect wavefront with incorrect aberrations present in it. Therefore, we will look at the more conventional approaches for phase unwrapping, their strengths and weaknesses.

In the previous section it was shown how the raw phase map can be calculated from multiple interferograms. If we observe Eq.(9), there is a significant problem that immediately occurs and is present in all PSAs regardless – the phase is calculated by an arctangent function.

$$\phi(x, y) = arctan \left[\frac{I_3 - I_1}{I_0 - I_2} \right]$$

The range of the principal values of the arctangent function is between $[\frac{-\pi}{2}; \frac{\pi}{2}]$, which introduces discontinuities in the calculated phase. We can use the properties of the tangent and arctangent and increase the length of the range to $[-\pi; \pi]$ using "Modulo 2π Phase Correction" [21], [22], or $[0; 2\pi]$ as done by other authors [3]. This comes from the fact that a tangent is the ratio of a sine and a cosine, where the signs of the sine and cosine functions are usually unknown. This is not true in our case, and we can use the knowledge of their signs to increase the range of the arctangent. However, this technique can only lower the number of discontinuities, not remove them altogether. The principle and conditions for Modulo 2π Phase Correction

are given in Eq. (21, 22) and Table 6. The table uses the range $[0;2\pi]$ as shown in the book of Malacara.

$$tan[\phi(x,y)] = \frac{sin[\phi(x,y)]}{cos[\phi(x,y)]} = \frac{\sum_{n=0}^{N-1} \alpha_n I_n}{\sum_{n=0}^{N-1} \beta_n I_n} \tag{21}$$

$$\therefore \sum_{n=0}^{N-1} \alpha_n I_n \propto \sin\left[\phi(x,y)\right] \qquad and \qquad \sum_{n=0}^{N-1} \beta_n I_n \propto \cos\left[\phi(x,y)\right] \tag{22}$$

Table 6: Modulo 2π Phase Correction [3]

Sine	Cosine	Corrected Phase $\phi(x,y)$	Phase Range
0	+	0	0
+	+	$\phi(x,y)$	0 to $\frac{\pi}{2}$
+	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$
+	−	$\phi(x,y) + \pi$	$\frac{\pi}{2}$ to π
0	−	π	π
−	−	$\phi(x,y) + \pi$	π to $\frac{3\pi}{2}$
−	0	$\frac{3\pi}{2}$	$\frac{3\pi}{2}$
−	+	$\phi(x,y) + 2\pi$	$\frac{3\pi}{2}$ to 2π

We call the calculated phase in the $[-\pi; \pi]$ range (or $[0; 2\pi]$) "wrapped". There is a need for an algorithm that can locate the discontinuities at the edges of the range and create a smooth phase surface. For example, assume we are moving along a certain direction on the raw phase map and the phase gradually builds up, with small increments, to the value of 2π at a certain pixel, the phase at the next pixel will drop down to 0 which is incorrect in the physical sense that a phase map/wavefront is a smooth continuous surface. Instead, the phase at the next pixel should be $2\pi + \eta$, where η is the small increment. To get around this problem, multiple methods for phase unwrapping have been devised. Based on their nature, they can be grouped into three categories – "sequential methods", "residues methods", and "least squares methods [23].

4.3.1 The Phase Unwrapping Problem on a Surface

In this section we will look at the mathematical definition of the phase unwrapping problem on a 2D surface. The problem of phase unwrapping has been discussed by various authors, however, in the paper of Shuvolov [21], a very clear general definition has been presented which will be shown here.

Let $\psi(x,y) \in \mathbb{R}$ be the absolute phase in radians that we would like to obtain and let it be a scalar piecewise-continuous function of the position arguments x and y. The CMOS sensor measures the irradiance over the exit pupil of the test lens, hence the absolute phase, corresponding to the measured irradiance must be defined over a closed domain $\Omega \subseteq \mathbb{R}^2$. The surface spanned by $\psi(x,y)$ is called the surface of absolute phase.

Let $W[.]$ denote the modulo 2π wrapping operator, which wraps the value of the absolute phase $\psi(x, y)$ between $[0; 2\pi]$ for every point $(x, y) \in \Omega$. Where we define a relative phase $\varphi(x, y)$, as directly obtained by the arctangent function.

$$W[\psi(x, y)] = \phi(x, y) \quad for \quad (x, y) \in \Omega \tag{23}$$

Interferograms usually consist of discrete data, where the measured irradiances per pixel are stored in an array, therefore, it is convenient to use discrete notation for the following operations. We can rewrite $\psi(x, y)$ and $\phi(x, y)$ as:

$$\Psi = \{\psi_{m,n}|_{m=0,...,M-1;n=0,...,N-1}\} \quad \Phi = \{\phi_{m,n}|_{m=0,...,M-1;n=0,...,N-1}\} \tag{24}$$

Here Ψ, Φ are arrays containing the values per pixel for the absolute and relative phase. The subscripts m, n correspond to the m-th horizontal and n-th vertical pixel, where the CMOS sensor has a total of M horizontal and N vertical pixels. It follows:

$$W[\psi_{m,n}] = \phi_{m,n} \quad for \quad m = 0, ..., M-1; \ n = 0, ..., N-1 \tag{25}$$

The absolute phase $\psi_{m,n}$ satisfies Eq.(25) if and only if the condition for congruence is satisfied. That is:

$$\psi_{m,n} = \phi_{m,n} + 2\pi \ k_{m,n}, \quad k_{m,n} \in \mathbb{Z} \tag{26}$$

If no boundary conditions are introduced, Eq.(25) has infinitely many solutions. This can be overcome by making use of the Nyquist sampling theorem. Here, the absolute difference of the phase between adjacent vertical or horizontal pixels must not exceed π radians.

$$|\psi_{m+1,n} - \psi_{m,n}| < \pi \quad and \quad |\psi_{m,n+1} - \psi_{m,n}| < \pi \tag{27}$$

We can denote the difference in the absolute phase between adjacent horizontal or vertical pixels in terms of finite differences as:

$$\Delta_X = \psi_{m+1,n} - \psi_{m,n} \quad and \quad \Delta_Y = \psi_{m,n+1} - \psi_{m,n} \tag{28}$$

And the finite difference in the relative phase between adjacent horizontal or vertical pixels as:

$$\delta_X = W[\phi_{m+1,n} - \phi_{m,n}] \quad and \quad \delta_Y = W[\phi_{m,n+1} - \phi_{m,n}] \tag{29}$$

The condition imposed by the Nyquist sampling theorem can be interpreted as the lack of discontinuities in the interferogram, which has the following consequences for the differences in the phase given by Eq.(28, 29):

$$\Delta_X(m, n) = \delta_X(m, n) \quad and \quad \Delta_Y(m, n) = \delta_Y(m, n) \tag{30}$$

Another boundary condition stems from the definition of $\psi(x, y)$. We mentioned that it is "a scalar piecewise-continuous function" and as such the curl of its gradient must be 0. For a 2D field this can be written as:

$$\nabla\psi = \left[\frac{\partial\psi}{\partial x}; \frac{\partial\psi}{\partial y}\right]^T = \left[\psi_{m+1,n} - \psi_{m,n}; \psi_{m,n+1} - \psi_{m,n}\right]^T \tag{31}$$

$$curl\left(\nabla\psi\right) = \frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial y^2} = 0 \tag{32}$$

Equation (25) can be rewritten in a discrete manner, in terms of the values per pixel:

$$\delta_X(m, n) + \delta_Y(m, n+1) - \delta_X(m+1, n) - \delta_Y(m, n) = 0 \tag{33}$$

Taking into account the various conditions mentioned thus far, the phase unwrapping problem, as defined by Eq.(23) has one unique solution if the phase at a certain starting point (m_0, n_0) is chosen as ψ_0. From the starting point, the phase at any other point can be calculated in terms of a sum of finite differences along a path P, leading to that point:

$$\psi(m, n) = \psi_0(m_0, n_0) + \sum_{i \in P} \Delta_i \tag{34}$$

So far, only the ideal case has been discussed with the two boundary conditions defined by the Nyquist sampling theorem Eq.(27) and the curl of the gradient Eq.(32). However, we must investigate the relationship between them. According to Shuvolov, satisfying Eq.(33) guarantees uniqueness of the solution but doesn't guarantee that Eq.(27) is also satisfied, where Eq.(27) is a requirement for the solution to be true.

In the cases where the interferogram contains linear discontinuities, whose end points lie outside of the interferogram or closed-loop discontinuities inside of it, then Eq.(33) will be satisfied everywhere, while Eq.(27) will be compromised at the points corresponding to the discontinuity. In many practical cases, including lens testing interferometry, the phase discontinuities lead to the so-called "singularity points", which are the endpoints (pixels) of the linear discontinuities where Eq.(33) is not satisfied. In these cases the phase unwrapping problem, as defined by Eq.(23, 27) stops existing. It doesn't present any difficulty to determine the locations of the linear discontinuities, however, they introduce ambiguity, which is related to their position on the interferogram.

The essence of phase unwrapping algorithms is to remove ambiguity in the phase. Ambiguity spans a set of various possible solutions for the phase map $\psi(x, y)$. The phase unwrapping algorithm is essentially a problem for the optimization of a certain functional over the set of all possible solutions. The optimization returns a single solution from the set.

4.3.2 Sequential Methods

Sequential methods rely on scanning. The algorithm starts at a certain point (pixel) and unwraps the phase from there by integrating the wrapped phase map. Depending on how the algorithm moves to the next pixels – integration path, the algorithm may be a "linear scanning", a "spiral scanning", or a "multiple direction scanning" one. There are also sequential algorithms that don't follow a path at all and rather rely on a weighting function that gives information about the reliability of the information for a certain pixel or region [24], [25], [26]. There are two main approaches, including the boundary following and the region growing approach. Both have their strengths and weaknesses.

The boundary following approach tries to find where the modulo 2π boundary is located and adds a multiple of 2π to maintain the phase continuity. The drawback is that these algorithms fail at regions with fast variations of the phase and where the phase information is corrupted by noise and is unreliable [25]. Such regions cause errors that gradually build up over the path of unwrapping.

The region growing approach relies on selected pixels, deemed "seeds", from which the phase map is unwrapped. Each "seed" pixel grows out a region, hence the name [27], [28]. This method suffers from the same drawback as the boundary following approach and cannot compensate for unreliable data. In addition, it is computationally heavier, although that in the case of lens testing, presented in this report, we can tolerate heavier algorithms.

4.3.3 Residues Methods

The residues methods can be viewed as a complementary technique to the sequential methods as they are meant to correct and compensate for the noise that is present in the raw phase map. Noise, as well as sampling errors, cause sinks and sources ("singularity points") to form in the map, known as "rotational residues". In the mathematical sense these residues occur when Eq.(32) is not satisfied, meaning that the $\nabla \psi(x, y)$ field has a non-zero rotational component. Residue methods strive to eliminate the singularity points, so that the raw phase may be unwrapped using any of the previously explained sequential methods. In the paper of Flynn [24] a useful expression for the residues map $r_{m,n}$ in terms of the relative phase $\phi_{m,n}$ has been given:

$$
\begin{aligned}
r_{m,n} = \frac{1}{2\pi} &\left[W\left[\varphi_{m+1,n} - \varphi_{m,n}\right] + W\left[\varphi_{m+1,n+1} - \varphi_{m+1,n}\right] \right] - \\
&- \frac{1}{2\pi} \left[W\left[\varphi_{m,n+1} - \varphi_{m,n}\right] + W\left[\varphi_{m+1,n+1} - \varphi_{m,n+1}\right] \right]
\end{aligned}
\tag{35}
$$

4.3.4 Least Squares Methods

Least squares methods integrate the whole field, instead of having to follow a certain path or use weighting function as in the sequential methods. There are two main types of least squares unwrapping problems – weighted and unweighted, both of which amount to solving partial differential equations in discrete form using different mathematics techniques. The idea behind is

28

that the phase differences for the absolute phase Δ_X, Δ_Y must agree with the phase differences for the relative phase δ_X, δ_Y as it was shown in Eq.(30). This is done using least squares, where the parameter ε is to be minimized for a choice of $\psi_{m,n}$.

$$\varepsilon^2 = \sum_{m}^{M-2} \sum_{n}^{N-1} (\Delta_X - \delta_X)^2 + \sum_{m}^{M-1} \sum_{n}^{N-2} (\Delta_Y - \delta_Y)^2 \tag{36}$$

Some examples include "Discrete Cosine Transform", "Finite element", and "Minimum L^0 norm" [29], [30], [31].

4.4 Zernike Polynomial Fitting

The OIS we are interested in, for this project, are circular symmetric, therefore only Zernike polynomials will be considered as they are orthogonal over the unit circle. The weighting factors per polynomial are called Zernike coefficients, and can tell us what and how much of a certain aberration is present in the wavefront. To calculate them, we must first solve an optimization problem. That is – the Zernike coefficients must be chosen such that we minimize the error (difference) between the actual wavefront, as given by the phase unwrapping algorithm, and the wavefront we construct by choosing the weighting factors (Zernike coefficients) for the individual polynomials and summing them up. This step is called "Zernike polynomial fitting". The formula for calculating the Zernike polynomials in terms of polar coordinates is given in Eq.(37).

$$
\begin{aligned}
Z_{n,m}(\rho,\theta) &= R_{n,m}(\rho) & for \quad m &= 0 \\
Z_{n,m}(\rho,\theta) &= R_{n,m}(\rho)\sin(m\theta) & for \quad m &< 0 \\
Z_{n,m}(\rho,\theta) &= R_{n,m}(\rho)\cos(m\theta) & for \quad m &> 0
\end{aligned}
\tag{37}
$$

Where $R_{n,m}(\rho)$ denotes the radial function and is given by:

$$R_{n,m}(\rho) = \sum_{s=0}^{\frac{n-m}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} \rho^{n-2s} \tag{38}$$

As with any other algorithm, there is always a small numerical error as the polynomials are not orthogonal over a discrete set of data points, such as the values per pixel as measured by the CMOS sensor. Another important limitation is the maximum number of polynomials that can be fit, as it is impossible to fit an infinite number of polynomials. This is due to the measurement data being discrete and aliasing becoming a problem when high order terms are fit. For example, the wavefront deviations caused by the production process of spherical cavities using "Single Point Diamond Turning" [33], cannot be represented with a reasonable number of

Zernike polynomials. Because of that, there are residual terms that are not present in the re-constructed wavefront obtained from fitting a finite number of polynomials. For ease of use we will refer to this wavefront as the "fitted wavefront" (FW). The different polynomial fitting techniques will be presented below.

4.4.1 Least Squares Fitting

A "Least Squares Fit" (LSF), is a common approach used in commercial software as it is easy to implement, and the calculations are executed using arrays (vectors and matrices) which is computationally efficient for the computer. As previously mentioned, we determine the wavefront aberrations by expanding the wavefront in terms of a set of polynomials, orthogonal over the pupil. The expansion coefficients (i.e. Zernike coefficients) can be determined by a minimization problem. That is, a finite number of Zernike polynomials are used to estimate the wavefront and the difference between the measured wavefront and the estimation is minimized by solving a linear system of equations. The procedure for fitting Zernike polynomials using Least squares is explained in the book of Mahajan [33] and is easy and straightforward to implement.

The wavefront aberration function $W(x, y)$, that is – our aberrated wavefront as obtained by the recorded interferograms, phase unwrapping, and scaling, can be expressed as a weighted sum of J number of Zernike polynomials.

$$W\left(x, y\right) = \sum_{j=1}^{J} a_j Z_j(x, y) \tag{39}$$

Here, $Z_j(x, y)$ is a scalar, which represents the jth Zernike polynomial evaluated at the location (x, y) and a_j is also a scalar, which is the Zernike coefficient. The variance of the aberration function is given as:

$$\sigma^2 = \sum_{j=5}^{J} (N_j a_j)^2 \tag{40}$$

Where N_j is a normalization coefficient, also called the RMS factor. It is important to note that in Eq.(40) the first 4 Zernike terms, corresponding to piston, x-tilt, y-tilt, and defocus, can be omitted. In an interferometer, these terms do not represent properties of the lens but instead arise from the alignment of the interferometer. As such, they are not considered in the calculation of the variance.

In a computer, the measured values per pixel are stored as data points in an array. For a 2D interferogram the stored array is also two-dimensional and can be converted into a column vector, containing all data points using the "flatten" function, depending on the programming language used. We will denote this column vector as D_{Nx1}, where N is the total number of data points. The column vector containing the estimated Zernike coefficients is denoted by $\hat{a}_{Jx1}$, where J is the number of Zernike polynomials. As indicated in Eq.(39), we need to evaluate all

Zernike polynomials at all data points, hence the need for a matrix that stores the calculated values – A_{NxJ}. Then we can write:

$$D_{Nx1} = A_{NxJ} \hat{a}_{Jx1} \tag{41}$$

To find an estimation for the Zernike coefficients we can multiply both sides by the inverse of matrix A_{NxJ}, however, the matrix is not square-invertible. This can be overcome by taking the pseudoinverse of it to obtain an expression for $\hat{a}_{Jx1}$.

$$\hat{a}_{Jx1} = \left(A_{NxJ}{}^T A_{NxJ} \right)^{-1} A_{NxJ}{}^T D_{Nx1} \tag{42}$$

A common approach to determine the quality of the wavefront fit is to look at the root-mean square difference between the fitted wavefront $\hat{W}(x, y)$ and the measured wavefront $W(x, y)$. For a perfectly estimated wavefront, the difference is zero.

$$Q = \frac{1}{N} \left\{ \sum_{i=1}^{N} \left[\hat{W}\left(x_i, y_i\right) - W\left(x_i, y_i\right) \right]^2 \right\}^{1/2} \tag{43}$$

4.4.2 Numerical Integration

Another approach for obtaining the Zernike expansion coefficients that is commonly found in literature is using integration [3], [33]. For a continuous set of data, the individual expansion coefficients can be calculated by integrating the measured wavefront, multiplied by the corresponding Zernike polynomial over the whole unit circle. This method is more computationally demanding and slower, however, it has certain benefits that will later be explained. The Zernike coefficients $a_{n,m}$ correspond to the $Z_{n,m}$ Zernike polynomial and measured wavefront W at the point (ρ, θ).

$$a_{n,m} = \frac{2(n+1)}{(1 + \delta_{m0})\pi} \int_0^{2\pi} \int_0^1 W\left(\rho, \theta\right) Z_{n,m}\left(\rho, \theta\right) \rho \, d\rho \, d\theta \tag{44}$$

As the Zernike polynomials are used for circular apertures, the formulas are often given in polar coordinates. The n, m notation follows from Born & Wolf [34], however, attention has to be paid as different authors tend to use different notations for the Zernike polynomials. The δ_{m0} function in Eq.(44) is the Kronecker delta function, which can be expressed as:

$$\delta_{m0} = \begin{cases} 1, & for \quad m = 0 \\ 0, & for \quad m \neq 0 \end{cases} \tag{45}$$

4.4.3   Comparison between Least Squares Fitting and Numerical Integration

According to Mahajan [33], the difference between the two methods is most noticeable for a small number of data points, and in the presence of noise. For a number of data points of up to a few hundred, without noise, the integration method yields better results compared to LSF, due to coupling between lower and higher-order modes in the LSF method. If, however, noise is present in the data set, the roles reverse. The integration method results worsen as noise increases, meaning that LSF performs better in that case. For a large number of data points, both methods perform approximately the same. The numerical error introduced by polynomial fitting, considering the other two previously-discussed algorithms – PSA and Phase Unwrapping, is expected to be the smallest and least relevant.

5   Foundations for Measurement Uncertainty Analysis

In the previous chapter, we examined the fundamental theory of interferometric measurements. This included recording interferograms and obtaining interpretable results using Zernike polynomial coefficients. However, each step in the process is subject to some error, which contributes to the total measurement uncertainty of the system.

This chapter discusses the various sources of error in the measurement process. These errors can arise from the computational aspect of the process, as well as from environmental and instrumentation factors. We will refer to these combined factors as "physical".

5.1   Measurement Uncertainty: Computational Aspects

There is a very strong relationship between the data acquisition and processing of the recorded interferometric data and the environmental and instrumentation factors that influence the measurements. In section 4.2.2 it was shown that PSAs require a certain number of recorded interferograms and may have the capabilities to compensate errors. These errors inevitably introduce uncertainties in the recorded data and subsequently – in the calculated Zernike coefficients.

5.1.1   Issues in Phase Unwrapping

In section 4.3, the general theory of phase unwrapping was presented. The seemingly trivial problem of locating the discontinuities originating from the principal value range of the arctangent function is perhaps one of the faster developing fields related to interferometry.

This section provides a brief overview of common issues encountered during phase unwrapping and reiterates the importance of robust phase unwrapping algorithms. It should be stated once more that the issue of phase unwrapping is not exclusive to the method used to determine the phase, but rather it is an inherent problem in numerically determining it. We'll begin

by discussing the fundamental principles of a phase unwrapping algorithm and why the basic example of 1D phase unwrapping, commonly found in textbooks, is not practical in real-world applications.

In almost every textbook that covers interferometry, phase unwrapping is usually explained in terms of a one dimensional graph, a cross-section of the wrapped phase, where the range of the values is the principal value range of the arctangent $[-\pi; +\pi]$. The most prevalent explanation is that the phase unwrapping algorithm identifies abrupt discontinuities in the wrapped phase and corrects for them by aligning the smooth regions, as demonstrated in the example in Fig.(10 - 13).
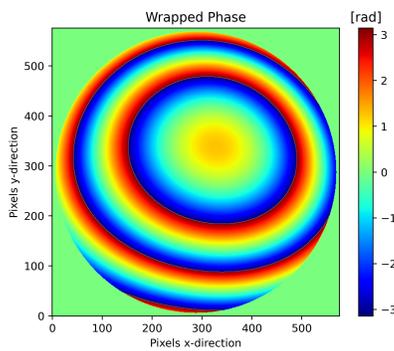
Figure 10: The wrapped phase obtained from computer-generated interferograms.
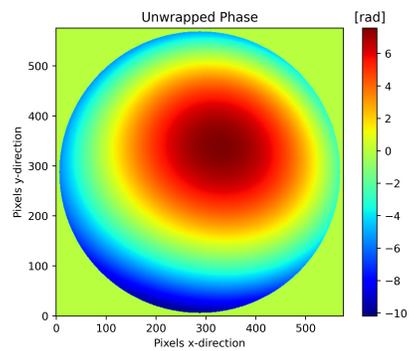
Figure 11: The unwrapped phase for the same example.
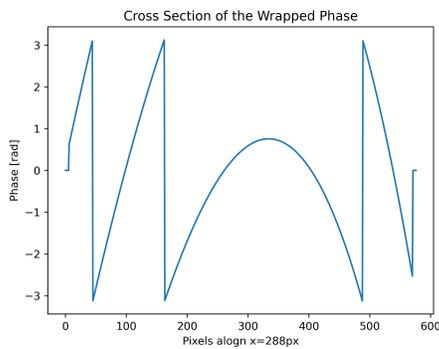
Figure 12: Cross-section of the wrapped phase wrapped between the principal value range of the arctangent
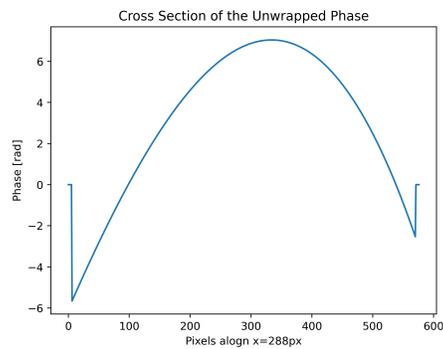
Figure 13: The unwrapped phase of the same cross-section

Here, an ideal case has been shown by analyzing computer-generated interferograms. The simplest example of a phase unwrapping algorithm unwraps both row- and column-wise to correct for the discontinuities. This algorithm, however, is extremely vulnerable to noise and errors as it does not have any built-in defenses to counter their effects. Suppose the phase contains a

simple noise spike. In the previously-mentioned example, the algorithm would mistakenly consider this point for a phase wrap and deliver incorrect result. However, the unwrapped phase at all consecutive points from there on is also going to be incorrect, ensuring an incorrect end result.

Therefore, the purpose of PU algorithms is to be insensitive to such disturbances, recognize them and correct for them. The 4 most common errors in phase unwrapping will be laid out.

1. Noise in the phase data: Even in a state-of-the-art laboratory environment, some amount of noise in the interferograms is unavoidable. In the case of Gaussian noise, the subsequently obtained phase data appears to have a grainy structure, where the boundaries between the wrapped regions appear less defined. This noise causes the aforementioned noise spikes, the effect of which can be detrimental. Therefore, a robust phase unwrapping algorithm may also act as a low-pass filter to an extent depending on the built-in defenses against noise corruption [48].

2. Discontinuities: On instances where the field we are trying to measure contains actual discontinuities, the PU algorithm may fall prey to its own compensation capabilities and incorrectly designate them as errors. This can especially be a problem when testing optical components that contain sharp edges in their surface profile, such as Fresnel lenses and diffraction gratings.

3. Regions of Invalid Data: Interferograms and phase maps with holes and gaps which do not contain any fringe information require algorithms that can adequately correct for them. Such empty regions may be caused by dust particles for example. In chapter 6, it will be shown that the edge detection of the pupil causes such small gaps to appear in areas where the data modulation of the pixels is below a certain threshold.

4. Undersampling: Although the Nyquist sampling condition, as stated in Eq.(27), is a fundamental theoretical limit, undersampling can become an issue before reaching this theoretical Nyquist limit due to the presence of noise. Nevertheless, the high density of pixels on a CMOS sensor should be sufficient for our application.

5.1.2   Polynomial Fit and Discretization Error

The problem of Zernike polynomial fitting over the pupil is that of a function recovery over a circular domain of discrete data points in the form of pixels. There exists extensive literature on the many shortcomings one may encounter in the process, however, this lies outside of the scope of the thesis. For detailed review of the theory on image reconstruction using Zernike polynomials and image moments, the reader is referred to M. Pawlak's book "Image Analysis by Moments: Reconstruction and Computational Aspects" [49]. In this section we will limit the focus to the two predominant errors one might expect when fitting Zernike polynomials over a phase map dataset – geometric and numerical error.

Geometric error refers to the way square pixels are used to represent the unit circle over which polynomials are fit. Constructing a circle out of squares leaves a small area, which either remains empty or extends beyond the circle's border. This error can be decreased by using smaller sized pixels, or otherwise said - increasing the pixel density inside the circle to create a finer grid. Especially in small pixel grids (i.e. 20x20px), it is the predominant error source. An example of the geometric error can be seen in Fig.(14, 15).
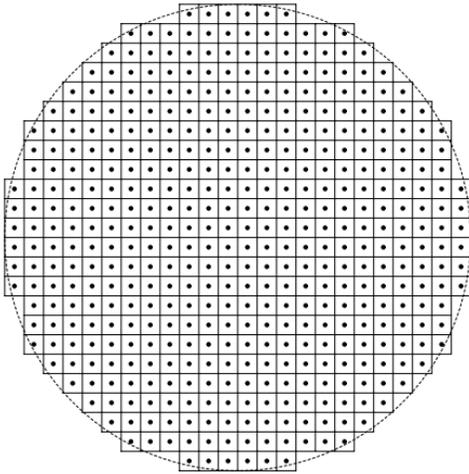
Figure 14: Illustration of the Cartesian pixel grid for calculating Zernike moments
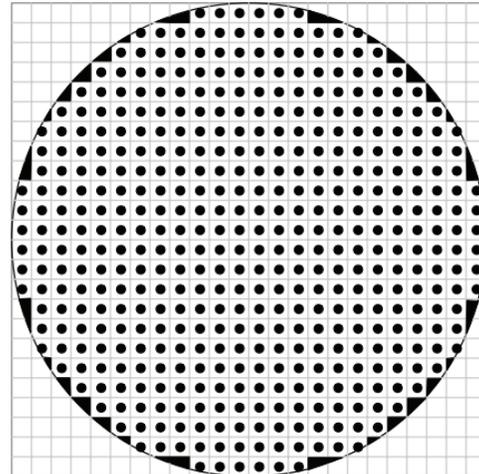
Figure 15: The boundary region defining the geometric error

Numerical error, on the other hand, originates from the need to accurately calculate the double integral used to find the Zernike coefficients, shown in Eq.(44). In the case where, a least squares fit algorithm is used, this translates to the error caused by the finite set of data samples, otherwise known as discretization error. However, the fit error is consistent and systematic, and can therefore be subtracted from the measurements. Nevertheless, there is some uncertainty in the fit coefficients, which can be estimated.

5.2   Measurement Uncertainty: Physical Aspects

Depending on their magnitude it is very possible that they can "throw-off" the measurements by a large amount or even create a situation where due to misalignment there are ghost-images present in the interferogram. Vibrations from the external environment are also a major concern as they introduce errors in the phase of the recorded interferometric images and depending on the choice of algorithms to be used, this may give far-from-realistic results for the Zernike coefficients. In the following subchapters the effects of these individual imperfections will be discussed.

5.2.1 Phase Shifter Errors

In section 4.2.2 we briefly explained how the behavior of different PSAs in the presence of phase-shift errors can be analyzed and how certain algorithms have the property of suppressing these errors. However, the source of such errors is physical as they originate in the phase-shifter. There are several errors related to the phase shifting process when conducting interferometric measurements – "phase-shifter miscalibration", "linear phase-shift error", "nonlinear phase-shift error". As the "phase-shifter miscalibration" and the "nonlinear phase-shift error" are systematic, they require special attention when precise interferometric measurements are required [35].

Errors arising from miscalibration, and nonlinearities can be reduced by simply purchasing a very linear phase-shifter, however, there are different ways that require attention. For example, it is possible to determine the voltage signal supplied to the phase-shifter, for which it provides a linear phase shift. [36], [16]. Another practical way of reducing the errors is by picking a phase shifting algorithm that takes multiple measurements, however, the algorithm in question needs to have this ability.

It is also possible to suppress nonlinear phase-shift errors by adding a linear bias to the movement of the phase-shifter. This can be achieved by making sure that the stepping is over a 2π range and by adding a 2π change in consecutive phase steps [37].

While the phase-shift errors may often present a problem, it is worth nothing that the phase errors introduced are usually very small, especially when phase shift algorithms better than the simple 3- and 4-step are used.

5.2.2 Detector Nonlinearity

Detector nonlinearities arise because there is a nonlinear relationship between the irradiance that is incident on the detector and the outputted voltage. In our case a CMOS detector is used. It is possible to adjust the gain of the camera, so that it will operate in the most linear part of its gain curve. However, it is important to look at the consequences that the nonlinearities in the detected irradiance can have on the measured interferograms. Recall Eq.(7):

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos\left[\delta(t) + \varphi(x, y)\right]$$

It can be rewritten in terms of the fringe contrast $\gamma(x, y)$ and the average intensity $I\prime(x, y)$.

$$I = I'\left[1 + \gamma \cos\left[\varphi + \delta\right]\right] \tag{46}$$

The nonlinear detected irradiance up to and including the second order term can be written as:

$$I_{det} = I + \varepsilon\ I^2 \tag{47}$$

Then if we substitute Eq.(46) into Eq.(47) we get an equation for the detected irradiance in terms of the actual irradiance I and the error coefficient ε. Here, third and higher order nonlinearity terms have been omitted to show the calculation, but they can be considered as well.

$$I_{det} = I'\left(1 + \varepsilon I'\right) + I'\left(1 + 2\varepsilon I'\right)\gamma\cos\left(\varphi + \delta\right) + \frac{\varepsilon}{2}\left(I'\gamma\right)^2\left\{1 + \cos\left[2\left(\varphi + \delta\right)\right]\right\} \qquad (48)$$

Here the $\cos\left[2\left(\varphi + \delta\right)\right]$ term follows from the formula:

$$cos^2\left(\theta\right) = \frac{1 + \cos\left(2\theta\right)}{2}$$

In the paper of Creath [37], she shows the simulated phase error and the peak-to-valley (PV) phase error caused by 10% second and third order nonlinearities in terms of the used phase stepping algorithm.
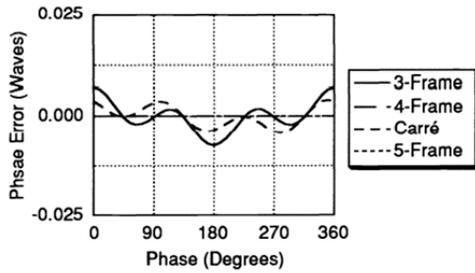
Figure 16: Phase Error 10% 2nd order detector nonlinearity [37]
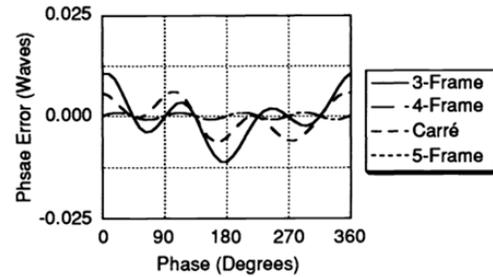
Figure 17: Phase Error 10% 3rd order detector nonlinearity [37]
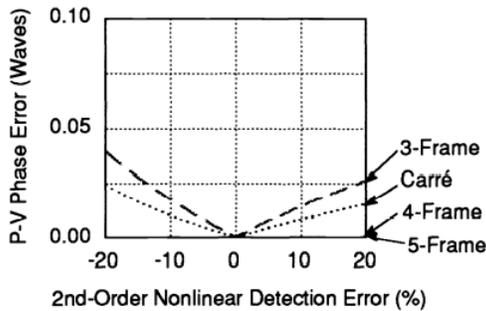
Figure 18: P-V 2nd order nonlinear detector error [37]
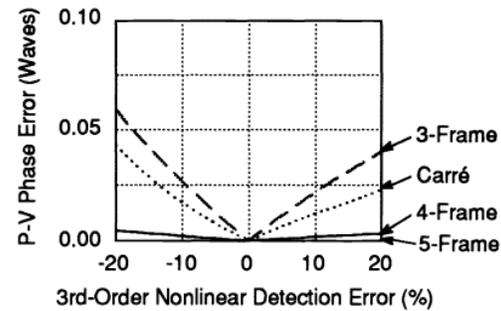
Figure 19: P-V 3rd order nonlinear detector error [37]

From Fig.(16-19) we can see that the phase errors introduced by nonlinearities in the detector for the 3-step algorithm [4] are in the order of a few $m\lambda$. Although, a nonlinearity of 10% is largely exagerated. Nevertheless, we can see that the phase error introduced by nonlinearities drops down fast as an algorithm with more steps is chosen. The Carré [38] and 4-step

algorithm [5] both use phase steps, however, in the algorithm of Carré it is not required for the phase step to be constant, while for the 4-step algorithm the phase step is known and equal to $\frac{\pi}{2}$. This can explain why one outperforms the other.

5.2.3 Camera Quantisation

Camera quantization error refers to the phase error introduced from the analog-to-digital conversion that occurs when the analog signal from the detector is converted into a digital one with discrete intensity levels. In his paper [39], Brophy derives a simple formula for the rms phase error caused by quantizing the irradiance for the family of phase stepping algorithms where the step is equal to $\frac{\pi}{2}$. This can be seen in Eq.(49). Furthermore, he also provides a general expression, from which Eq.(49) is derived, that can be used for analyzing the phase error caused by intensity noise given a model and a choice for PSA.

$$\sigma_\varphi = \frac{1}{\sqrt{3}\gamma K} \tag{49}$$

Here σ_φ denotes the rms phase error, γ – the fringe contrast, and K is the number of quantized irradiance/intensity levels. For example, for an 8-bit camera $K = 2^8 = 256$ discrete irradiance levels, where the lowest recorded irradiance level in the interferogram corresponds to the integer value of 0 and the highest to 255. The fringe contrast also plays an important role, as low fringe contrast increases the rms-phase error.

In the case of our metrology setup, the camera's analog-to-digital converter (ADC) has 12 bits, which, as seen from Fig.(20), exhibits small rms phase error. However, in the data file, containing the discrete irradiance measurements per pixel, the values are in the range $[0; 255]$ which is the range for 8-bits. Additionally, it is worth investigating what the fringe contrast from the recorded interferograms is, as the simulation in Fig.(20) uses maximum fringe contrast, therefore a real-world measurement is expected to have a lower value for γ and thus a higher rms phase error. Nevertheless, it is expected that the contribution of this error source to the overall uncertainty of the interferometric measurements should be small if not negligible.
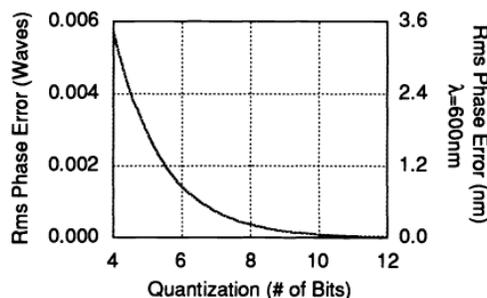

Figure 20: RMS phase error due to intensity quantization for $\gamma = 1$ [37]

5.2.4   Camera Noise

To acquire good optical metrology data, it is important to consider how different noise sources affect the measurements. When considering the output signal of the CMOS sensor, the goal is to achieve the best signal-to-noise ratio (SNR) possible. In the case of our metrology setup, two different cameras are used – "IDS UI-3060SE-M-GL" and "SWIR1300KMA". Their respective sensors are the Sony IMX174LQJ-C, and the Sony IMX990. The larger the output signal, the higher the SNR, however, there are certain limitations to that. To examine the different noise sources, we must first consider the measurement principle of the sensor.

In linear array detectors, incoming light falls on the many pixels of the sensor, called photosites. The falling photons excite electrons on the photosite and over a short period of time current is accumulated (integrated). This period is called the integration time. Once the integration time has passed, the pixel is emptied of the accumulated current during its read-out. However, this process is not ideal, and several factors have to be taken into account, such as "photosite saturation", "dark current", "quantum efficiency", "conversion efficiency", etc.

Every pixel has a maximum amount of charge it can store. This is called the "full well capacity". An absorbed photon generates an electron on the pixel. Once the charge storage capacity of it has been reached, the photon can no longer generate an electron. This occurs in the case of too many photons falling on the surface. A different phenomenon happens in the instances of too little photons – "shot noise", where the number of photons arriving on a pixel follows a Poisson distribution.

Not every photon manages to excite an electron. This is called "quantum efficiency" – Q, and is usually provided in the datasheet for the sensor. For example, a quantum efficiency of 75% means that only 75% of the fallen electrons do generate an electron. A similar parameter to keep in mind is the conversion efficiency α. It relates to the percentage of accumulated electrons that are converted to digital counts when reading-out the photosite.

An interesting phenomenon is the "dark current". It occurs from thermally excited electrons. Therefore, even in the absence of light current will be accumulated on the photosite. The amount depends on the temperature. A measurement of the current in a pixel therefore consists of the following components [40] as indicated by Eq.(50).

$$m = s + d + b + r \tag{50}$$

Where m is the measured current, s is the "photocurrent", d is the "dark current", b is the fixed voltage offset added to the analog-to-digital converter, and r is the readout noise. Expressions for the photocurrent and dark field are provided by Eq.(51) and Eq.(52).

$$s = \alpha Q N \tag{51}$$

$$d = \alpha \beta (T) t \tag{52}$$

Here, N is the number of photons, $\beta(T)$ is the temperature dependent dark current coefficient, and t is the chosen integration time. The noise in the photocurrent is the shot noise, and it was previously mentioned that for large number of photons the shot noise follows a normal distribution. Therefore, the width of the distribution will be $\sqrt{QN}$. We can also assume that the noise originating from the dark current also follows a normal distribution with width $\sqrt{\beta(T)t}$. Therefore, the noise from the two sources can be written as:

$$n_{phot} = \alpha\sqrt{QN} \qquad n_{dark} = \alpha\sqrt{\beta(T)t} \tag{53}$$

The formula for the total noise in the measurement in the case where it is assumed that the noise sources are independent and normally distributed is:

$$n_{tot} = \sqrt{n_{phot}^2 + n_{dark}^2 + n_{read}^2 + n_{bias}^2} \tag{54}$$

In Eq.(54) n_{read} refers to the read-out noise which is unavoidable and depends on the read-out speed, while n_{bias} is the bias noise and depends on the read-out electronics. The formula for the SNR in terms of the number of photons and different noise sources is provided in Eq.(55).

$$SNR = \frac{\alpha QN}{\sqrt{n_{phot}^2 + 2n_{dark}^2 + 2n_{read}^2 + 2n_{bias}^2}} \tag{55}$$

The reason for the factor "2" is because the signal "s" is defined by two separate measurements – the first measurement being done with the laser turned ON and the second, dark measurement, being done with the laser turned OFF [40]. In this way we can remove the contribution of the dark current and baseline to the measurements.

We have shown that the measurement process is not ideal and introduces certain errors arising from its physical limitations. So far, the conducted initial measurements haven't indicated a possible corruption of the measurement data due to noise errors originating from the camera. It is also worth investigating how much freedom the two camera models allow in terms of adjusting their data-acquisition process and measurement settings and whether such adjustments would considerably influence the result, obtained from the measured interferograms.

5.2.5 Vibrations

External vibrations and air turbulences can be a significant error source. However, the measurements are conducted in a clean room environment and the setup is located on an air-damped table. To further limit the effects of air turbulences, the setup is enclosed in a black box made from polycarbonate, which also serves to isolate it from the lighting in the room.

Vibrations, on the other hand, are currently a problem especially when using the spherical mirror, as shown in Fig.(3). The observed fringes are vibrating with a frequency that hasn't been yet

determined. In the paper of Ai [41], the effects of vibrations are considered. It was found that in the case that vibrations can be approximated by a single frequency oscillation, the phase-error introduced by said oscillation is periodic and has a frequency of twice the spatial frequency of the interference fringes.

Different attempts can be made to limit the effect of vibrations. For example, changing the currently used 3D printed plastic support of the spherical mirror with a metal one, mounting the OIS under test and the spherical mirror on the same X, Y, Z-stage, analyzing the oscillation frequency of the fringes and compensating for it.

5.3 Measurement Uncertainty Evaluation Model

There exist two main approaches to evaluate the uncertainty of the acquired phase map – analytically, and through Monte Carlo simulations, none of which has been the object of extensive studies [50], [51]. Both analyses consider the uncertainty of the calculated phase obtained from the phase stepping algorithm of choice in the presence of different errors. This analysis is incomplete as it doesn't proceed to consider the subsequently introduced uncertainties through the phase unwrapping and polynomial regression algorithms. Moreover, it assumes the researcher has knowledge about the uncertainty originating from individual error sources, which is often not the case. For example, the camera noise floor may be stated in the datasheet, but other quantities such as the background irradiance drift or the mid-spatial frequencies (MSF) present in the reference mirror may require separate measurements using additional equipment. This introduces its own difficulties. Hence, an uncertainty analysis model that limits the requirement for extra procedures is favored.

We propose two Measurement Uncertainty Evaluation Models. The first combines the already established analytical approaches [50] to obtain the measurement uncertainty per calculated Zernike coefficient. The second model uses Monte Carlo simulations that, as far as the author is aware, has not been utilized before. In both models, we strive to extract as much information as possible from actual measurements in order to limit the need for additional tests and equipment. In chapter 6 the results will be presented, however, the procedure will be outlined here.

5.3.1 Analytical Uncertainty Evaluation Model

The Analytical Uncertainty Analysis of linear PSAs has been developed largely by E. Hack [50] in accordance with the "Guide to the Expression of Uncertainty in Measurement" (GUM) [52]. His proposed method builds up on Freischlad's and Koliopoulos' early attempts at evaluating the uncertainties in the measured phase map originating from random errors by applying Gaussian error propagation to different PSAs [10].

The analytical formulas provided by Freischlad and Koliopoulos give an insight into the variance introduced by a small number of error sources. However, the given formulas are limited

in their applicability as they cannot directly be used to extract information from actual measurements. Instead, they offer an understanding of how errors affect the phase measurement. The subsequent expansion of the theory by E. Hack considers many more error sources, however, the downside is that the provided formulas assume the user has knowledge on the variation of important measurement parameters a priori. In some cases the information must be taken from the datasheet of an instrument (i.e. camera gain fluctuations), where we cannot immediately verify that the stated value is correct for our particular instrument, unless extensive troubleshooting and testing of the individual instrument is employed. Nevertheless, it is a useful tool if the obtained results are taken with a grain of salt. The following equations show the necessary preliminaries for acquiring the combined variance of the phase measurement, according to Hack. To obtain the measurement uncertainty in terms of the calculated Zernike coefficients an additional technique must be used, where the phase measurement uncertainty is translated to uncertainty of the fitted Zernike polynomials.

Let us first recall the expression for the calculated phase in terms of the arctangent function and the sampling weights that is given by Eq.(17).

$$\phi(x, y) = \arctan \frac{\sum_{n=0}^{N-1} \alpha_n I_n}{\sum_{n=0}^{N-1} \beta_n I_n} = \arctan \frac{G \sin(\phi)}{G \cos(\phi)} \tag{56}$$

Where α_n and β_n are real coefficients, previously referred to as the sampling weights of the algorithm, and G is a scaling factor, also referred to as "gain". We can rewrite the coefficients in vector form and introduce two additional vectors $\vec{A}$ and $\vec{B}$, which depend on the phase step $\Delta_{0 \dots N-1}$.

$$\vec{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{N-1} \end{bmatrix} \qquad \vec{A} = \begin{bmatrix} \sin(\Delta_0) \\ \sin(\Delta_1) \\ \vdots \\ \sin(\Delta_{N-1}) \end{bmatrix} \qquad \vec{B} = \begin{bmatrix} \cos(\Delta_0) \\ \cos(\Delta_1) \\ \vdots \\ \cos(\Delta_{N-1}) \end{bmatrix} \qquad \vec{I} = \begin{bmatrix} I_0 \\ I_1 \\ \vdots \\ I_{N-1} \end{bmatrix} \tag{57}$$

In order to simplify the expressions for the phase measurement uncertainty, Hack rewrites the sampling vectors $\vec{\alpha}$ and $\vec{\beta}$ as the complex vector $\vec{c}$, as well as the vectors $\vec{A}$ and $\vec{B}$ as the complex vector $\vec{C}$, where:

$$\begin{aligned} c_k &= b_k + i a_k = |c_k| exp(i\delta_k) \\ C_k &= B_k + i A_k = exp(i\Delta_k) \end{aligned} \tag{58}$$

Here, Δ_k are the phase step angles and δ_k are called coefficient angles, which can be calculated from Eq.(58) and obey a similar linear relation as the phase step angles. The formula for calculating the combined variance of the phase map is shown in Eq.(59).

$$u^2(\varphi,\ S_k) = \frac{4}{I''^2 \left| \vec{c} \cdot \vec{C} \right|^2} \left\{ \begin{array}{l} \sum\limits_{j=0}^{N-1} |c_j|^2 \sin^2(\varphi - \delta_j) u^2(S_j) + \\ + \sum\limits_{j=0}^{N-1} \sum\limits_{k>j}^{N-1} |c_j||c_k| sin(\varphi - \delta_j) sin(\varphi - \delta_k) u(S_j, S_k) \end{array} \right\} \tag{59}$$

Here, $u(S_j)$ is the individual signal uncertainties and $u(S_j, S_k)$ are covariances between the individual signals. The formula calculates the variance of the phase ϕ per pixel, hence, the values of ϕ and I''^2 are per pixel, or otherwise said - per datapoint. However, we require information on many individual signal uncertainties, which may be a downside to this model. Hack provides a table with formulas for quantitative evaluation that is used to find said uncertainties. The full table can be found in Appendix A.

The expression above gives the uncertainty per data point of the phase map, originating from various error sources. The phase map has to then be unwrapped using a phase unwrapping algorithm. This operation introduces some error, however, its nature is that of a systematic one that is more noticeable around edges and discontinuities, rather than a random error acting over the whole phase map. That is not to say that it can be ignored altogether, however, analytical evaluation of the uncertainty that is introduced by the phase unwrapping algorithm is beyond the scope of this report. To the author's knowledge, there doesn't exist a single unified theory or extensive research on the uncertainty caused by phase unwrapping algorithms as they are very diverse in their principles of operation and often times unpredictable. However, if such a theory were to exist, it may be incorporated in the current model. We will proceed under the assumption that no errors or uncertainties are introduced by phase unwrapping.

The following step, as outlined by section 4.4, is the fitting of Zernike polynomials. The most popular methods are the least-squares fitting techniques as they offer high computational efficiency due to their convenient matrix forms. Throughout the project, we have only employed linear least-squares fitting to evaluate the Zernike polynomials as the number of datapoints in the recorded interferograms are of the order of several hundred thousand and the number of fitted polynomials ranges up to 80.

It was already mentioned that polynomial regression techniques introduce errors, hence the fit coefficients have an uncertainty that doesn't stem from any uncertainties in the dataset but rather occurs from the fitting process itself. The usual way to obtain said uncertainties of the fit coefficients is to take the diagonal elements of the covariance matrix. However, in order to acquire the confidence intervals for the fitted coefficients, we must employ a different albeit similar approach called the "Maximum Likelihood Method" [53]. It has the added benefit that in the case the datapoints in our set have a known uncertainty/variance, the least-squares fitting algorithm properly accounts for it by using the same variance as a scaling factor. This means that datapoints, for which the uncertainty is high, will be deemed "untrustworthy" by the algorithm.

Hence, in regions with noisy data, such as those where pixel values exhibit medium to high-amplitude oscillatory behavior, the least-squares fit algorithm is likely to incorrectly attempt to fit high-order Zernike polynomial terms over the region. The expression for the quantity to be

minimized $\chi^2(\hat{a})$ in terms of the set of N datapoints $\hat{y}$ and the fitting function $\hat{y}(x; \hat{a})$, where x is an independent variable responsible for the uncertainty in the fit data and $\hat{a}$ is the vector containing the coefficients used to minimize χ^2 is given by Eq.(60).

$$\chi^2(\vec{a}) = \sum_{i=1}^{N} \frac{[y_i - \hat{y}(x_i; \hat{a})]^2}{u^2(\phi_i, S_k)} \tag{60}$$

In the expression above, you can notice that we have employed the variance of the phase map, per pixel i, that is given by Eq.(59). In section 4.4, the expression for calculating the fitted Zernike coefficients $\hat{a}_J$ for J number of Zernike polynomials was shown as:

$$\hat{a}_{Jx1} = \left(A_{NxJ}{}^T A_{NxJ}\right)^{-1} A_{NxJ}{}^T D_{Nx1}$$

Where D_{Nx1} is a vector containing the datapoints and A_{NxJ} is a matrix of the fitting function, in our case - J number of Zernike polynomials evaluated at each datapoint. We can rewrite this equation in terms of the covariance matrix C_{JxJ}, where $C_{JxJ} = \left(A_{NxJ}{}^T A_{NxJ}\right)^{-1}$. This gives us Eq.(61). Here, if variance of the phase map $u^2(\phi, S_k)$ is not employed, the elements along the diagonal of C_{JxJ} give us the variance of the coefficients due to the fit error.

$$\hat{a}_{Jx1} = C_{JxJ} A_{NxJ}{}^T D_{Nx1} \tag{61}$$

In the method of maximum likelihood, proposed by Richter [53], if we scale the matrix A_{NxJ} and the vector D_{Nx1} element-wise by dividing a particular datapoint by the standard deviation for said datapoint, we can obtain the uncertainty of the fitted Zernike coefficients $\delta\hat{a}$. The new matrix A'_{NxJ} and vector D'_{Nx1} would therefore contain the following components:

$$A'_{ij} = \frac{A_{ij}}{u(\phi_{ij}, S_k)} \qquad D'_{ij} = \frac{D_{ij}}{u(\phi_{ij}, S_k)} \tag{62}$$

The uncertainties in the calculated Zernike coefficients, originating from the fit error, as well as the various errors present throughout the interferogram measurement process, can then be found using Eq.(63).

$$\delta\hat{a}_{Jx1} = C'_{JxJ} A'^{T}_{NxJ} D'_{Nx1} \tag{63}$$

5.3.2 Monte Carlo Uncertainty Evaluation Model

Compared to the analytical one, the Monte Carlo Uncertainty Evaluation Model can be described in terms of a set number of procedures that need to be executed in a certain order, rather than be represented by a compact set of formulas and equations. The main principle of this model revolves around the idea that we can make a realistic simulation of the measurement process using a computer model. By coupling the results from real measurements with results obtained from the simulation model, we can extract information on several systematic errors, which we can correct for. The model then requires statistical information about the random errors present in our measurements. However, unlike in the case of the analytical model, this information need not be translated to uncertainty in the phase map. It is also not always

necessary to quantify individual noise sources, such as the noise originating from the camera, the backround irradiance, laser source instability, etc. As long as the combined effect of said errors over the pupil can be determined, this can easily be implemented in the simulation model, where we can run Monte Carlo simulations to obtain the variance per calculated Zernike coefficient.

There are two necessary preliminaries that are part of the model – an interferogram analyzer, and an artificial interferogram generator. Both modules will be thoroughly explained in chapter 6. There are several companies that offer commercial software for interferogram analysis. However, in this project we have designed our own using the previously outlined principles of interferometric measurements.

The procedure is as follows:

Step 1: Record real phase-stepped interferograms using an interferometer of choice. These can be any number recordings depending on the phase-stepping algorithm of choice.

Step 2: Use the interferogram analyzer to acquire the wavefront and fit a set number of Zernike polynomials over it. The results produced by the analyzer are the fit coefficients for the chosen Zernike polynomials, which are our initial result.

Step 3: Reconstruct the wavefront using the same calculated coefficients and subtract the reconstructed wavefront from the actual wavefront, as obtained after the phase-stepping algorithm, the phase unwrapping algorithm, and wavefront scaling. This gives us the mid-spatial frequencies (MSF).

Step 4: Plug the initially calculated Zernike coefficients, the obtained MSF into the artificial interferogram generator. Generate artificial interferograms.

Step 5: Analyze the generated artificial interferograms using the interferogram analyzer. The newly calculated Zernike coefficients will differ from the initial ones by a few $m\lambda$. This is due to the effects of the MSF and the linear fit error. However, this is a systematic error which we can store and remove later on. We will refer to this error as the "Systematic Error".

Step 6: Plug any available statistical information on the error sources into the artificial interferogram generator together with the initially calculated Zernike coefficients. This is the heart of the Monte Carlo simulation. A set number of cycles N is ran, i.e 500, where we generate and subsequently analyze synthetic interferograms with added noise, taken from their respective distributions. Each set of calculated Zernike coefficients is stored.

Step 7: Remove the systematic "Fit Error" from the datasets. Next, analyze the data by employing techniques like computing the sample's mean and variance, and creating histograms and boxplots.

Flowcharts of the full process can be seen in Fig.(21 - 23). More detailed information on the generation and analysis of the interferograms will follow in section 6.1. This also includes the procedures through which statistical information on the error sources was obtained, as well as the software implementation in terms of the used modules and finally - the obtained results.
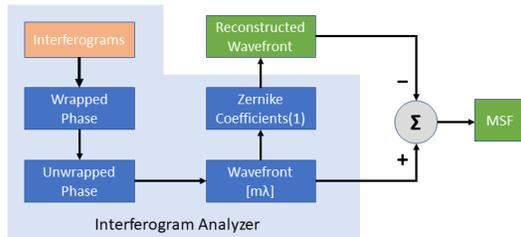

Figure 21: Step 1-3: Description of the "Interferogram Analyzer" module as it is used to extract the mid-spatial frequencies. The input interferograms in the current case must originate from real measurements.
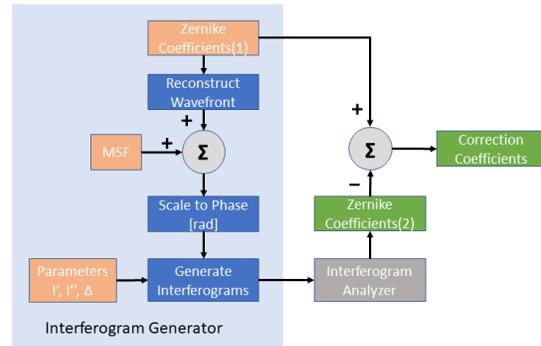
Figure 22: Step 4-5: Description of the "Interferogram Generator" module as it is used to extract systematic "Fit Error". The parameters depend on the PSA of choice.
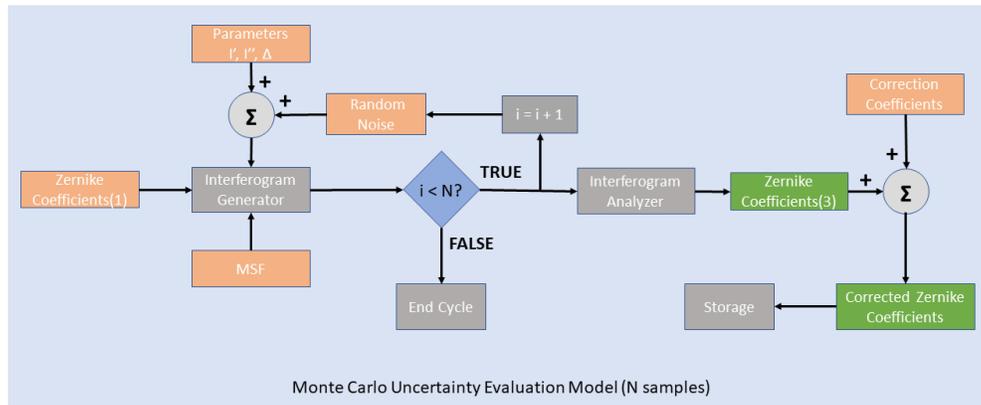
Figure 23: Step 6-7: Overview of the general process of obtaining statistical datasets per Zernike coefficient using Monte Carlo simulations and available information on the random noise sources.

In the above flowcharts, the peach-colored blocks represent inputs, the dark blue ones represent individual operations inside a module, the green blocks stand for intermediate or final results, the gray blocks may either be full modules, i.e. "Interferogram Generator", "Interferogram Analyzer", or represent miscellaneous operations such as the i counter. The light blue

background defines a full module. In the flowchart shown in Fig.(23), the acquired Zernike co-efficients per iteration are stored in a storage as a statistical dataset. It is up to the user to choose an appropriate statistics technique for the evaluation of said dataset (i.e. histograms, boxplots, etc.).

6 Experiments & Results

In this chapter the results from the conducted experiments will be presented. First, the simulation model used in the uncertainty evaluation of the interferometric measurements will be presented in detail. Next, the acquired results for the systematic errors will be shown. The systematic error in our measurements consists of the added aberrations present in the reference spherical mirror, the fit error originating from the polynomial regression model, and the mid-spatial frequencies (MSF). Then, the procedure for obtaining the mid-spatial frequencies will be outlined. The final section concludes the chapter with the results of the Monte Carlo simulation for various amounts of random noise in terms of the individual uncertainty per calculated Zernike coefficient.

6.1 Interferometric Measurements: Simulation Model

Perhaps the most important module is the one used for the analysis of the interferograms. It takes real or artificial interferograms, computes the wrapped phase map using a PSA of choice. We have only considered two PSAs – 4-step Wyant and 5-step Schwieder-Hariharan, both of which use interferograms that are phase-stepped at 90°. However, the program can be used with any PSA regardless of the number of interferograms required or the phase step. The wrapped phase is then unwrapped using a PU algorithm. Two readily-available PU algorithms were imported as modules in the Python script. The first one is a region-growing algorithm that was designed by M. Herraez et al. The second one uses discrete cosine transform (DCT), was designed by D. Ghiglia and L. Romero, and requires a weighting mask for best performance. The fitting of Zernike polynomials is done using the least squares technique outlined in section 4.4.1. An overview of the work principle of the full module is provided in pseudocode by Alg.(1).

Algorithm 1 Interferogram Analyzer
_____

1: Load interferograms, Include libraries, Include PU module, Include Cartesian Zernike polynomials Z[j]
2: **procedure** CALCULATE WRAPPED PHASE(Interferograms: $intf_1, intf_2, \ldots, intf_n$)
3:      Apply a PSA: $\phi_{wrapped} = atan2(PSA_{numerator}, PSA_{denominator})$
4: **end procedure**
5: **procedure** PHASE UNWRAPPING(wrapped phase: $\phi_{wrapped}$)
6:      Unwrap phase: $\phi_{unwrapped} = PU(\phi_{wrapped})$
7:      Scale $rad => m\lambda$: Wavefront WF = $\frac{\phi_{unwrapped}}{4\pi * 10^{-3}}$
8: **end procedure**
9: **procedure** POLYNOMIAL FITTING(WF, Number of polynomials: J)
10:      Flatten the wavefront: D_{Nx1} = WF.flatten()
11:      Create empty A_{NxJ} array and fill it
12:      **for** $x = 1, 2, \ldots, x_{size}$ of WF **do**
13:          **for** $y = 1, 2, \ldots, y_{size}$ of WF **do**
14:              **for** $j = 1, 2, \ldots, J$ **do**
15:                  $A[y * x_{size} + x, j] = Z[j](x, y)$;
16:              **end for**
17:          **end for**
18:      **end for**
19:      Calculated Zernike Coefficients: $\hat{a}_{Jx1} = \left(A_{NxJ}{}^T A_{NxJ}\right)^{-1} A_{NxJ}{}^T D_{Nx1}$
20: **end procedure**
_____

The second important module, the Interferogram Generator, is described by Alg.(2). It generates a wavefront from a chosen set of Zernike coefficients – for example, the ones calculated by the Interferogram Analyzer. The same applies for the average irradiance $I^{'}$ and the fringe modulation $I^{"}$. We may either select values for them or use values as extracted by the Interferogram Analyzer.

Here it needs to be reminded once more that there are 3 unknowns per pixel/datapoint – the phase/wavefront ϕ, average intensity I', and fringe modulation $I^{"}$. As most PSAs use more than 3 interferograms, the problem of finding the unknowns becomes overdetermined. Therefore, I' and $I^{"}$ can be found using a least squares fit. Once found, they can be input in the Interferogram Generator.

Algorithm 2 Interferogram Generator

1: Load Zernike Coefficients $C_1, C_2, \ldots, C_n$, Include Cartesian Zernike polynomials as functions $Z[1,2,\ldots,n](x,y)$
2: Optional: Load MSF, Load I' and I'' as arrays
3: **procedure** RECONSTRUCT WAVEFRONT(Zernike coefficients)
4:     Choose Wavefront/Plot dimensions: $WF_{size} = MSF_{size} = (x_{size} \, x \, y_{size})$ datapoints
5:     **for** $x = 1, 2, \ldots, x_{size}$ **do**
6:         **for** $y = 1, 2, \ldots, y_{size}$ **do**
7:             WF$[x,y] = C_1 Z_1[x,y] + C_2 Z_2[x,y] + \ldots + C_n Z_n[x,y]$
8:         **end for**
9:     **end for**
10:     Add the MSF: $WF = WF + MSF$
11: **end procedure**
12: Scale $m\lambda => rad$: $\phi = WF * (4\pi * 10^{-3})$
13: **procedure** GENERATE INTERFEROGRAMS(ϕ, I', I'', $\Delta_{1,2,\ldots,K}$)
14:     **for** $k = 1, 2, \ldots, K$ **do**
15:         **for** $x = 1, 2, \ldots, x_{size}$ **do**
16:             **for** $y = 1, 2, \ldots, y_{size}$ **do**
17:                 $I_k[x,y] = I'[x,y] + I''[x,y] * cos(\phi[x,y] + \Delta_k)$
18:             **end for**
19:         **end for**
20:     **end for**
21: **end procedure**

The third module described by Alg.(3) uses the two previously-described modules where random noise is added. As it is not an analytical uncertainty model, the effects of some noise sources may be grouped together. For example, the laser instability, camera noise and shot noise are grouped as simply Gaussian noise over the whole interferogram. The shot noise follows a Poisson distribution, however, for a high photon count, the distribution approaches a normal one.

In the general theory of phase stepping interferometry, it is assumed that I' and I'' are constant over the recorded interferograms. In reality that might not be the case and small variations are very likely.

Finally, the phase stepping can never be ideal. There are certain self-calibrating PSAs that do not require a constant phase step. However, they usually lack in their error-compensation abilities. Also, this is not the case for Wyant's 4-step algorithm or Schwieder-Hariharan's 5 step one as they require a constant phase step of 90°. The 5-step algorithm compensates fairly well for phase step errors, however, this inevitably adds some amount of uncertainty in the measurements. The maximum amount of phase-step error can be approximated by looking at the datasheet of the piezo transducer of the flat reference mirror. Of interest is the repeatability along the axis of movement, as well as the possible pitch and yaw rotation about the other two axes that may occur during phase-stepping.

Algorithm 3 Monte Carlo Uncertainty Evaluation Module Using Schwieder-Hariharan

1: Load Interferogram Analyzer, Load Interferogram Generator
2: Load Zernike Coefficients $C_1, C_2, \ldots, C_n$, Include Cartesian Zernike polynomials as functions $Z[1, 2, \ldots, n](x, y)$
3: Load MSF, Load I' and I'', Load Correction Coefficients $cc_1, cc_2, \ldots, cc_n$
4: **for** $k = 1, 2, \ldots, K$ **do**
5:   Add noise: $I' + \delta_k I'$, $I'' + \delta_k I''$, $\Delta + \delta_k \Delta$, Gaussian Img Noise: G
6:   **procedure** INTERFEROGRAM GENERATOR(Z Coeff, MSF, $I' + \delta_k I'$, $I'' + \delta_k I''$, $\Delta + \delta_k \Delta$, G)
7:    **return** Artificial Interferograms: $intf_{k,1}, intf_{k,2}, \ldots, intf_{k,5}$
8:   **end procedure**
9:   **procedure** INTERFEROGRAM ANALYZER($intf_{k,1}, intf_{k,2}, \ldots, intf_{k,5}$)
10:    **return** Calculated Zernike Coefficients: $C'_1, C'_2, \ldots, C'_n$
11:   **end procedure**
12:   **procedure** CORRECT FOR SYSTEMATIC ERROR(Calc Z Coeff: $C'_1, C'_2, \ldots, C'_n$, Correction Coeff: $cc_1, cc_2, \ldots, cc_n$)
13:    **return** $C''_{k,1}, C''_{k,2}, \ldots, C''_{k,n} = C'_1 + cc_1, C'_2 + cc_2, \ldots, C'_n + cc_n$
14:   **end procedure**
15: **end for**
16: Plot Boxplots or Histograms Per Zernike Coefficient from $C''_{k,1}, C''_{k,2}, \ldots, C''_{k,n}$

6.2 Systematic Error Analysis

This section deals with the identification and removal of systematic errors in our measurements. The systematic errors consist of the additional aberrations introduced in our measurements by the spherical reference mirror of NA ≤ 0.985, the error originating from the polynomial regression, and the effect of the mid-spatial frequencies (MSF) on the calculated Zernike coefficients.

6.2.1 Aberration Profile of the Spherical Reference Mirror

A major downside of Twyman-Green interferometers is the need for an optical reference to test against. In our case that is the spherical mirror of NA ≤ 0.985 and the planar mirror. The spherical mirror contains optical aberrations that inevitably make part of the measurements. However, an advantage is that we can rotate it by 180°. The uneven aberration terms of the spherical mirror would then change sign and can be subtracted from the measurement. The principle is explained below and can be seen in Fig.(24).

The wavefront W(x,y) can be expressed in terms of a sum of weighted polynomials (i.e. Zernike polynomials). The individual polynomials can either be angularly odd or even. This can be expressed as:
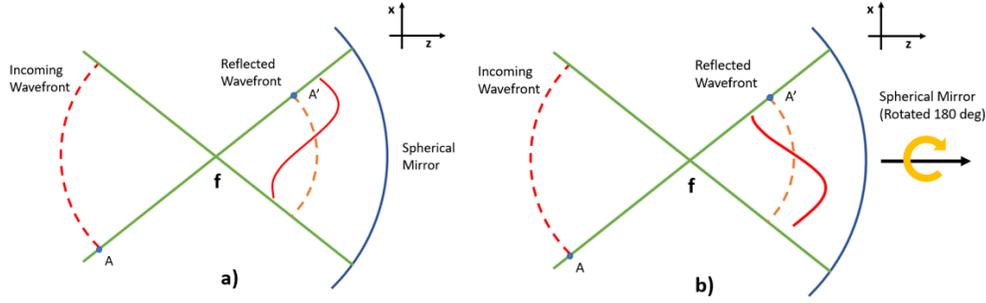
Figure 24: Aberrated wavefront after reflection from the spherical mirror: a) spherical mirror at 0° b) spherical mirror rotated by 180°

$$W(x, y) = W_{\text{even}}(x, y) + W_{\text{odd}}(x, y) =$$
$$= \sum_{i=0}^{\infty} c_{\text{even},i} P_{\text{even},i}(x, y) + \sum_{i=0}^{\infty} c_{\text{odd},i} P_{\text{odd},i}(x, y) \tag{64}$$

Rotating the spherical mirror by 180° about the z-axis has the effect of reversing the positive and negative x- and y-axes as shown in Fig.(25).
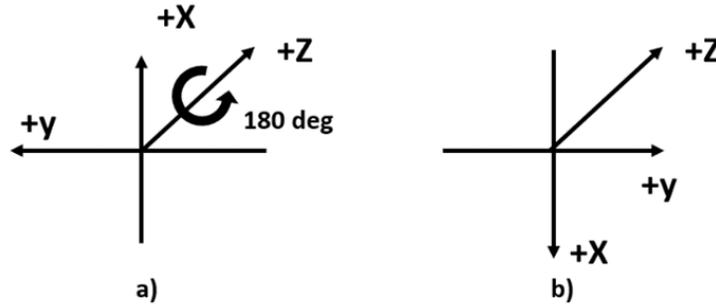
Figure 25: Axes: a) prior to rotation of the spherical mirror b) after rotation of the spherical mirror. Odd aberrations change their signs in the measurement.

Therefore, the wavefront after rotation can be expressed in terms of the coordinates of the initial non-rotated axial system as shown in Eq.(65).

$$W_{\text{rotated}}(x, y) = W(-x, -y) = \sum_{i=0}^{\infty} c_{\text{even},i} P_{\text{even},i}(-x, -y) + \sum_{i=0}^{\infty} c_{\text{odd},i} P_{\text{odd},i}(-x, -y) =$$
$$= \sum_{i=0}^{\infty} c_{\text{even},i} P_{\text{even},i}(x, y) - \sum_{i=0}^{\infty} c_{\text{odd},i} P_{\text{odd},i}(x, y) \tag{65}$$

A measurement using the spherical mirror, $W_{\text{with}_\jmath}(x, y)$, contains the even and odd aberrations introduced by the microscope objective, as well as the even and odd aberrations introduced by the spherical mirror. Therefore we can write the expression for the measured wavefront in terms of the individual components as:

$$W_{\text{with}_\jmath}(x, y) = W_{\text{even}_{\text{obj}}}(x, y) + W_{\text{odd}_{\text{obj}}}(x, y) + W_{\text{even}_\jmath}(x, y) + W_{\text{odd}_\jmath}(x, y) \tag{66}$$

By conducting two measurements – the first with the spherical mirror at 0 degrees and the second with the spherical mirror rotated by 180 degrees, and subtracting the two wavefronts, then dividing by two, we can quantify the odd aberrations of the spherical mirror.

$$W_{\text{odd}_\jmath}(x, y) = \frac{W_{\text{with}_\jmath}(x, y) - W_{\text{with}_\jmath}(-x, -y)}{2} \tag{67}$$

Then an on-axis measurement is coducted using a planar mirror, $W_{\text{with}_|}(x, y)$, where the mirror is placed at the focal point of the lens/objective under test. We can assume the planar mirror doesn't introduce additional aberrations when the wavefront is reflected. As light falls onto a small point – a few microns in diameter, we assume the surface at that particular location to be ideally flat. However, a planar mirror cannot be used to measure the odd aberrations of the microscope objective. Upon reflection, the wavefront is flipped. This means that odd aberrations change sign and are cancelled after passing through the microscope objective for a second time. The measured aberrations are the even aberrations of the microscope objective.

$$W_{\text{with}_|}(x, y) = W_{\text{even}_{\text{obj}}}(x, y) \tag{68}$$

The even aberrations of the spherical mirror can be found from Eq.(69). The result consists of the odd aberrations of the objective and the even aberrations of the spherical mirror. We can then represent the result in terms of Zernike polynomials, where the Zernike polynomials can be separated into even and odd polynomials in terms of their angular dependence. The angularly odd Zernike polynomials represent the odd aberrations of the microscope objective and the angularly even Zernike polynomials the even aberrations of the spherical mirror.

$$W_{\text{with}_\jmath}(x, y) - W_{\text{with}_|}(x, y) - W_{\text{odd}_\jmath}(x, y) = W_{\text{odd}_{\text{obj}}}(x, y) + W_{\text{even}_\jmath}(x, y) \tag{69}$$

In this way the aberration profile of the spherical mirror can be obtained. It can then be used to measure over the field. Such a measurement is straightforward when the lens under test is a telecentric lens (as in most ASML lenses of interest). The telecentricity guarantees that the pupil is completely filled with light also for off-axis field points. Moreover, when a spherical mirror is used in off-axis measurements, the telecentricity guarantees that for all field points one always uses the same part of the spherical mirror. This much simplifies the correction of the measurements for the aberrations in the spherical mirror. In Fig.(26, 27) the wavefront corresponding to the odd and even aberrations present in the spherical mirror is shown.
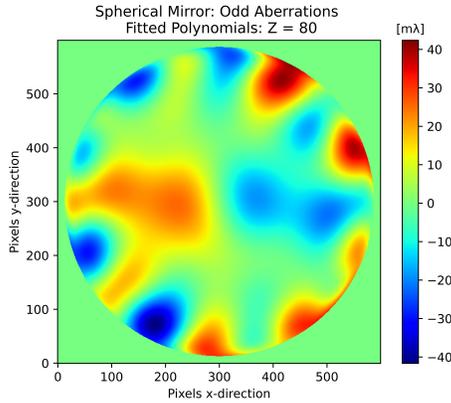
Figure 26: Wavefront corresponding to the odd abberations present in the spherical mirror. Tilt is removed.
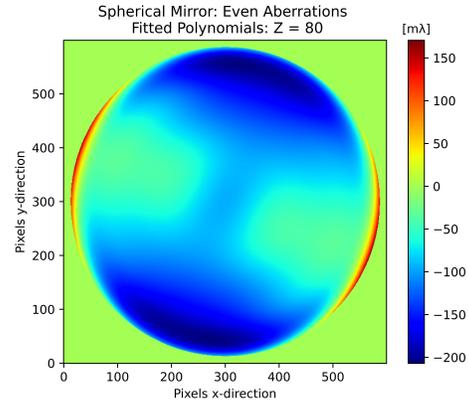

Figure 27: Wavefront corresponding to the even abberations present in the spherical mirror. Piston and defocus are removed.

To obtain the whole aberration profile of the spherical mirror we can sum the wavefront contributions from the odd and even aberrations. This aberration profile can later be subtracted from subsequent measurements over the field.

6.2.2 Mid-Spatial Frequencies

Mid-Spatial Frequencies can often times be detrimental to the performance of a precision optical instrument. In our measurements, their presence can be recognized in the concentric patterns of the recorded interferograms, which originate from the reference sphere. They correspond to the spatial frequencies, which are higher than the first 80 Zernike polynomials but lower than the surface roughness. Their origin lies in the production process of spherical mirrors and their presence is largely unavoidable. Often times their effect is neglected and trivialized, however, as the demand for more precise optical instruments increases, various attempts have been made to correctly measure them and consider their effects in optical designs.

As we are investigating various potential culprits, that can introduce measurement uncertainty in a Twyman-Green interferometer, we cannot leave the topic of MSF untouched. During various experiments, where data from real measurements was purposefully introduced into simulation models, it was found out that introducing MSF to synthetic interferograms has a noticeable, yet predictable effect on the calculated Zernike coefficients.

In this section, first the procedure to separate MSF from interferometric measurements (Fig.21) will be presented in detail. Then, the effect of the MSF on the calculated Zernike coefficients will be considered.

Obtaining the Mid-Spatial Frequencies

The MSF have a higher spatial frequency as compared to Zernike polynomials, which can be used to our advantage. This means that the least squares fit algorithms wouldn't be able to represent them in terms of Zernike polynomials, unless an unreasonably-large number of polynomials is fit. Not only is this unnecessary, but it also presents itself as a problem, since a recorded interferogram is effectively a discrete data array, meaning that there is a limit to the highest polynomial order that can be fit before overfitting begins to take place due to discretization. Therefore, if a reasonably-large number of Zernike polynomials is fit (i.e. 80), the MSFs would not be represented in terms of Zernike coefficients and can therefore be obtained by subtracting the reconstructed wavefront from the original one.

It is worth mentioning that this procedure does not yield the exact MSF, but rather the residual wavefront that couldn't be represented in terms of the first 80 Zernike polynomials. This residual contains the MSF together with the high-spatial frequencies. The high-spatial frequencies originate from the surface roughness of the optical reference, which in our case is less than $1nm$ for polished surfaces. We choose to refer to the residual wavefront as MSF as they have a predominant effect due to the clearly visible concentric grooves in the recorded interferograms. Nevertheless, we do not omit the effects of high-spatial frequencies in our analysis.

The obtained MSF can be seen in Fig.(30). The piston, tilt, and defocus terms have priorly been removed from the original and reconstructed wavefront for better visualization.
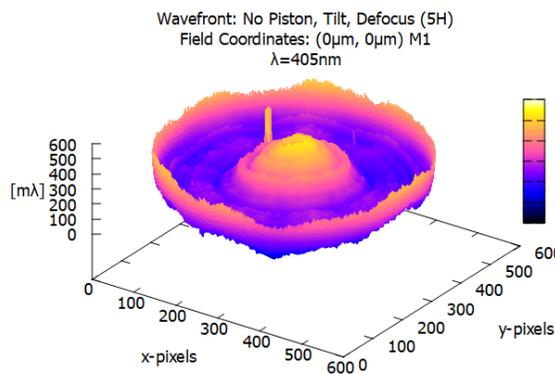
Figure 28: Wavefront as aquired by real interferograms using Schwider-Hariharan's 5 step PSA and Herraez' PU algorithm. On-axis measurement.
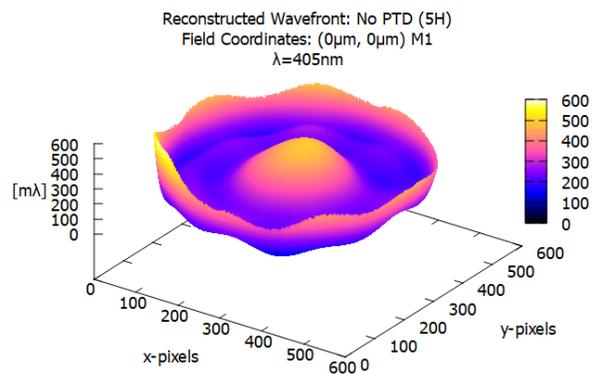
Figure 29: The reconstructed wavefront from the first 80 fitted Zernike coefficients.
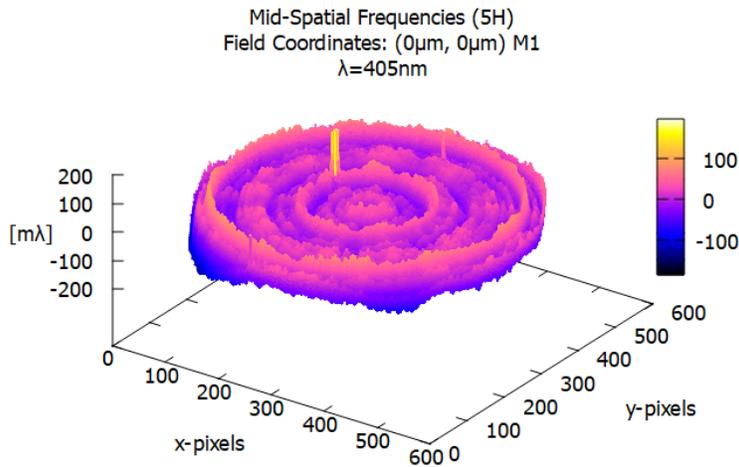
Figure 30: The residual wavefront as obtained from subtracting the reconstructed one from the original. This we refer to as MSF.

The figures above show the wavefront and the MSF after fitting of the first 80 Zernike polynomials for an on-axis measurement as described in Fig.(2). The results for different field coordinate measurements show that the MSF is constant for the most part where the peak-to-peak amplitude is of the order of $\sim 150 - 200m\lambda$. In Fig.(31) and Fig.(32) you can see a comparison of the MSF as obtained from an on-axis measurement and a measurement at field coordinates $(-25\mu m, -25\mu m)$.
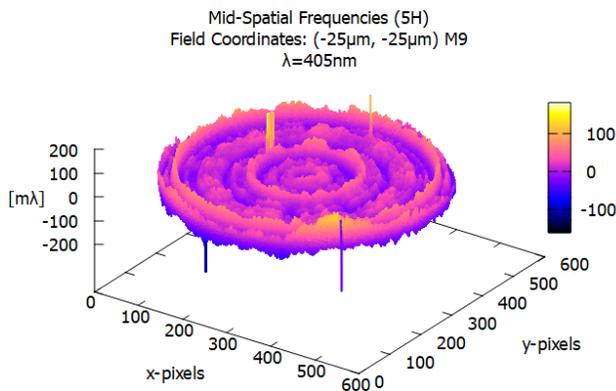
Figure 31: The MSF obtained from the fitting of the first 80 Zernike coefficients for an off-axis measurement with field coordinates $(-25\mu m, -25\mu m)$.
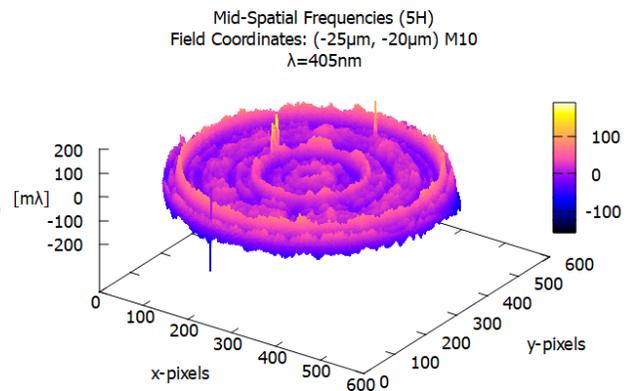
Figure 32: The MSF obtained from the fitting of the first 80 Zernike coefficients for an off-axis measurement with field coordinates $(-25\mu m, -20\mu m)$.

Effect of the Mid-Spatial Frequencies on the Calculated Zernike Coefficients

It is difficult to estimate beforehand the effect the MSF would have on interferometric measurements. The reason for that are the two main algorithms employed – the phase unwrapping algorithm and the polynomial regression model. We can predict that the MSF would have some effect on the unwrapping of the phase. However, to what extent depends largely on the PU algorithm of choice and how well it can compensate for noise. A similar situation arises with the polynomial regression model. We know that spatial noise, as introduced by the MSF, would have an effect on the calculated fit coefficients, however, it depends on the individual polynomial we are fitting over the wavefront. For example, the results show that the calculated fit coefficient for triangular astigmatism along "Y" (Z_{-3}^3) is less affected by fit error and the MSF originating from the spherical mirror, as compared to tilt or defocus. In order to test the effects of MSF, Fit and PU error on the calculated Zernike coefficients an experiment was conducted using a simulation model. The individual procedures have been described by Alg.(4-6).

Algorithm (4) is used to find the Fit and PU error in the absence of any MSF. First, a wavefront is generated using a chosen set of Zernike coefficients, then the wavefront is scaled and wrapped in the principal value range of the arctangent function $[-\pi; +\pi]$. This operation resembles the way the wrapped phase is usually obtained using a PSA and interferograms. The wrapped wavefront is later unwrapped using a PU algorithm of choice, which inevitably introduces some additional errors. Finally, the same polynomials used for the initial generation are fit using a least squares approach. The calculated Zernike coefficients can be subtracted from the original ones used in the generation step to arrive at the error per coefficient that the used operations introduce.

Algorithm 4 Evaluate Systematic Error: Fit & PU

1: **procedure** GENERATE WAVEFRONT(Zernike coefficients, Wavefront dimensions)
2:     Choose n Zernike coefficients: $C_1, C_2, \ldots, C_n$
3:     Choose Wavefront/Plot dimensions: $WF_{size} = (x_{size} \text{ x } y_{size})$ datapoints
4:     **for** $x = 1, 2, \ldots, x_{size}$ **do**
5:         **for** $y = 1, 2, \ldots, y_{size}$ **do**
6:             WF$[x, y] = C_1 Z_1[x, y] + C_2 Z_2[x, y] + \ldots + C_n Z_n[x, y]$
7:         **end for**
8:     **end for**
9:     Scale to rad & Wrap the wavefront in the range $[-\pi; +\pi]$
10:     Unwrap the wavefront: PU algorithm of choice (i.e. Herraez, Ghiglia)
11:     Scale back to $m\lambda$
12:     Fit the n Zernike polynomials: $C_1', C_2', \ldots, C_n'$
13:     Calculate the difference between original & calculated: $C_1 - C_1', C_2 - C_2', \ldots, C_n - C_n'$
14:     Plot the results
15: **end procedure**

Algorithm ([5](#)) differs from Alg.([4](#)) in that it adds MSF to the generated wavefront. The full procedure is otherwise identical.

Algorithm 5 Evaluate Systematic Error: Fit, PU, MSF Added

1: **procedure** GENERATE WAVEFRONT(Zernike coefficients, Wavefront dimensions)
2:   Choose n Zernike coefficients: $C_1, C_2, \ldots, C_n$
3:   Choose Wavefront/Plot dimensions: $WF_{size} = (x_{size} \times y_{size})$ datapoints
4:   **for** $x = 1, 2, \ldots, x_{size}$ **do**
5:     **for** $y = 1, 2, \ldots, y_{size}$ **do**
6:       $WF[x, y] = C_1 Z_1[x, y] + C_2 Z_2[x, y] + \ldots + C_n Z_n[x, y]$
7:     **end for**
8:   **end for**
9:   Add the previously-found MSF to the Wavefront: $WF_{MSF} = WF + MSF$
10:   Scale to rad & Wrap the wavefront in the range $[-\pi; +\pi]$
11:   Unwrap the wavefront: PU algorithm of choice (i.e. Herraez, Ghiglia)
12:   Scale back to $m\lambda$
13:   Fit the n Zernike polynomials: $C_1', C_2', \ldots, C_n'$
14:   Calculate the difference between original & calculated: $C_1 - C_1', C_2 - C_2', \ldots, C_n - C_n'$
15:   Plot the results
16: **end procedure**

In Alg.([6](#)), the MSF are added to the generated wavefront, then after phase wrapping and unwrapping they are subtracted.

Algorithm 6 Evaluate Systematic Error: Fit, PU, MSF Added & Subtracted

---

1: **procedure** GENERATE WAVEFRONT(Zernike coefficients, Wavefront dimensions)
2:   Choose n Zernike coefficients: $C_1, C_2, \ldots, C_n$
3:   Choose Wavefront/Plot dimensions: $WF_{size} = (x_{size} \times y_{size})$ datapoints
4:   **for** $x = 1, 2, \ldots, x_{size}$ **do**
5:     **for** $y = 1, 2, \ldots, y_{size}$ **do**
6:       $WF[x, y] = C_1 Z_1[x, y] + C_2 Z_2[x, y] + \ldots + C_n Z_n[x, y]$
7:     **end for**
8:   **end for**
9:   Add the previously-found MSF to the Wavefront: $WF_{MSF} = WF + MSF$
10:   Scale to rad & Wrap the wavefront in the range $[-\pi; +\pi]$
11:   Unwrap the wavefront: PU algorithm of choice (i.e. Herraez, Ghiglia)
12:   Scale back to $m\lambda$
13:   Subtract the previously-added MSF from the Wavefront: $WF' = WF_{MSF} - MSF$
14:   Fit the n Zernike polynomials: $C_1', C_2', \ldots, C_n'$
15:   Calculate the difference between original & calculated: $C_1 - C_1', C_2 - C_2', \ldots, C_n - C_n'$
16:   Plot the results
17: **end procedure**

Figures ([33]) and ([34]) show the results for the systematic error per Zernike coefficient for two used PU algorithms – Herraez' and Ghiglia's. The MSF is obtained from real measurements, where it was shown that it is equivalent to the wavefront residual left after the fitting of the first 80 Zernike polynomials. In the current experiment, the effects on the first 10 Zernike polynomials were considered. Hence, adding a residual should in theory have no influence on the calculated coefficients. However, the results show there is some, which most likely comes from the PU algorithm not performing as well. Nevertheless, the exact effect per coefficient can be found using Alg.([4-6]). This systematic error can then be removed from actual measurements.
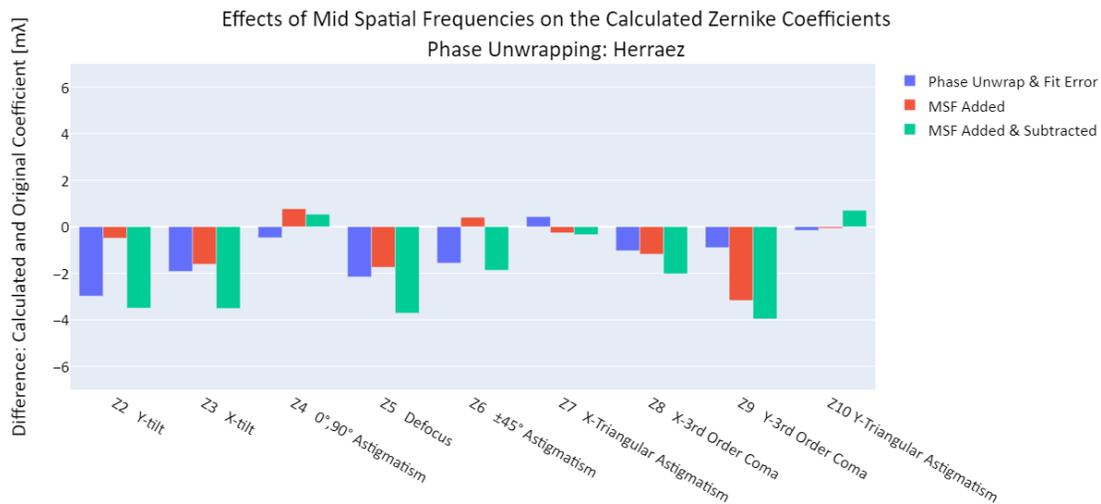
Figure 33: The systematic error in the evaluated first 10 Zernike coefficients as outlined in Alg.([4] - [6]). Used PU algorithm: Herraez
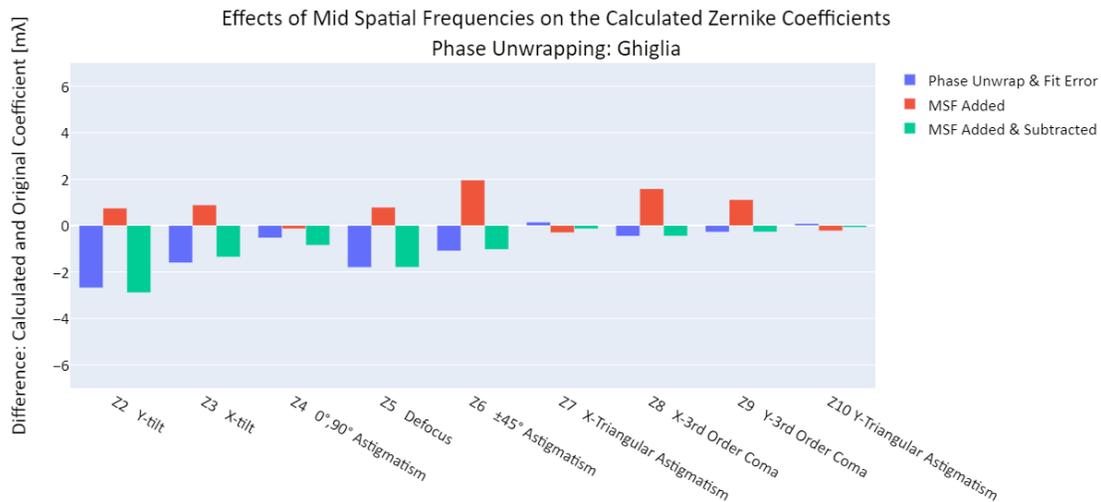
Figure 34: The systematic error in the evaluated first 10 Zernike coefficients as outlined in Alg.([4] - [6]). Used PU algorithm: Ghiglia

It is not recommended to subtract the MSF after adding them to the generated wavefront as the PU algorithm has built-in defenses against noise, which act as a filter. This means that some of the noise present in the MSF will be filtered out during PU. When we later subtract the MSF from the unwrapped phase, we are adding the same noise back with a negative sign, hence why in Fig.(33) and Fig.(34) for some coefficients it is noticeable that addition and subtraction of MSF performs worse than addition only. Nevertheless, some form of error-compensation must occur as we can see that the errors behave unpredictably.

6.3   Laser Instability and Camera Noise

There are various sources of error in interferometric wavefront measurements. One point of concern is the stability of the used light source. In the conducted measurements various monochromatic CW lasers were used, as already outlined by Table 1. During live-video observation of the interferogram image on the computer screen, small image intensity fluctuations were observed. There are possible explanations for this, such as a fault in the electronics of the laser, unstable power supply, faulty PID controller of the temperature control for the laser diode, or thermal noise in the cavity of the laser diode. Lasers are complicated devices and laser troubleshooting is very well a research area on its own. Therefore the exact causes of the instability fall outside of the scope of this research. However, as the instability of the laser source is an observable phenomenon, we can record it, quantify it, and use the results to update the simulation model, from which measurement uncertainties are derived.

For the purpose of analyzing the irradiance fluctuations, the camera was positioned directly after the collimator lens as shown in the (figure below) and a recording was made. The $8bit$ values from a grid of pixels were recorded over a duration of $100s$ with a frame rate of $75fps$. A total of 5000 frames with a size of ($96px$ X $96px$) were recorded. The frames were then processed using the statistical analysis proposed in this subsection.

It is worth noting that the analysis assumes Gaussian distribution of the random variables. This assumption can be validated by taking a look at how the pixel values change over the span of the recording. In Fig.(35) the distribution of the central pixel value over the recorded 5000 frames has been shown. We can see that the values follow a Gaussian distribution and are centered around a mean. Subsequently, the mean values per pixel over all 5000 frames were calculated and the distribution over the grid has been plotted. This can be seen in Fig.(36). Again, a familiar Gaussian distribution appears.
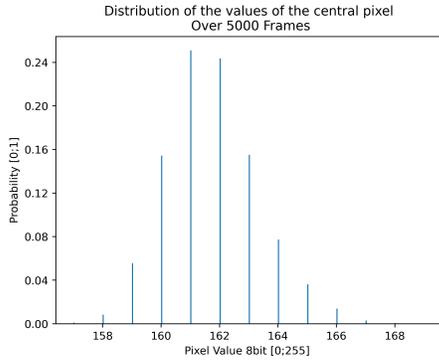
Figure 35: Distribution of the values of the central pixel over the 5000 recorded frames.
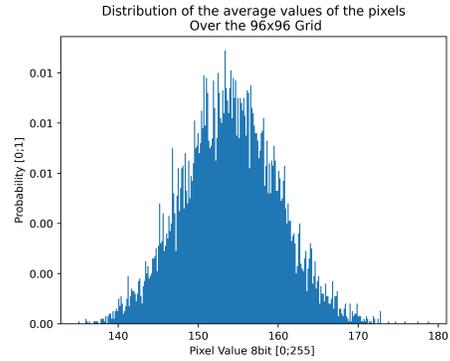

Figure 36: Distribution of the mean values of the pixels on the grid over the 5000 recorded frames. (Ensemble of the mean values as seen in Fig.(35))

6.3.1 Statistical Analysis of Recorded Irradiance

The experiment was conducted in a dark room to minimize the effects of stray light. Two types of recordings were created – with the laser ON, and the laser OFF. We will refer to the former as "Light Frames", and to the later as "Dark Frames". In the ideal scenario, the $8bit$ pixel value, representing the irradiance is constant, however, due to laser instability, dark current, camera electronic noise, etc., the irradiance values tend to oscillate around a mean, and over the observed time period of $66s$ even the mean is shown to drift. This can be seen in Fig.(37) and Fig.(38). The first figure is a good indicator of the irradiance drift present in our measurements.
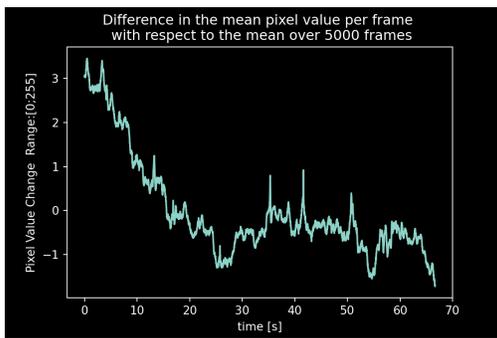
Figure 37: The figure shows the difference between the mean pixel value per frame and the mean over all 5000 frames.
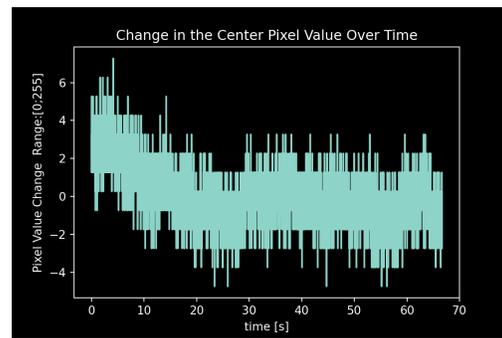
Figure 38: Change in the pixel value of the central pixel on the grid over all 5000 frames.

Before the statistical analysis model is explained, it is worth noting the relationship between the duration of the recording and the recording of interferograms while lens testing. When testing a microscope objective, for example, the camera captures 5 interferograms, each 90° out of phase compared to the previous one. The images are taken over a period of $1000ms$, where the

delay time between adjacent images is $200ms$. Therefore, the observed drift in Fig.(37), over the duration of 70 seconds is not relevant to us, unless we increase the delay time between adjacent image captures. For this purpose, it is worth conducting the statistical analysis in terms of "windows", where each window has a length of $1000ms$. Fig.(39) and Fig.(40) show how the value of the central pixel on the grid varies inside the 1st, 5th and 10th window for the "Light Frames" and "Dark Frames".
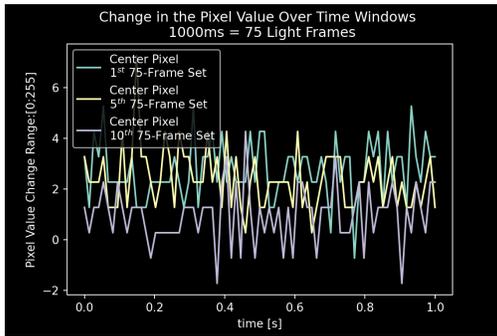
Figure 39: Laser ON: Change of the values of the central pixel with respect to the mean, shown for the 1st, 5th, and 10th measurement window.
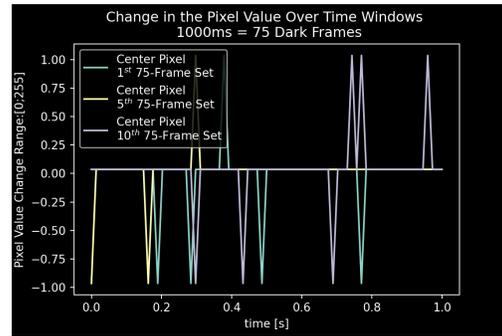
Figure 40: Laser OFF: Change of the values of the central pixel with respect to the mean, shown for the 1st, 5th, and 10th measurement window.

By knowing how the irradiance changes on average per pixel, during said timespan, we can model this change as a Gaussian noise with a certain mean and standard deviation which will be quantified at the end of this section. We can later implement this noise in the Monte Carlo simulations to find the uncertainty for the calculated Zernike coefficients and hence bring our simulation model closer to reality. A graph of the full process for the analysis of the recorded frames is provided in Fig.(41).
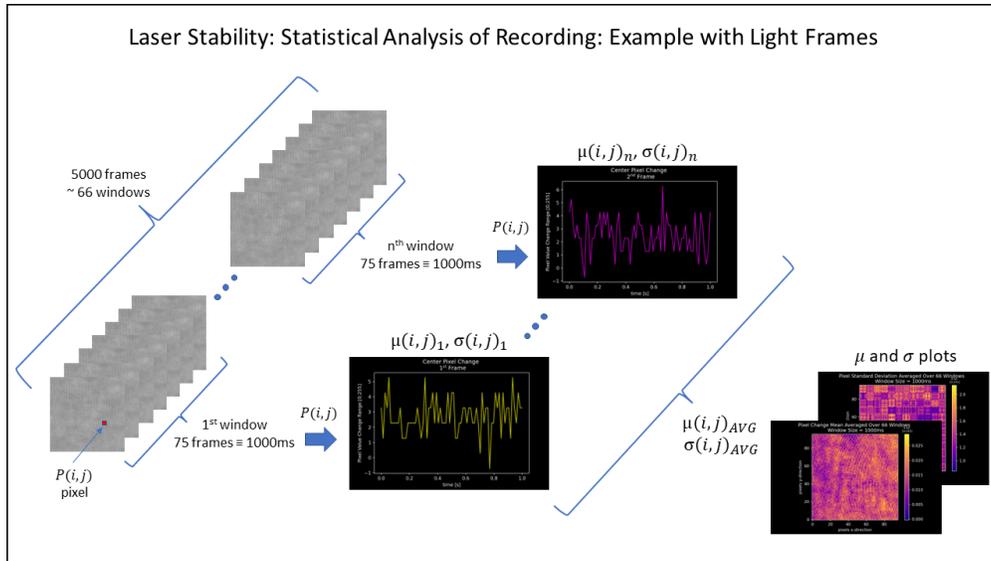
Figure 41: Overview of the procedure for conducting statistical analysis on recordings in order to acquire information about the laser instability, camera noise, shot noise, and irradiance drift. The figure shows an example analysis with the Laser ON (Light Frames).

We start by grouping frames together into time intervals we already referred to as "windows". A total of 5000 frames were recorded, where the camera recorded 75 frames per second. We group the frames into 66 windows and calculate the mean and standard deviation $\mu_n(i,j), \sigma_n(i,j)$ per individual pixel on the grid for each individual window n, as shown. For a selected pixel at a location (i,j), we then calculate the average mean and standard deviation $\mu_{AVG}(i,j), \sigma_{AVG}(i,j)$ from the individual means and standard deviations for all 66 windows $- \mu_{(1\ldots66)}(i,j), \sigma_{(1\ldots66)}(i,j)$ to arrive at the four plots shown in Fig.(42-45).
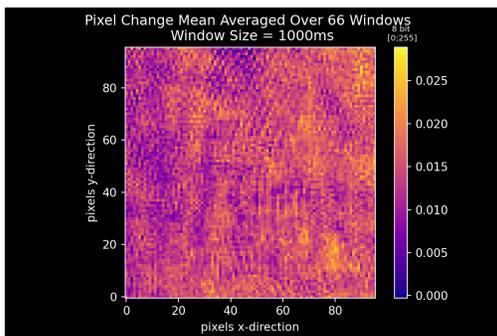
Figure 42: Laser ON: Mean of the change of the pixel values, averaged over the windows. Due to the irradiance drift, the mean is not 0.
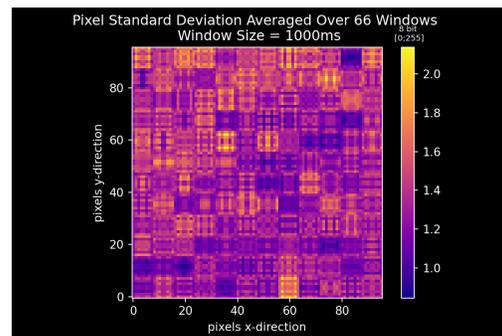
Figure 43: Laser ON: Standard deviation of the change of the pixel values, averaged over all the windows.
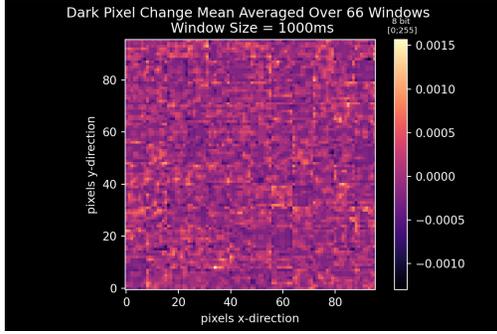
Figure 44: Laser OFF: Mean of the change of the pixel values, averaged over the windows. Due to the irradiance drift, the mean is not 0.
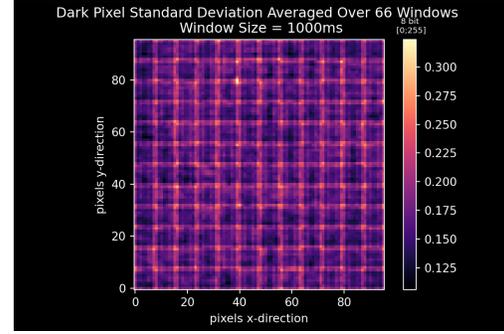

Figure 45: Laser OFF: Standard deviation of the change of the pixel values, averaged over all the windows.

In the standard deviation plot we recognize a beautiful box pattern which originates from how the pixels on the CMOS sensor are grouped together in 8x8 cells. The domino-like pattern of the standard deviation from the "Light Frames" may come from electronic noise in the cell, causing certain pixels from the cell to be more or less sensitive to incoming photons. In the "Dark Frames" case, the same pattern is not observed. However, a possible explanation is that for low light levels the difference in the sensitivity for individual pixels in the cell is negligible and a possible averaging for the whole cell takes place. Only at the edge pixels of the cell do we see a higher standard deviation/variance as there would some amount of uncertainty whether an incoming photon is registered in the current cell or the cell adjacent to the cell's edge.

The fine patterns shown above are an interesting thing to observe, however, we need to translate these results into useful parameters. The values for the mean and standard deviation are given in the $8bit$ value range $[0; 255]$. Therefore we can already see that the amount of noise corrupting our image, originating from the spontaneous change in the pixel values over time is small.

$$\mu_{AVG, plot} = \frac{0.01426}{256} = 5 * 10^{-5} \approx 0$$

$$\sigma_{AVG, plot} = \frac{1.345}{256} = 0.00525$$

From the two plots for the standard deviation, we can see that there are certain locations where the values are higher than the average value for the whole plot, namely, of the order of 2. The maximum recorded standard deviation is 2.138.

$$\sigma_{MAX, plot} = \frac{2.138}{256} = 0.00835$$

The difference between the average and the maximum value is therefore negligible and we needn't concern ourselves with implementing the domino-like pattern in the simulations. A simple noise model, where Gaussian noise with ($\mu = 0, \sigma = 0.00835$) is added to the pixels of the computer-generated interferograms would suffice.

6.4   Error Analysis of Background Irradiance I' and Fringe Modulation I''

In the PSI theory, the background irradiance I' and fringe modulation I'' are assumed to be constant. In section (4.2.2), it was discussed that many of the PSAs have built-in defenses against incorrect phase-stepping, for example. However, errors in terms of I' and I'' are often overlooked, where it is assumed that I' and I'' are constant over the course of the measurements. In this section we will show that this is not case for real data.

6.4.1   Obtaining I' and I'' Using System of Equations

In section (4.2), the following system of equations was given, where we explained that a minimum of 3 interferograms are required to solve for the 3 unknowns - $I'(x, y)$, $I''(x, y)$, and $\phi(x, y)$. An expression for $\phi(x, y)$ was provided in terms of a PSA inside an arctangent function, however, no expressions for finding $I'(x, y)$, $I''(x, y)$ were given. Obtaining them from the system of equations has been trivialized due to the assumption that they are constant over the recorded interferograms: $I'_1(x, y) = I'_2(x, y) = \ldots = I'_n(x, y)$, and $I''_1(x, y) = I''_2(x, y) = \ldots = I''_n(x, y)$, where n is the nth recorded interferogram. Using more than 3 interferograms should have the effect of overdetermining the system. For example, with 5 interferograms the system of equations can be written as:

$$
\begin{aligned}
I_1(x, y) &= I'(x, y) + I''(x, y)\cos[\phi(x, y) + \Delta_1] \\
I_2(x, y) &= I'(x, y) + I''(x, y)\cos[\phi(x, y) + \Delta_2] \\
I_3(x, y) &= I'(x, y) + I''(x, y)\cos[\phi(x, y) + \Delta_3] \\
I_4(x, y) &= I'(x, y) + I''(x, y)\cos[\phi(x, y) + \Delta_4] \\
I_5(x, y) &= I'(x, y) + I''(x, y)\cos[\phi(x, y) + \Delta_5]
\end{aligned}
\tag{70}
$$

Once the phase $\phi(x, y)$ has been calculated using the PSA of choice, if the assumption that I' and I'' remain constant over the recorded interferograms, then any two equations taken from the system shown in Eq.(70) could, in theory, be used to calculate them. This is not entirely correct for two reasons. One reason is that in the analytical expression for calculating I' and I'' division by 0 may occur in certain datapoints which causes the solution to explode. This can be seen in the denominator in Eq.(71), where the first and second equations were used to obtain an expression for I' and I''. Another reason is that both the background irradiance and fringe modulation depend on external factors such as the laser stability, for example, which is not necessarily constant over all measurements.

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \Delta_1]$$
$$I_2(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \Delta_2]$$

$$I'(x, y) = \frac{I_2 \cos(\phi + \Delta_1) - I_1 \cos(\phi + \Delta_2)}{\cos(\phi + \Delta_1) - \cos(\phi + \Delta_2)} = I'_{12} \tag{71}$$

$$I''(x, y) = \frac{I_2 - I_1}{\cos(\phi + \Delta_2) - \cos(\phi + \Delta_1)} = I''_{12}$$

We can calculate $I'(x, y)$ and $I''(x, y)$ from the following combinations of equations/recorded interferograms: $1\&2, 2\&3, \ldots, 4\&5$. Each time we calculate $I'(x, y)$ and $I''(x, y)$ from a set of two equations (i.e. 1&2), we have slightly different results due to errors. This can also be used to testify the hypothesis that in real interferograms the two coefficients cannot be constant but are rather approximately constant over the measurements. However, this leads to one major problem. Solving the system of equations given by Eq.(70) requires that we have less unknowns than we have equations. If we consider the background irradiance and fringe modulation to vary for each recorded interferogram - $I'_{1\ldots5}(x, y)$ and $I''_{1\ldots5}(x, y)$, then we arrive at 11 unknowns and 5 equations, which is a problem.

There is a way to bypass this hindrance by first assuming $I'(x, y)$ and $I''(x, y)$ are constant over the recorded interferograms and using a least squares solution to find the average value for all interferograms - $I'_{AVG}(x, y)$ and $I''_{AVG}(x, y)$. Then we can calculate $I'_{12, 23, \ldots, 45}(x, y)$ and $I''_{12, 23, \ldots, 45}(x, y)$ from pairs of equations and observe the difference between the average and individual value for the pair. First, the least squares solution will be explained.

$$\begin{bmatrix} I_1(x, y) \\ I_2(x, y) \\ \vdots \\ I_n(x, y) \end{bmatrix} = \begin{bmatrix} 1 & \cos(\phi + \Delta_1) \\ 1 & \cos(\phi + \Delta_2) \\ \vdots & \vdots \\ 1 & \cos(\phi + \Delta_n) \end{bmatrix} \begin{bmatrix} I' \\ I'' \end{bmatrix} = A\overrightarrow{I', ''} \tag{72}$$

The results for the calculated I'_{AVG} and I''_{AVG} from a set of 5 recorded interferograms with a 90° phase step between adjacent interferograms can be seen in Fig.(46, 47).
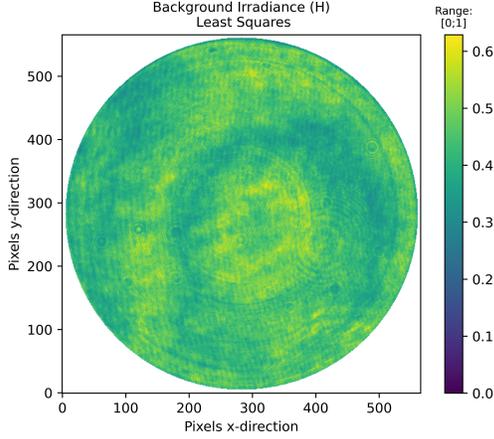
Figure 46: The average background iradiance I'_{AVG} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.
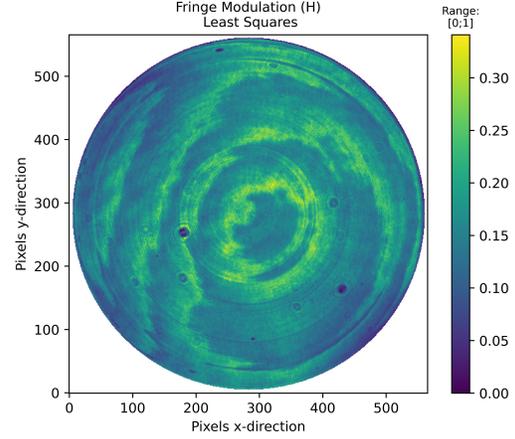

Figure 47: The average fringe modulation I''_{AVG} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.

We can see that for $n > 2$ the A matrix from Eq.(73) is not square invertible. Therefore, we can use its pseudoinverse to find I' and I''. Using all 5 interferograms as indicated in Eq.(70) gives us $I'_{AVG}(x, y)$ and $I''_{AVG}(x, y)$.

$$\overrightarrow{I',''} = (A^T A)^{-1} A^T \overrightarrow{I} \tag{73}$$

If we use two equations to compute I' and I'', the matrix is square invertible and we arrive at the formulas given by Eq.(71). One drawback to this technique is that $I'_{12, 23, \ldots, 45}(x, y)$ and $I''_{12, 23, \ldots, 45}(x, y)$ often contain values in plus and minus infinity. This can be noticed from the denominator of Eq.(71). For certain combinations of ϕ and Δ_1, Δ_2, the denominator can converge to 0. It is not entirely clear whether this is a limitation of the proposed technique or whether there is a deeper meaning behind. In chapter (7) the reasoning as to the physical explanation of this phenomenon will be given, however, this topic is left to future research. In the current analysis, the datapoints where the values exceed 1 or fall below 0 are set to 'NaN' and are therefore discarded from the calculations. An example of the calculated $I'_{12, 45}(x, y)$ and $I''_{12, 45}(x, y)$ has been shown in Fig.(48-51).
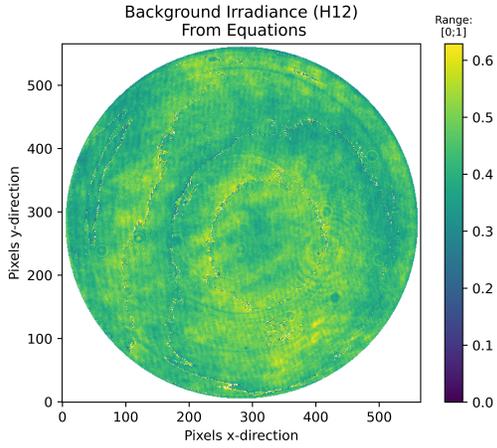
Figure 48: The average background iradiance I'_{12} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.
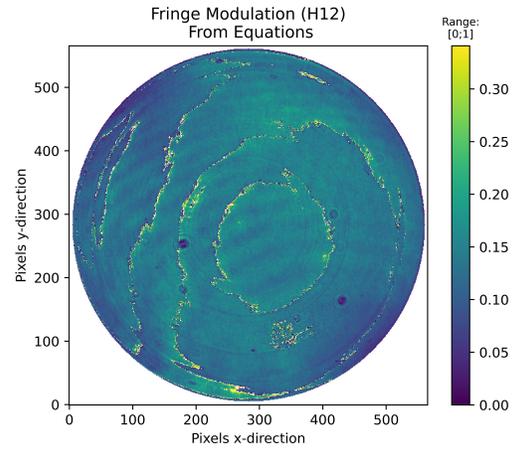

Figure 49: The average fringe modulation I''_{12} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.
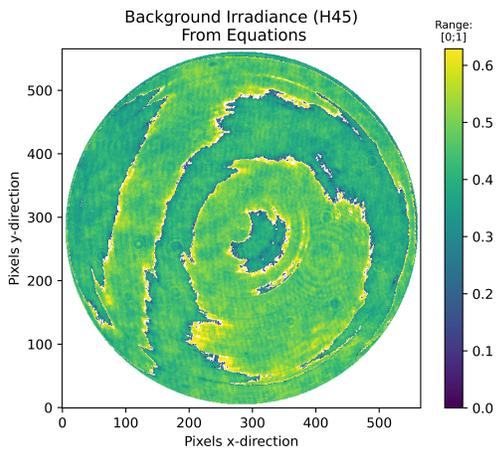

Figure 50: The average background iradiance I'_{45} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.
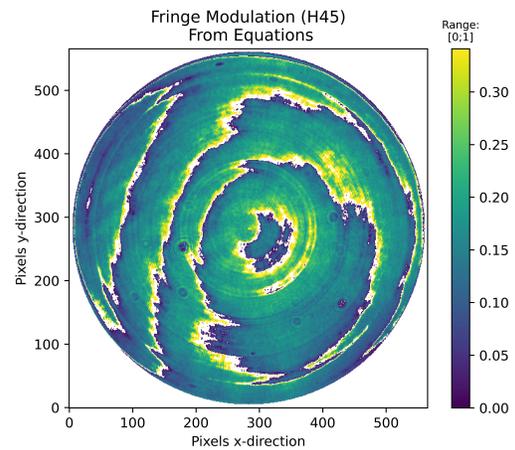

Figure 51: The fringe modulation from I''_{45} as calculated from Eq.(72). (H) indicates Schwieder-Hariharan PSA.

Instead of using the least squares method to find $I'_{AVG}(x, y)$ and $I''_{AVG}(x, y)$, we can also average the individual background irradiances and fringe modulations calculated from the pairs of equations - $I'_{12, 23, \ldots, 45}(x, y)$ and $I''_{12, 23, \ldots, 45}(x, y)$. The results are shown in Fig.(52, 53).
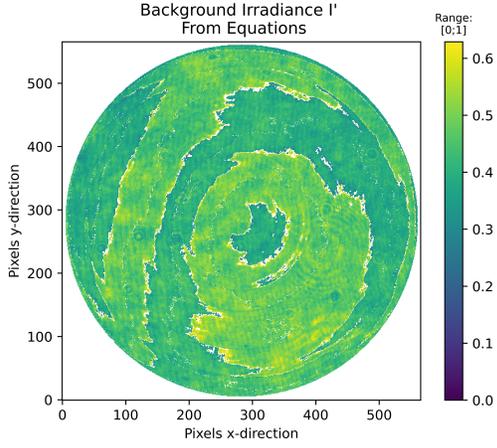
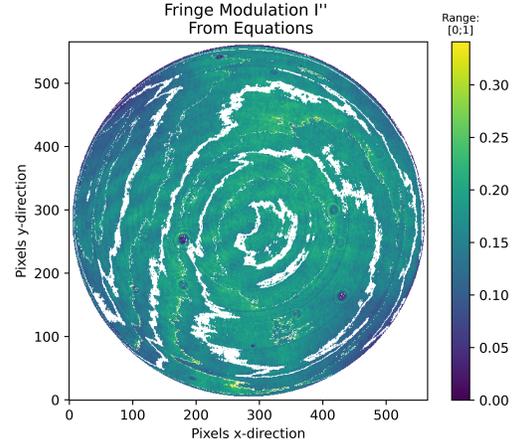Figure 52: The average background iradiance $I'_{AVG}(x, y)$.

Figure 53: The average fringe modulation $I''_{AVG}(x, y)$.

Using $I'_{12, 23, \ldots, 45}(x, y)$ and $I''_{12, 23, \ldots, 45}(x, y)$, we can compute the differences between the averaged coefficients and the coefficients per equation pair as shown in Eq.(74).

$$
\begin{aligned}
I'_{\delta 12} &= I'_{AVG} - I'_{12} & \qquad I''_{\delta 12} &= I''_{AVG} - I''_{12} \\
&\;\;\vdots \\
I'_{\delta mn} &= I'_{AVG} - I'_{mn} & \qquad I''_{\delta mn} &= I''_{AVG} - I''_{mn}
\end{aligned}
\tag{74}
$$

The individual errors/differences can then be averaged to find the average error per pixel for the background irradiance and fringe modulation. This is given in Eq.(75).

$$
\begin{aligned}
I'_{AVG\,Error}(x, y) &= \sum_{i=1}^{n-1} \sum_{j>n}^{n} I'_{\delta i j}(x, y) \\
I''_{AVG\,Error}(x, y) &= \sum_{i=1}^{n-1} \sum_{j>n}^{n} I''_{\delta i j}(x, y)
\end{aligned}
\tag{75}
$$

In Fig.(54, 55) the calculated average error for the background irradiance and the fringe modulation has been shown.
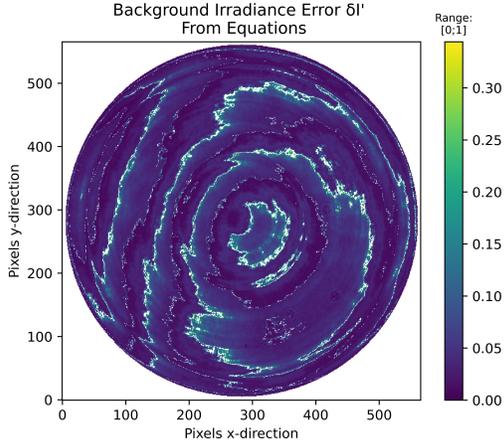
Figure 54: The average background iradiance error $I'_{AVG\,Error}(x,y)$ as calculated from Eq.(75). (H) indicates Schwieder-Hariharan PSA.
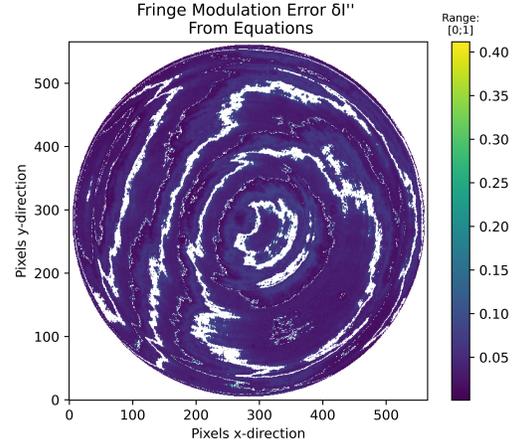

Figure 55: The average fringe modulation error $I''_{AVG\,Error}(x,y)$ as calculated from Eq.(75). (H) indicates Schwieder-Hariharan PSA.

6.4.2 Obtaining I' and I'' Using Least Squares Solution

In section (6.4.1) a method was shown where equations taken from the whole system of equations were grouped into pairs in order to find average values for the background irradiance and fringe modulation. We also considered the error arising from $I'(x,y)$ and $I''(x,y)$ not being constant over the set of recorded interferograms.

In this section we propose a different method, where we make use of the least squares solution shown in Eq.(72). Inserting all interferograms into the equation provides the average background irradiance and fringe modulation $I'_{AVG}(x,y)$ and $I''_{AVG}(x,y)$. The results thereof were shown in Fig.(46, 47). However, the variation of $I'(x,y)$ and $I''(x,y)$ across the interferograms is also of interest to us and we can modify Eq.(72) to obtain it.

By omitting certain interferograms in Eq.(72) we can observe how the result for the calculated $I'(x,y)$ and $I''(x,y)$ changes in terms of the result for all interferograms. Equation (76) is an example, where the third interferogram is omitted. Figures (56) and (57) show the effects of omitting the third interferogram on the calculated background irradiance and fringe modulation using least squares.

$$
\begin{bmatrix} I_1(x,y) \\ I_2(x,y) \\ I_4(x,y) \\ \vdots \\ I_n(x,y) \end{bmatrix} = \begin{bmatrix} 1 & cos(\phi + \Delta_1) \\ 1 & cos(\phi + \Delta_2) \\ 1 & cos(\phi + \Delta_4) \\ \vdots & \vdots \\ 1 & cos(\phi + \Delta_n) \end{bmatrix} \begin{bmatrix} I' \\ I'' \end{bmatrix} = A \overrightarrow{I'^{,''}_{(3)}} \tag{76}
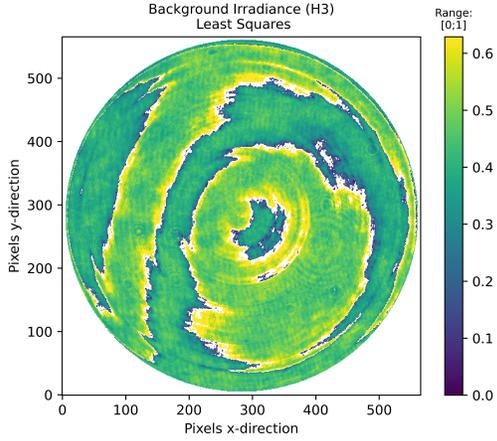$$

69

Figure 56: The average background iradiance with the 3rd interferogram omitted $I'_{(3)}(x, y)$ as calculated from Eq.(76). (H) indicates Schwieder-Hariharan PSA.
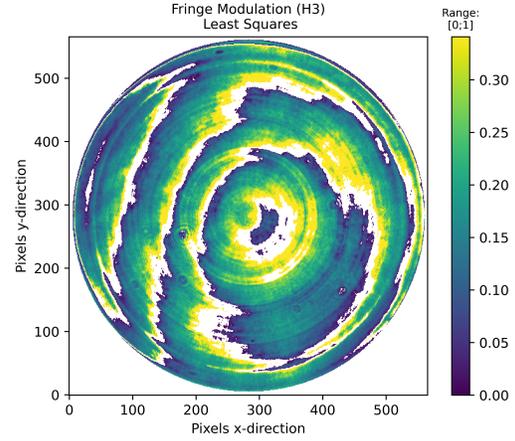

Figure 57: The average fringe modulation with the 3rd interferogram omitted $I''_{(3)}(x, y)$ as calculated from Eq.(76). (H) indicates Schwieder-Hariharan PSA.

Using the same approach as indicated by Eq.(74), we can estimate the average background irradiance error $I'_{AVG Error}(x, y)$ and the average fringe modulation error $I''_{AVG Error}(x, y)$. The results are plotted in Fig.(58, 59).
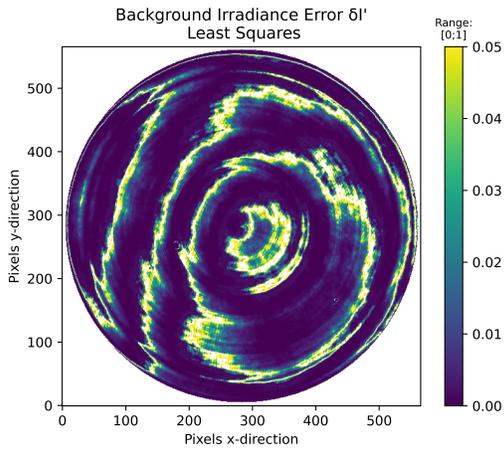

Figure 58: The average background iradiance error using least squares $I'_{AVG Error}(x, y)$.
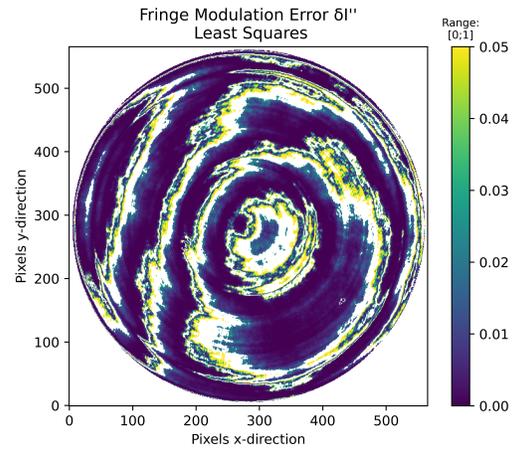

Figure 59: The average fringe modulation error using least squares $I''_{AVG Error}(x, y)$.

6.4.3 Interpreting the Results

In this section (6.4.1) and (6.4.2) two method were proposed. Here we will briefly compare the results obtained from both of them. The results are integral part of the Monte Carlo simulations used to evaluate the uncertainty.

First the average values for the background irradiance and fringe modulation can be compared for the two methods – pairing equations and the least squares approach. Let us first consider the I'_{AVG}. Despite Fig.(52) and (46) clearly appearing distinct, the mean value across the pupil ('NaN' values excluded) is close to identical. The value is calculated according to Eq.(77). The results are shown in Table (7).

$$I'_{GRID\ MEAN} = \frac{\sum\limits_{x=1}^{x_{size}} \sum\limits_{y=1}^{y_{size}} I'_{AVG}(x, y)}{x_{size} * y_{size}}$$

(77)

$$I'_{GRID\ MEAN\ ERROR} = \frac{\sum\limits_{x=1}^{x_{size}} \sum\limits_{y=1}^{y_{size}} I'_{AVG\ Error}(x, y)}{x_{size} * y_{size}}$$

Table 7: Results from the two techniques for the mean I', I'' and their error, first averaged over the measurements and then averaged over all datapoints on the grid.

	Grid Mean I'_{AVG}	Grid Mean I''_{AVG}	Grid Mean $I'_{AVG\ Error}$	Grid Mean $I''_{AVG\ Error}$
EQN Pairs	0.442	0.165	0.0225	0.0378
LSTSQ	0.438	0.164	0.0104	0.0135

In Table (7) one can notice that the mean values for the errors in the background irradiance and fringe modulation, albeit small compared to I'_{AVG} and I''_{AVG}, differ considerably between the two methods. Especially in the case of the fringe modulation error for the Equation Pairs (EQN Pairs) technique. It is unclear where this might come from but one hypothesis is that the Least Squares Solution (LSTSQ) may be less susceptible to outliers in the dataset and therefore more accurate.

However, we must also look at the distribution of the error values for I' and I'' for the two methods. This is the point at which the difference between the two becomes obvious and things start to seem peculiar. In Fig.(60-63) histograms of the error values have been shown. The results show that the error distribution does not follow a Gaussian distribution but rather a Gamma distribution. For the Least Squares method the distribution is exponential for $\delta I''$ and

71

somewhat exponential for $\delta I'$. The exponential distribution is a special case of the Gamma distribution, however, it is unclear what causes the different distributions and we leave this topic for future research. A discussion will follow in chapter (7).
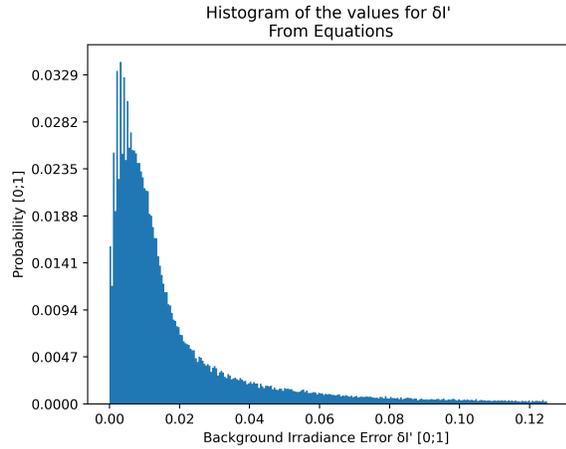

Figure 60: Histogram of the average absolute error values for the background irradiance $\delta I'$ for the EQN Pairs method.
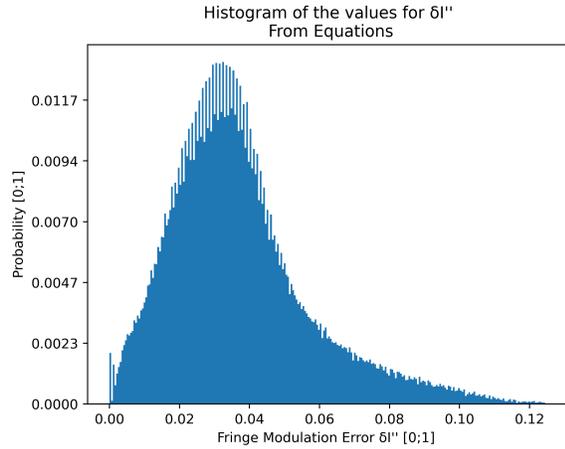

Figure 61: Histogram of the average absolute error values for the fringe modulation $\delta I''$ for the EQN Pairs method.
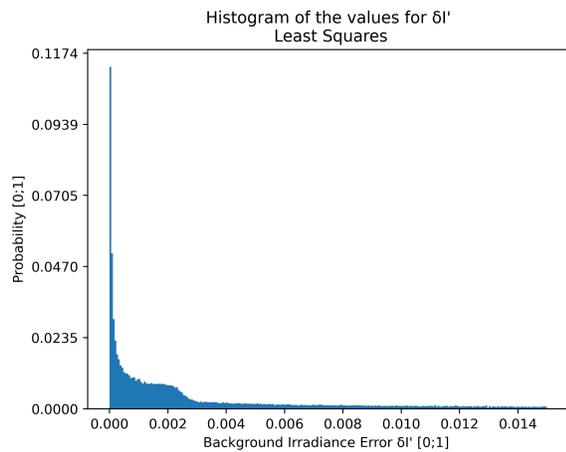

Figure 62: Histogram of the average absolute error values for the background irradiance $\delta I'$ for the LSTSQ method.

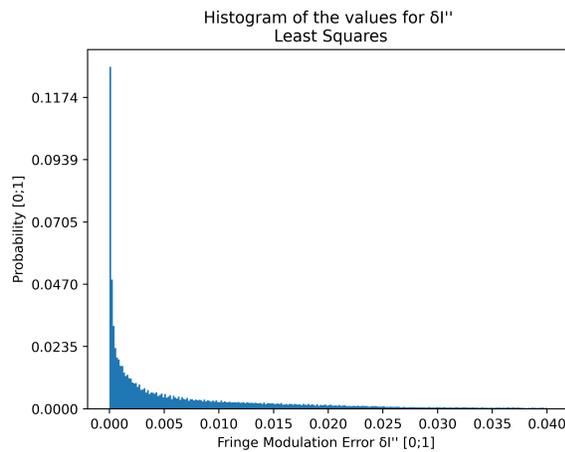Figure 63: Histogram of the average absolute error values for the fringe modulation $\delta I''$ for the LSTSQ method.

Beware that in Fig.(60-63) the shown distributions are not symmetrical, hence the calculated mean values given in Table (7) are the expected values and the middle point of the distributions is the medial, which is different from the mean in this case. Table (8) gives the calculated

mean and median points for the background irradiance error and the fringe modulation error. In section (6.5.3) they will be used to approximate the original distributions in order to generate realistic random noise.

Table 8: Calculated mean and median values for the $\delta I'$ and $\delta I''$ random noise.

	Mean $\delta I'$	Median $\delta I'$	Mean $\delta I''$	Median $\delta I''$
EQN Pairs	0.0225	0.011	0.0378	0.0342
LSTSQ	0.0104	0.00212	0.0135	0.00369

6.5 Error Analysis of Phase Stepping Δ

Phase stepping, as previously explained, is the process of introducing a know phase shift to the recorded spatial phase in terms of discrete steps - $\Delta_{1,\dots,n}$. Many PSAs, such as the two used – 4-step Wyant and 5-step Schwieder-Hariharan, require a constant phase step, for example 90°. In our Twyman-Green interferometer setup this is done by translating the flat reference mirror along the axis of the laser beam. A translation of 1λ corresponds to a phase shift between the beams in the reference and test arm of 4π due to our interferometer being double-pass.

The translation is done using a piezo actuator. However, no actuator can make a perfect translation each time. Therefore, in the datasheet one can find parameters, such as the actuator's repeatability over a given range of motion and the maximum tilt along the other two principal axes, expressed in terms yaw and pitch angles (assuming that translation happens along the z-axis).

In this section the phase step error due to the repeatability of the translation and the further introduced x- and y-tilt will be evaluated. The specifications of the used actuator are as follows:
1: The actuator has a repeatability of $1nm$ over the full range of motion of $15\mu m$, as per the datasheet.
2: The maximum tilt along the x- and y-axis expressed as yaw and pitch angles is $1\mu rad$, as per the datasheet.

6.5.1 Translation Error

A phase step of 90° for a wavelength of $\lambda = 405nm$ corresponds to a translation of $50.625nm$ along the z-axis, which is considerably less than the full range of motion. However, it is not advised to interpolate the repeatability over the much shorter motion range as we cannot be certain whether the positioning error is accumulated, in a linear fashion, over the distance of the motion or whether it is largely due to nonlinear effects occurring at the start and finish of the motion. Therefore, we will be extra strict and assume the repeatability to be equal to $1nm$ per phase step. Hence, each time the flat reference mirror phase-steps, we assume that there is a positioning error in the range $[-1nm; +1nm]$.

We can express the maximum translation error in terms of percentage or phase angle:

$$x_\% = \frac{1nm * 100}{50.625nm} = 1.97\%$$

$$\frac{x_{deg}}{1nm} = \frac{90°}{50.625nm}$$

$$x_{deg} = 1.78°$$

6.5.2 Tilt Error

The tilt error can be modeled by considering the flat reference mirror to be a plane with an arbitrary orientation, where the maximum values for the tilt along the x- and y-axis for the pupil can be calculated from the maximum yaw and pitch angles stated in the datasheet. From the tilt we can find how much certain points across the pupil are displaced along the z-axis. Hence, the datapoints are not going to experience the same phase shift as compared to the translation error case, where the error is equal for all datapoints.

Example: Given: Pupil magnification equal to 1. Entry pupil diameter of high-NA microscope objective as projected on the planar reference mirror is $3.5mm$. Both yaw and pitch angles are equal to $\alpha = 1\mu rad$. Figure (64) illustrates the example. The maximum x- and y-tilt are equal to:

$$sin(\alpha) \approx \alpha$$
$$x_{tilt} = y_{tilt} = 1\mu rad * \frac{3.5mm}{2} = 1.75nm$$
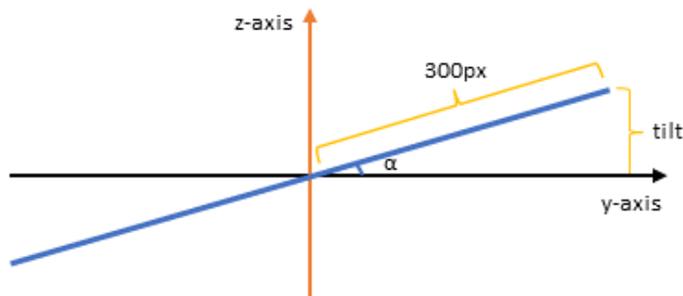

Figure 64: Illustration of how the maximum allowable x- or y-tilt is calculated given the example.

Knowing the maximum allowable tilt, we can define the range of tilt values for the x- and y-axis that our plane can assume in 3D space as $[-1.75nm; +1.75nm]$. Furthermore, we are interested in the z-displacement the added tilt causes. Hence, we can use the equation of a tilted plane in space Eq.(78):

$$
\begin{aligned}
z(x, y) &= a * x + b * y \\
a, b &\in [-1.75nm; +1.75nm]
\end{aligned}
\tag{78}
$$

In the Monte Carlo Simulation, a and b are random variables taken from the specified range. To obtain the phase step error caused by the random tilt, the result simply needs to be scaled to radians by multiplying by $\frac{4\pi}{\lambda}$. An example of a plane with random tilt orientation using Eq.(78) can be seen in Fig.(65).
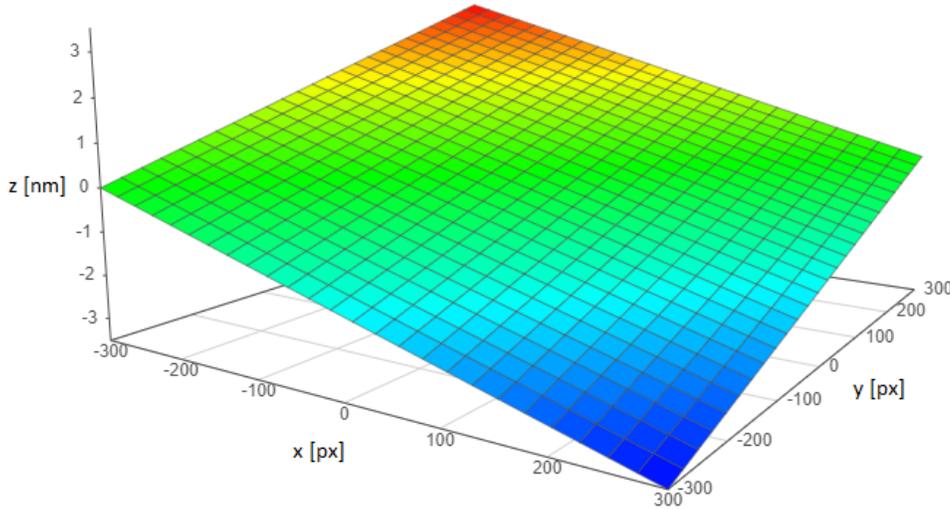

Figure 65: An example of a plane with tilt orientation $a, b = [-1, 75nm; +1.75nm]$ using Eq.(78).

6.5.3 Monte Carlo Uncertainty Evaluation

In section (6.1) we introduced and explained the working principle of the Monte Carlo Uncertainty Evaluation Model. Throughout the previous sections, different error sources were defined, introducing both systematic and random errors in the measurements. In this section we will explain how random perturbations were introduced into the simulation and will present the results for the calculated uncertainty per Zernike coefficient. The simulation model is quite computationally heavy, especially for a large number of iterations/samples, hence only the uncertainty results for the first 10 Zernike polynomials will be shown.

Table (9) contains an overview of the random errors that follow a Gaussian distribution and their respective mean and standard deviation in the Monte Carlo (MC) simulation. All random

errors, with the exception of the background irradiance $\delta I'$ and the fringe modulation $\delta I''$, were modeled using a Gaussian distribution, where the maximum permissible errors, as calculated in the previous sections, were chosen as the upper limit for the 3σ interval. That is to say, the maximum permissible tilt was calculated as $+1.75nm$, hence the magnitude of any random tilt introduced in the MC simulation must fall within the $[0\,;\,1.75nm]$ interval with 99.7% confidence. This means that the standard deviation for tilt would be: $\sigma = \frac{1.75nm}{3}$.

Table 9: Overview of the mean and standard deviation used to model the random noise that follows a Gaussian distribution.

Random Errors	μ	σ
x-tilt	0.0 [nm]	0.587 [nm]
y-tilt	0.0 [nm]	0.587 [nm]
z-translation	0.0 [nm]	0.333 [nm]
Gaussian	0 [8bit]	2.138 [8bit]

This approach of modelling random errors doesn't apply to $\delta I'$ and $\delta I''$ as their distributions are not Gaussian. Due to the time constraints on the project, a distribution of the exact shape couldn't be introduced in the model. However, we were able to approximate the shape using a Gamma distribution with certain shape parameter k and scale parameter θ. The values of the used distribution parameters are given in Table(10). A side-by-side comparison of the original and the approximated Gamma distributions for $\delta I'$ and $\delta I''$ can be seen in Fig.(66-69).

Table 10: Overview of the shape parameter k and scale parameter θ used to model the random noise that follows a Gamma distribution. Two cases are shown: the random noise as calculated from the EQN Pairs approach, and the random noise as calculated from the least squares solution.

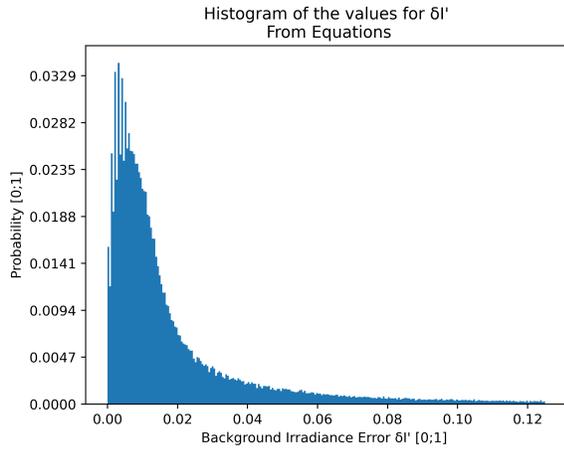| Random Errors	k	θ
$\delta I'$ EQN Pairs	1.3	0.0145
$\delta I''$ EQN Pairs	3	0.0133
$\delta I'$ LSTSQ	0.5	0.05
$\delta I''$ LSTSQ	0.5	0.019

Figure 66: Distribution of the average absolute error values for the background irradiance $\delta I'$ for the EQN Pairs method.
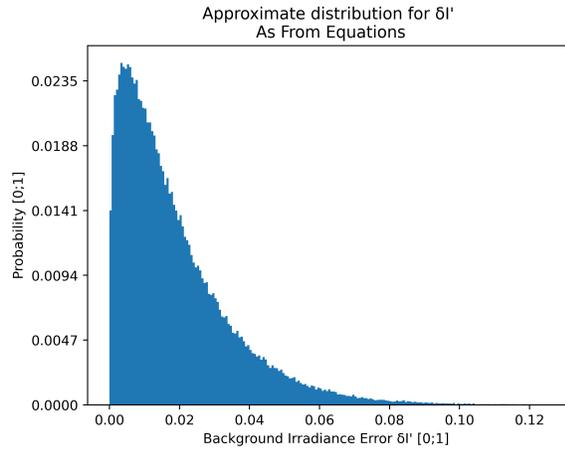

Figure 67: Approximated distribution of the average absolute error values for the background irradiance $\delta I'$ for the EQN Pairs method.
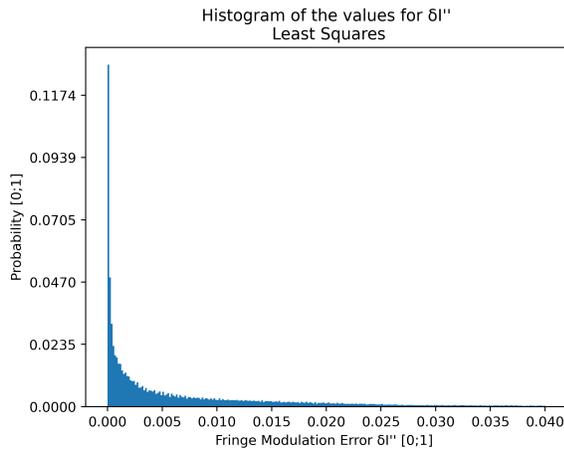

Figure 68: Distribution of the average absolute error values for the fringe modulation $\delta I''$ for the LSTSQ method.
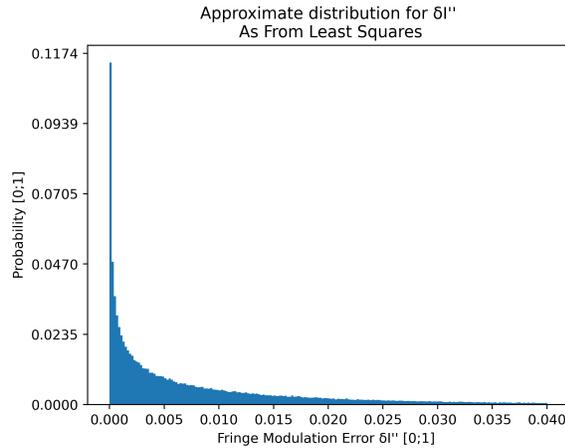
Figure 69: Approximated distribution of the average absolute error values for the fringe modulation $\delta I''$ for the LSTSQ method.

Results of the Monte Carlo Simulations

The model uses a set of Zernike polynomials, the average background irradiance $I'_{AVG}(x, y)$, the fringe modulation $I''_{AVG}(x, y)$, and the MSF to generate random test interferograms which are then evaluated. The four parameters come from evaluated real interferograms after correcting for the aberrations in the spherical mirror. In order to evaluate the measurement uncertainty

77

per Zernike coefficient, the previously found random perturbations have been inserted into the model described in Fig.(23).

Multiple sets of calculated Zernike coefficients are obtained from the multiple sets of test interferograms. We then form datasets per individual Zernike polynomial. For example, all calculated x-tilt coefficients are stored in a single array, all y-tilt coefficient, all defocus coefficients, and so on.

The datasets can later be evaluated using a technique of choice. Here, we have used "box and whiskers" plot, also known as "boxplot", as it provides better insight than simply calculating the standard deviation of the dataset. Boxplots use the median to split the dataset in two subsets. The median of the complete dataset is called the second quartile Q_2. The medians of the two subsets are referred to as Q_1 and Q_3. The range between Q_1 and Q_3 is called the interquartile range (IQR) and is used to estimate whether a point is an outlier or not. In the boxplots of the results the outliers are marked by a dot. The criterion is as follows:

$$low\ outlier\ if\ <\ Q_1 - 1.5 * IQR$$
$$high\ outlier\ if\ >\ Q_3 + 1.5 * IQR$$

Apart from the random perturbations and systematic errors, two other problems are common to occur. Both of them relate to the size, shape, and location of the analysis mask used to define the pupil on the screen. We will refer to the process of obtaining the analysis mask as "edge detection". The principle is explained in Appendix (B) and makes use of the interferometric visibility $\gamma(x, y)$. In the current project circular masks were used. The two problems at hand are namely the possibly incorrect location of the analysis mask the possible cropping of the few outermost pixels of the pupil.

Using the edge detection approach outlined in Appendix (B), it was observed that for a 1217px X 1936px grid, the center of the analysis mask coincides with the exact center of the pupil with an uncertainty of $\pm 1 px$ in the x and y direction. The diameter of the analysis mask tends to slightly underestimate the actual diameter of the pupil with an error of up to 4px.

The results provided by the boxplots can be grouped in 3 sets. The first set corresponds to Fig.(70-73), where only random perturbations are present and no errors arise from the analysis mask. The calculated uncertainties are lower than the uncertainties in sets 2 and 3.

In set 2, Fig.(74-77), we introduce the possibility for the analysis mask to be taken equal or smaller than the diameter of the pupil. The value range is between $[0; 4 px]$, which is discrete as compared to the value range of the random perturbations which are continuous. Due to the time constraint on the project, the exact effects of combining continuous and discrete random variables on the MC simulation haven't been studied in detail and are left for future research. Nevertheless, we believe that the uncertainty range, as calculated through the boxplots is a good approximation.

In set 3, Fig.(78-81), an additional feature is introduced, where the analysis mask may be off-center by $\pm 1 px$ with respect to the center of the pupil. This set is considered to be the most complete one from the 3 and can be used to determine an uncertainty budget for the measured coefficients. It can be observed that certain Zernike coefficients are more affected by the errors, especially the ones caused by an incorrect analysis mask. This is because incorrectly truncating a polynomial that has a high gradient towards the edge of the pupil, makes it possible for the polynomial to be represented in terms of a combination of lower order ones. [33]

The 3 sets are shown in the indicated order as it helps to visualize the impact the correct edge detection and centering of the analysis mask have on the final results. It is also worth noting that various techniques for edge detection exist. Most are used for object recognition in computer vision and do not have a direct connection to interferometry. The edge detection used in this project is based on the interferometric visibility $\gamma(x, y)$ and is explained in more detail in Appendix (B).

Set 1: Realistic Perturbation From EQN Pairs

Figure 70: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Set 1.

Figure 71: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Set 1. Corrected for systematic errors.

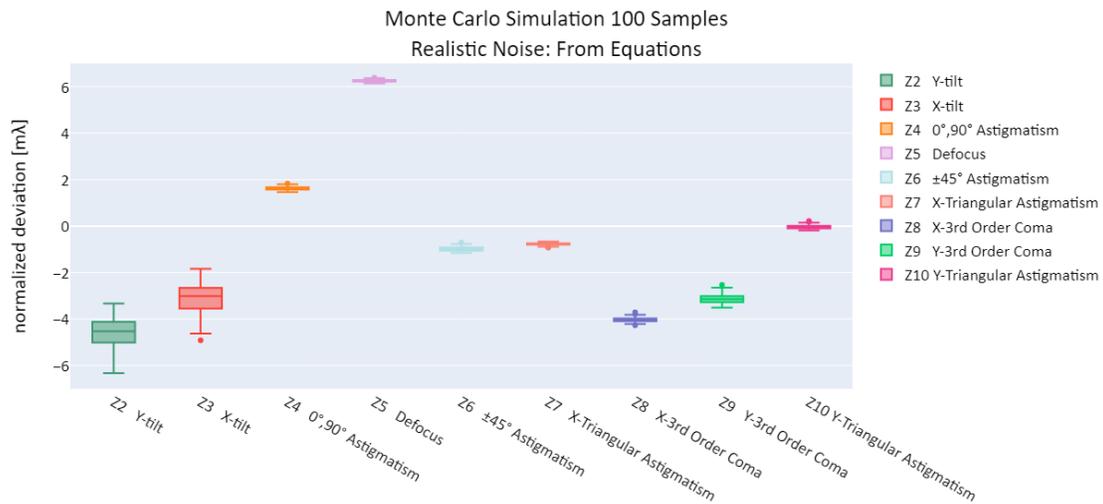## Set 1: Realistic Perturbation From LSTSQ

Figure 72: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Set 1.
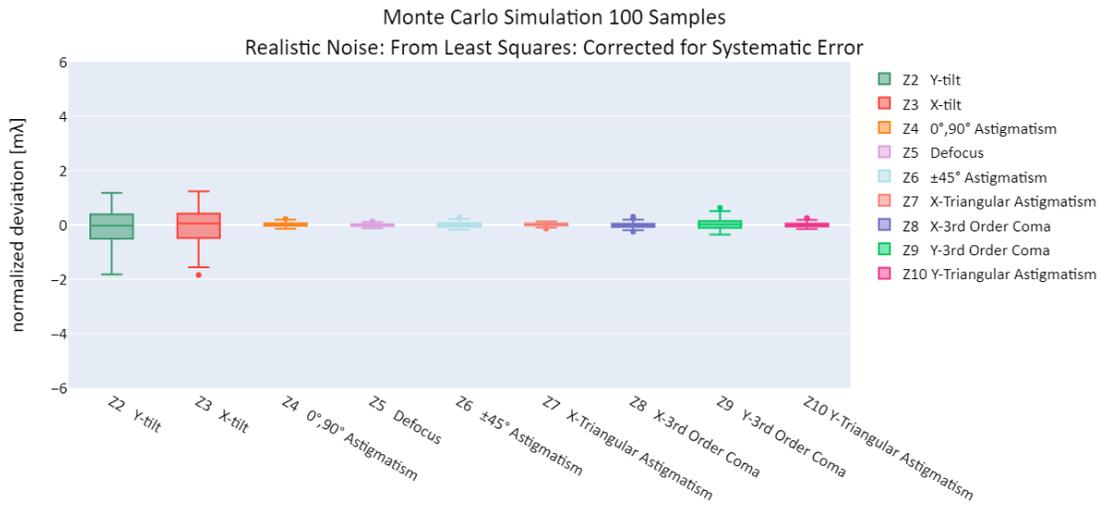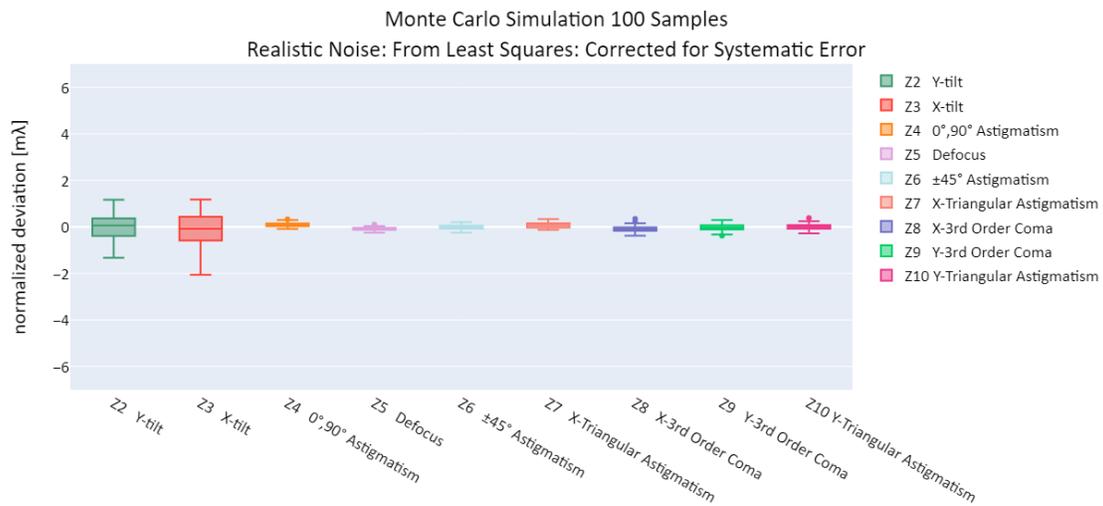
Figure 73: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Set 1. Corrected for systematic errors.
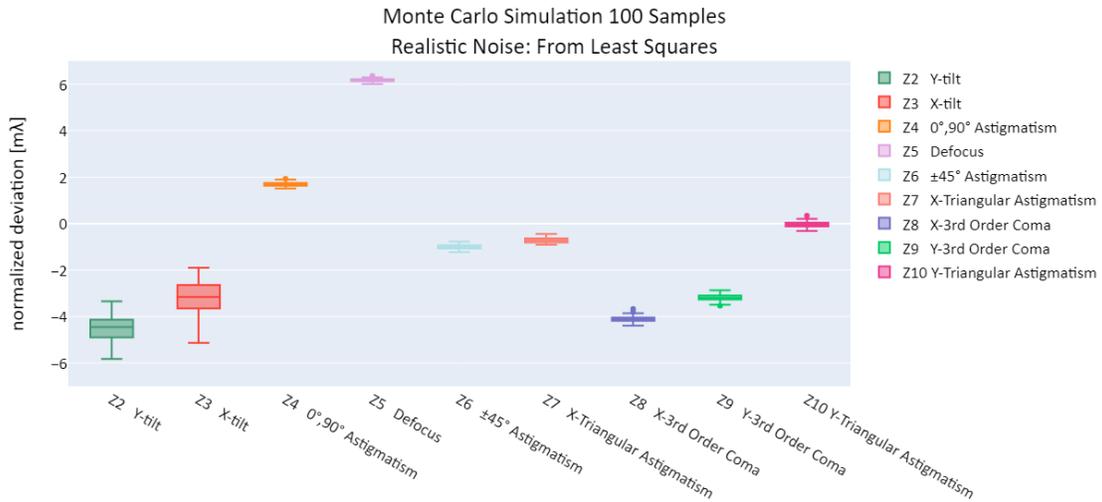
Set 2: Realistic Perturbation From EQN Pairs, $[0; 4px]$ Diameter Crop

Figure 74: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations and $[0; 4px]$ cropped pupil mask. Set 2.

Figure 75: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations and $[0; 4px]$ cropped pupil mask. Set 2. Corrected for systematic errors.

Set 2: Realistic Perturbation From LSTSQ, $[0; 4px]$ Diameter Crop

Figure 76: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations and $[0; 4px]$ cropped pupil mask. Set 2.

Figure 77: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations and $[0; 4px]$ cropped pupil mask. Set 2. Corrected for systematic errors.

Set 3: Realistic Perturbation From EQN Pairs, $[0; 4px]$ Diameter Crop and $\pm 1px$ Mask Off-Center

Figure 78: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Pupil mask cropped $[0; 4px]$ and off-center $\pm 1px$. Set 3.
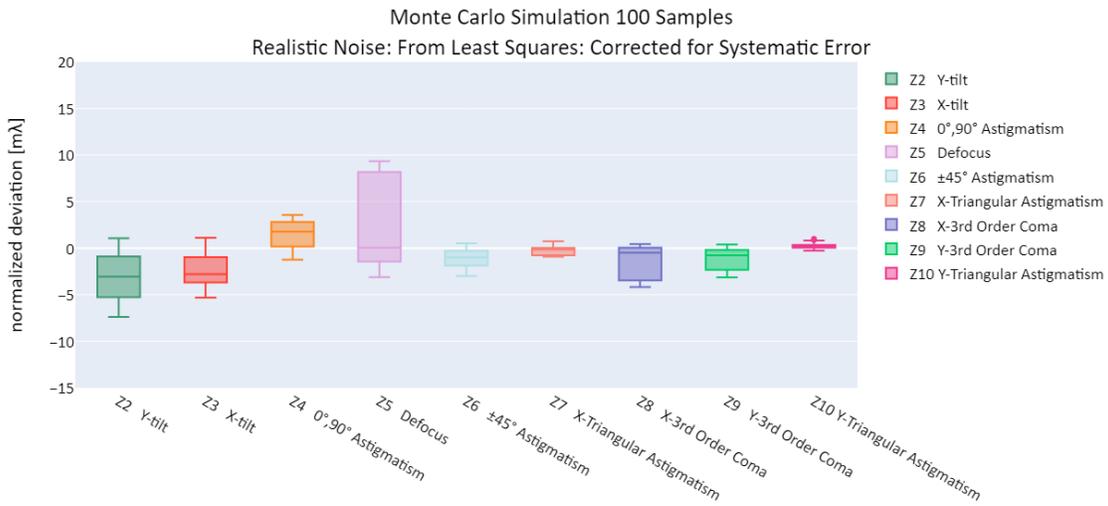
Figure 79: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Pupil mask cropped $[0; 4px]$ and off-center $\pm 1px$. Set 3. Corrected for systematic errors.

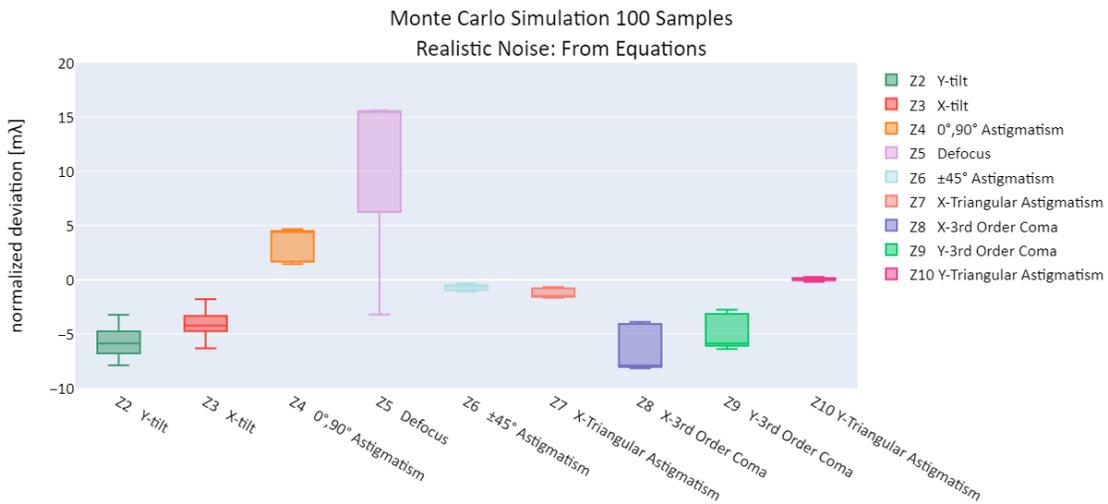**Set 3: Realistic Perturbation From LSTSQ, $[0; 4px]$ Diameter Crop and $\pm 1px$ Mask Off-Center**
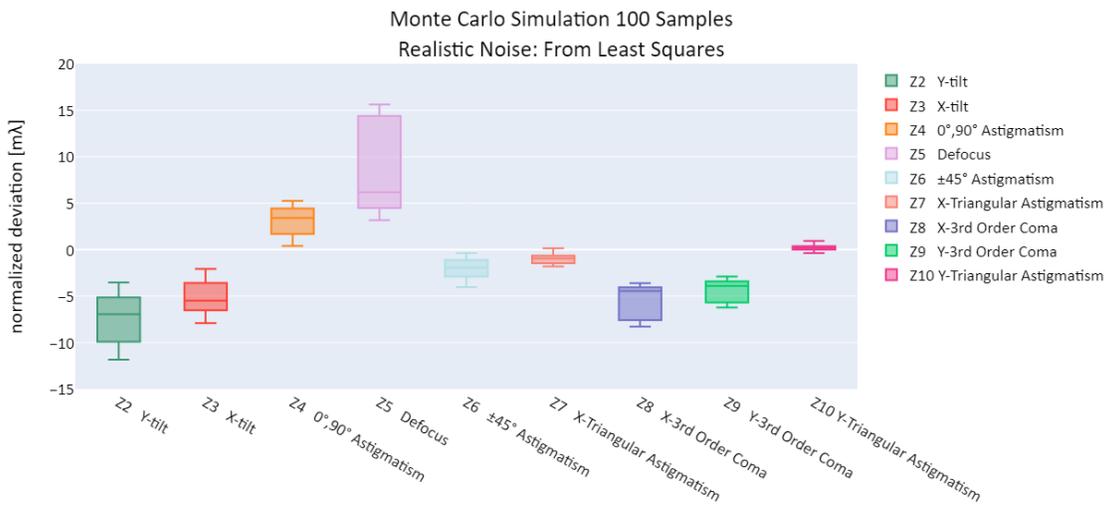
Figure 80: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Pupil mask cropped $[0; 4px]$ and off-center $\pm 1px$. Set 3.
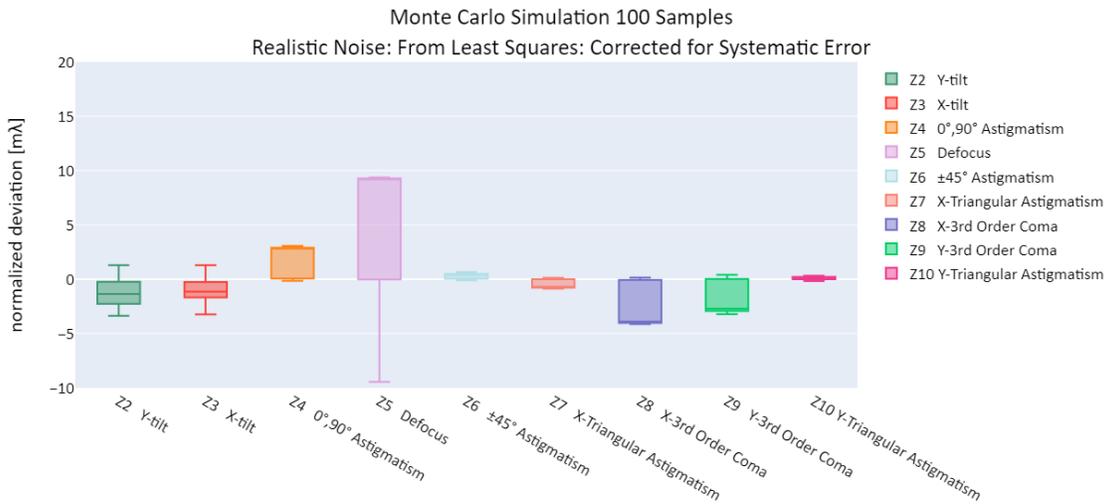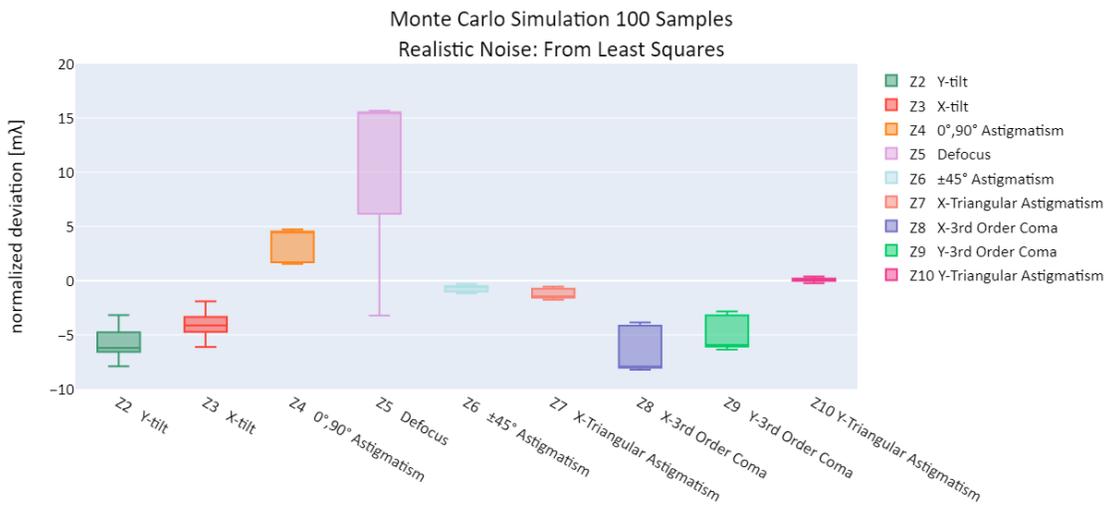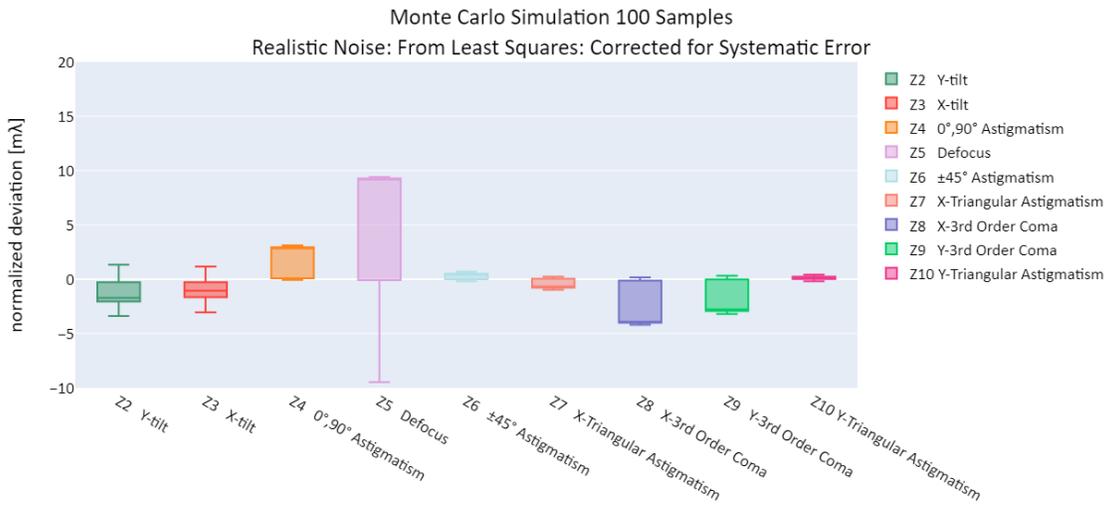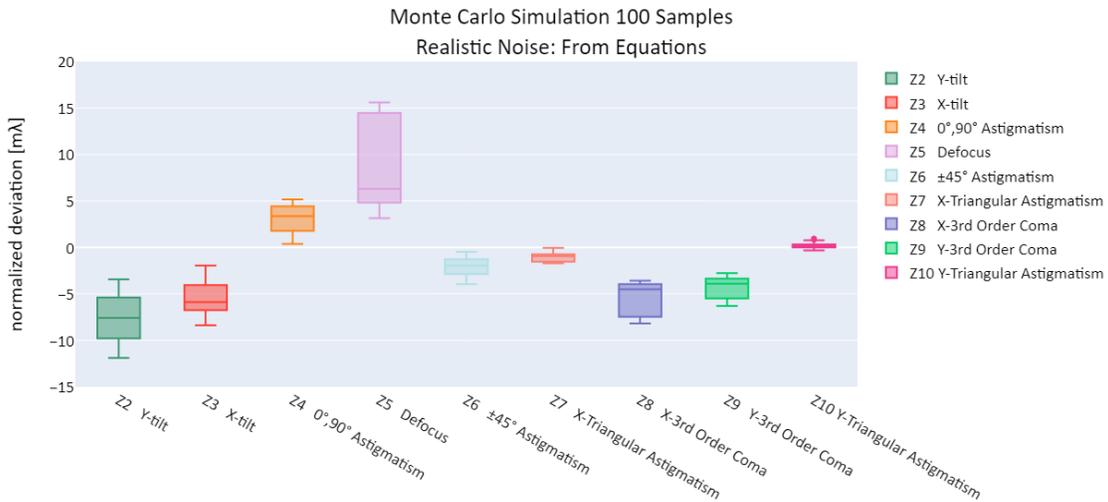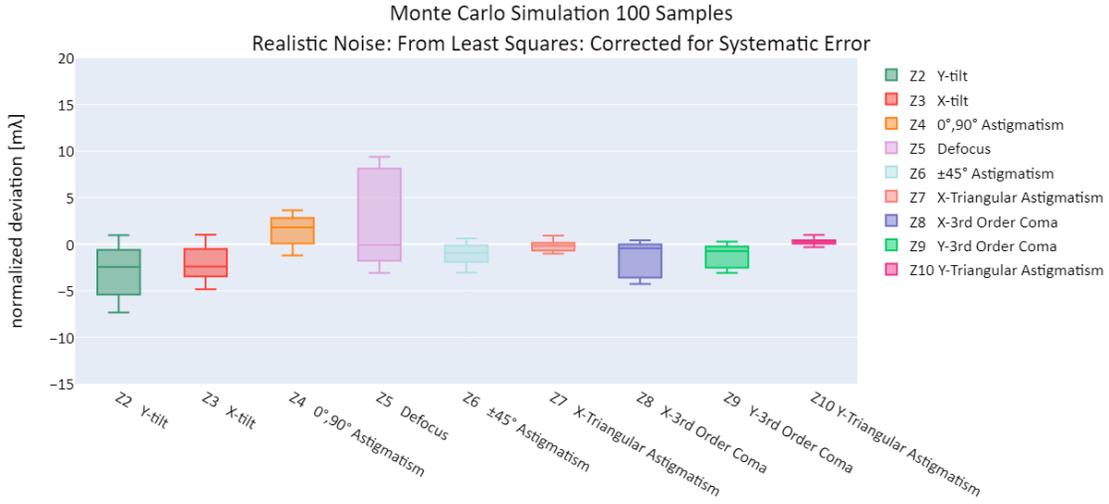
Figure 81: Boxplots showing the uncertainty per calculated Zernike coefficient for realistic perturbations. Pupil mask cropped $[0; 4px]$ and off-center $\pm 1px$. Set 3. Corrected for systematic errors.

In the boxplots one can notice that the uncertainty range varies greatly between the Zernike polynomials. Therefore, it makes sense that the individual uncertainties of the coefficients should be considered. Still, we can use the "worst case" (Z_5 Defocus) and the "average of the uncertainty range" of the boxplots as a rough estimate of the overall measurement uncertainty. By taking the uncertainty range of defocus, and the average of the uncertainty ranges from Fig(78) and Fig(81), and assuming symmetry of the dataset, we get the following approximate of the uncertainty budget.

$$u_{\text{worst} \atop \text{case}}(Z) = \pm 9.40 m\lambda$$

$$u_{avg}(Z) = \pm 2.33 m\lambda$$

7 Discussion & Recomendations

This section discusses the strengths and weaknesses of the proposed Monte Carlo Uncertainty Evaluation Model, as well as some aspects, where the observed phenomena could not be explained. Due to multitude of factors that contribute to the uncertainty, it was necessary to have a broad scope for the research. This meant that certain features of the uncertainty evaluation process couldn't be studied in detail.

Throughout the project we considered various errors and error sources and their influence on interferometric measurements. Different ways for their identification and compensation were discussed, as well as their origins in the system. Finally, a model for the evaluation of the uncertainty per calculated Zernike coefficient was proposed. The model requires statistical information of the error sources which can be extracted from the interferograms themselves, be taken from the datasheets of components used in the interferometer, or be found through additional measurements, such as the one described in section (6.3).

First, we will address the potential limitations and suggestions for future research regarding the three primary algorithms: the phase stepping algorithm, the phase-unwrapping algorithm, and the least squares fit algorithm. Then, we will discuss the two proposed uncertainty evaluation models and possible improvements that can be made. The chapter concludes with a showcase of the observed phenomena to which a full explanaition couldn't be provided.

Phase Step Algorithms

One aspect that requires further attention is the error compensation ability of PSAs. The existing literature on the topic is quite broad, however, much remains to be done. The compensating abilities of PSAs are often evaluated by the algorithm designers from a purely analytical standpoint. On other instances simulation models are used to estimate the algorithm's performance. However, the used models may lack realism and are often not tested against real measurement data. Furthermore, many authors tend to overlook the importance of the physical meaning behind certain errors and instead describe them in terms of a mathematical expression. This is an area which can be expanded on and can benefit from a single unified convention.

Phase Unwrapping

Phase unwrapping algorithms, on the other hand, can be described using the theory presented in section (4.3). However, the uncertainty they introduce to the unwrapped data remains a largely untouched topic. Lens testing doesn't necessarily imply strict requirements for the PU algorithm, as compared to the requirements for InSAR measurements. Nevertheless, a case study of prevalent PU errors in lens testing interferometry is worth considering. Throughout this project only two PU algorithms were used, hence, no definite conclusions for their performance can be made.

Polynomial Fitting

Another thing to consider is the polynomial fitting of high order terms. A major shortcoming of the Monte Carlo Uncertainty Evaluation Model from Fig.(23) is that it is computationally heavy due to the algorithm having to evaluate the Zernike polynomials pointwise. It is possible to use a Cholesky decomposition to speed up the process by reducing the number of datapoints over which the Zernike polynomials need to be fit. However, the effects this operation might have on higher order polynomials, where the discrete data sampling is insufficient, requires more attention.

Analytical Uncertainty Evaluation Model

Due to the time constraints on the project, the analytical uncertainty evaluation model defined in section (5.3.1) couldn't be tested. It remains a theoretical concept that requires further ex-

perimentation to prove or disprove it. It is likely that the greatest challenge would be to acquire trustworthy information on the uncertainty in the calculated phase map caused by individual error sources. One such error source is the phase unwrapping.

Monte Carlo Uncertainty Evaluation Model

A major advantage of the Monte Carlo method is that it is straightforward and easy to implement. The downsides, however, are that it too requires extensive knowledge of the statistical distribution of random errors and is also computationally heavy. The greatest challenge thus far was to obtain information on the random errors. These include the phase shift error, the light-source instability, the background irradiance I', and the fringe modulation I''.

Notice that the effects of electronic noise, dark current, laser instability, shot noise, etc., were bunched together in a single noise perturbation. Unlike the analytical model, the MC one can take their combined effect as an input. This can be considered both a strength and a weakness. On one hand, it is not necessary to find the individual random error signals and consider their correlation. On the other hand, we cannot know for certain which error source contributes the most to the overall instability. This means that troubleshooting cannot happen unless we investigate the individual error sources.

The background irradiance I' and the fringe modulation I'' are also points of concern. The PSAs assume they are constant over the recorded interferograms in order to solve a system of equations for 3 unknowns, but in reality they don't necessarily need to be constant which is a not too well researched topic. The physical explanation of I' and I'' is provided in textbooks using Eq.(5), however, the generalized example of two-beam interference may appear more complex when investigating a real interferometer due to multiple reflections, scattering, polarization effects, etc.
The analysis conducted on I' and I'' testifies that the textbook explanation is lacking as we observe discontinuities, evident from Eq.(71) and Fig.(56,57). What is peculiar is that these discontinuities bear resemblance to the interfaces separating light and dark fringes in the interferograms. A possible explanation is that the visibility/data modulation at the interface between light and the dark fringes may be undefined, hence why for a small number of equations/interferograms this causes such undefined regions to appear.

References

[1] Global and United States Optical Lens Groover Market Report & Forecast 2022-2028, 2022

[2] E. Hecht, "Optics", Essex: Pearson, 2017

[3] D. Malacara, "Optical Shop Testing", 2nd Edition, New Jersey: John Wiley & Sons Inc, 2007

[4] J. Wyant, "Use of an ac heterodyne lateral shear interferometer with real-time wavefront correction systems", Applied Optics, 14, 2622-2626, p. 5, 1975

[5] J. Wyant, C. L. Koliopoulos, B. Bhushan and O. E. George, "An optical profilometer for surface characterization of magnetic media", ASLE transactions, vol. 27, no. 2, pp. 101-113, 1984

[6] W. T. Welford, "Aberrations of Optical Systems", Bristol: J. W. Arrowsmith, 1986

[7] S. Kim , J. Jeon, K. Yangjin, S. Naohiko, and M. Mamoru, "Design and Assessment of Phase-Shifting Algorithms in Optical Interferometer", International Journal of Precision Engineering and Manufacturing-Green Technology, p. 25, 2022

[8] A. J. Lewis, "Absolute length measurement using multiple-wavelength phase-stepping interferometry", University of London, Imperial College of Science, Technology, and Medicine, Department of Physics and Applied Optics Group, London, 1993

[9] L. L. Deck, "Suppressing phase errors from vibration in phase-shifting interferometry", Applied Optics, vol. 48, no. 3948-3960, p. 13, 2009

[10] K. Freischlad and C. L. Koliopolous, "Fourier description of digital phase-measuring interferometry", Journal of the Optical Society of America A, vol. 7, no. 542-551, p. 10, 1990

[11] K. G. Larkin and B. F. Oreb, "Design and assessment of symmetrical phase-shifting algorithms", Journal of the Optical Society of America A, vol. 9, no. 1740-1748, p. 9, 1992

[12] J. E. Greivenkamp, "Generalized Data Reduction For Heterodyne Interferometry", Optical Engineering, vol. 23, no. 234350, p. 3, 1984

[13] C. J. Morgan, "Least-squares estimation in phase-measurement interferometry", Optics Letters, vol. 7, no. 368-370, p. 3, 1982

[14] K. Hibino, B. F. Oreb and P. S. Fairman, "Wavelength-scanning interferometry of a transparent parallel plate with refractive-index dispersion", Applied Optics, vol. 42, no. 3888-3895, p. 8, 2003

[15] K. Hibino, B. F. Oreb, P. S. Fairman, and J. Burke, "Simultaneous measurement of surface shape and variation in optical thickness of a transparent parallel plate in wavelength-scanning Fizeau interferometer", Applied Optics, vol. 43, no. 6, p. 9, 2004

[16] J. B. Hayes, "Linear Methods of Computer Controlled Optical Figuring", The University of Arizona, Tucson, Arizona, USA, 1984

[17] K. Hibino, "Error-Compensating Phase Measuring Algorithms in a Fizeau Interferometer", Optical Review, vol. 6, no. 6, p. 10, 1999

[18] Y. Kim, K. Hibino, N. Sugita and M. Mitsuishi, "Design of phase shifting algorithms: fringe contrast maximum", Optics Express, vol. 22, no. 15, p. 11, 2014

[19] P. J. De Groot, "Correlated errors in phase-shifting laser Fizeau interferometry", Applied Optics, vol. 53, no. 19, p. 9, 2014

[20] C. Zuo, J. Qian, S. Feng, W. Yin, Y. Li, P. Fan, J. Han, K. Qian, and Q. Chen, "Deep learning in optical metrology: a review", Light: Science & Applications, vol. 11, no. 39, p. 54, 2022

[21] R. Shuvolov, "Algorithm of the Green's Function Method for Two-Dimensional Phase Unwrapping", Journal of Mining Sciences and Technologies, vol.2, ID: 528.856.044.1, p. 13, 2011, (in russian)

[22] D. Ghiglia and M. Pritt, "Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software", Michigan, USA: John Wiley & Sons Inc., 1998

[23] A. Baldi, F. Bertolino, and F. Ginesu, "Phase Unwrapping Algorithms: A Comparison", in "Interferometry in Speckle Light: Theory and Applications", Lausanne, Switzerland, 2000

[24] T. J. Flynn, "Consistent 2-D phase unwrapping guided by a quality map", in International Geoscience and Remote Sensing Symposium, Lincoln, USA, 1996

[25] D. Rosenfield, N. Ching and M. Braun, "Two-Dimensional Phase Unwrapping Using a Minimum Spanning Tree Algorithm", IEEE Transactions of Image Processing, vol. 1, no. 3, p. 11, 1992

[26] M. Takeda and T. Abe, "Phase unwrapping based on maximum cross-amplitude spanning tree algorithm: a comparative study", in Interferometry VII: Techniques and Analysis, San Diego, 1995

[27] J. Gierloff, "Phase Unwrapping by Regions", in "31st Annual Technical Symposium on Optical and Optoelectronic Applied Sciences and Engineering", San Diego, USA, 1987

[28] K. M. Hung and T. Yamada, "Phase unwrapping by regions using least-squares approach", Optical Engineering, vol. 37, no. 11, p. 6, 1998

[29] D. Ghiglia and L. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods", Journal of the Optical Society of America A, vol. 11, no. 1, p. 11, 1994

[30] M. Pritt, "Phase Unwrapping by Means of Multigrid Techniques for Interferometric SAR", IEEE Transactions on Geoscience and Remote Sensing, vol. 34, no. 3, p. 11, 1996

[31] D. Ghiglia and L. Romero, "Minimum L^p-norm two-dimensional phase unwrapping", Journal of the Optical Society of America A, vol. 13, no. 10, p. 15, 1996

[32] J. Wyant and K. Creath, "Basic Wavefront Aberration Theory for Optical Metrology", Applied Optics and Optical Engineering, vol. 11, p. 53, 1992

[33] V. Mahajan, "Optical imaging and aberrations: Part III: Wavefront Analysis", Bellingham, Washington USA: SPIE Press, 2013

[34] M. Born and E. Wolf, "Principles of Optics", 6th Edition, Oxford, UK: Pergamon Press, 1980

[35] J. Schwider, R. Burow, K. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel, "Digital wavefront measuring interferometry: some systematic error sources", Applied Optics, vol. 22, no. 21, p. 12, 1983

[36] J. Wyant and C. Ai, "Effect of Piezoelectric Transducer Nonlinearity on Phase Shift Interferometry", Applied Optics, vol. 26, no. 6, p. 5, 1987

[37] K. Creath, "Phase Measurement Interferometry: Beware These Errors", in Laser Interferometry IV: Computer-Aided Interferometry, San Diego, USA, 1992

[38] P. Carré, "Installation et utilisation du comparateur photoélectrique et interférentiel du Bureau International des Poids et Mesures", Bureau International des Poids et Mesures, Sèvres, 1966

[39] C. Brophy, "Effect of intensity error correlation on the computed phase of phase-shifting interferometry", Journal of the Optical Society of America A, vol. 7, no. 4, p. 5, 1990

[40] J. Prangsma, "Noise in detectors for spectroscopy", Ibsen Photonics, 2015

[41] C. Ai, "Phase measurement accuracy limitation in phase-shifting interferometry", University of Arizona, Tucson, USA, 1987

[42] Y. Chen and Q. Kemao, "Phase-shifting algorithms with known and unknown phase shifts: comparison and hybrid", Optics Express, vol. 30, no. 5, p. 28, 2022

[43] M. Takeda, T. Aoki, T. Sotomaru, T. Ozawa, T. Komiyama and Y. Miyamoto, "Two-Dimensional Phase unwrapping by Direct Elimination of Rotational Vector Fields for Phase Gradients Obtained by Heterodyne Techniques", Optical Review, vol. 5, no. 6, p. 6, 1998

[44] R. M. Goldstein, H. A. Zebker and C. L. Werner, "Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping", Radio Science, vol. 23, no. 4, pp. 713-720, 1988

[45] J. M. Huntley, R. Cusack, and H. T. Goldrein, "Improved noise-immune phase-unwrapping", Applied Optics, vol. 345, no. 5, pp. 781-789, 1995

[46] J. M. Huntley, "Noise-immune phase unwrapping algorithm", Applied Optics, vol. 28, no. 15, p. 3, 1989

[47] J. M. Huntley, J. R. Buckland, and S. R. Turner, "Unwrapping noisy phase maps by use of a minimum-cost-matching algorithm", Applied Optics, vol. 34, no. 23, p. 9, 1995

[48] P. Rastogi, E. Hack, "Phase Estimation in Optical Interferometry", CRC Press, 2015

[49] M. Pawlak, "Image Analysis by Moments: Reconstruction and Computational Aspects", Oficyna Wydawnicza Politechniki Wroclawskiej, Wroclaw 2006

[50] E. Hack, J. Burke, "Invited Review Article: Measurement uncertainty of linear phase-stepping algorithms", AIP Publishing, 2011

[51] R. Cordero et al., "Uncertainty analysis of temporal phase-stepping algorithms for interferometry", Optics Communications, vol. 275, no. 1, p.144-155, 2007

[52] Joint Committee for Guides in Metrology (JCGM/WG1), "Guide to the Expression of Uncertainty in Measurement", 2008

[53] P. H. Richter, "Estimating Errors in Least-Squares Fitting", The Telecommunications and Data Acquisition Progress Report, 1995

A Noise Quantities: Formulas

TABLE I. Influence quantities and references to the respective formula for quantitative evaluation. Explicit expressions for algorithms with $\sin(\beta_k + \delta_k) = 0 \ \forall k$ and for the N-bucket algorithms are included.

Influence quantity to be estimated[a]	General case	Phase angle average for cases with $\sin(\beta_k + \delta_k) = 0 \ \forall k$	N-bucket algorithms
Camera additive noise and quantization levels	Eq. (26)	Eq. (27) $$(u^2(\varphi, S_k))_\varphi = 2\frac{u^2(S)}{S_F^2} F$$ $$F = \frac{|c|^2}{|c C^+ c^+|^2}$$	Eq. (28) $$u_N^2(\varphi, S_k) = \frac{u^2(S)}{S_F^2} \times \frac{2}{N}$$ $$F = \frac{1}{N}$$
$u(S) = S_{\text{camnoise}}$ (GL)			
$u(S) = \frac{1}{\sqrt{12}} S_{\text{LSB}}$ (GL)			
Illumination instability $r(I)$ (%)	Eq. (31)	Eq. (32)	$$u_N^2(\varphi, \text{fluct}) = \frac{r^2(\iota)}{N}\left[\frac{2}{\nu^2} + \frac{1}{2} - \left(\frac{2}{\nu}\delta_{N3} + \frac{1}{2}\delta_{N4}\right)\cos(N\varphi)\right]$$
Camera gain fluctuations $r(\gamma)$ (%)		$(u^2(\varphi, \text{fluct}))_\varphi = r^2(\iota)\left(\frac{2}{\nu^2} + \frac{1}{2}\right) F$	
Phase step jitter $u(\beta)$ (rad)	Eq. (34)	Eq. (35) $(u^2(\varphi, \beta))_\varphi = u^2(\beta)\frac{3}{2} F$	$u_N^2(\varphi, \beta) = \frac{u^2(\beta)}{2N}(3 + \delta_{N4}\cos(N\varphi))$
Poisson count statistics: number of photons	Eq. (37)	Eq. (38)	$$u_N^2(\varphi, \text{Poisson}) = \frac{2}{S_F^2 N}(S_B - \delta_{N3}\frac{1}{3}S_\Gamma\cos(N\varphi))$$
S_B, S_Γ (cnts)		$(u^2(\varphi, \text{Poisson}))_\varphi = \frac{2}{S_B\nu^2} F$	$= \frac{2}{S_B N \nu^2}(1 - \delta_{N3}\frac{1}{3}\nu\cos(N\varphi))$
Linear phase step miscalibration $\beta(1+\varepsilon)$	Eq. (41)	Eq. (42)	$$u_N(\varphi, \varepsilon) = \varepsilon\frac{\pi}{N}\left((N-1) - \frac{\sin(2\varphi - \frac{2\pi}{N})}{\sin(\frac{2\pi}{N})}\right)$$
$\varepsilon[\%]$		$(u^2(\varphi, \varepsilon))_\varphi = \beta^2\varepsilon^2\frac{1}{4}N(N-1)(2N-1)F$	
General phase step miscalibration	Eq. (44)	Eq. (45)	$$u_N(\varphi, e_p) = \frac{e_p}{N}\sum_{k=0}^{N-1} k^p\left(1 - \cos\left(k\frac{4\pi}{N} + 2\varphi\right)\right)$$
$\varepsilon_k = k\varepsilon\beta + k^2 e_2 + \cdots + k^P e_P \ e_k$ (rad)		$(u^2(\varphi, e_p))_\varphi \le e_p^2 \frac{3}{2} \sum_{k=0}^{N-1} k^{2p} F$	
Detector nonlinearity $\tilde{S}_k = S_k(1 + r_2 S_k + \cdots + r_J S_k^{J-1})$	Eq. (48)		lowest contributing order $N - 1$ $$u_N^2(\varphi, r_{N-1}) = (r_{N-1}S_B^{N-2})^2\left(\frac{1}{2}\right)^{2N-4}\sin^2(N\varphi)$$
r_n (GL^{1-n})			
Higher harmonics $\frac{S_k}{S_\Gamma}$ (%)	Eq. (52)	Eq. (53) $(u^2(\varphi, S_q))_\varphi \le \frac{S_q^2}{S_\Gamma^2} N F$	$u_N^2(\varphi, S_{N-1}) = \frac{S_{N-1}^2}{S_\Gamma^2}\sin^2(N\varphi)$
Background irradiance drift δ_{drift} (%)	Eq. (56)	Eq. (57) $(u^2(\varphi, \delta_{\text{drift}}))_\varphi \le \frac{\delta_{\text{drift}}^2}{\nu^2}\frac{1}{3}N(N-1)(2N-1)F$	$u_N^2(\varphi, \delta_{\text{drift}}) = \frac{\delta_{\text{drift}}^2}{\nu^2}\left\{\frac{\cos^2(\varphi - \frac{\pi}{N})}{\sin^2(\frac{\pi}{N})}\right\}$
Source irradiance drift $\delta_{\text{illumdrift}}$ (%)	Eq. (60)	Eq. (61) $(u^2(\varphi, \delta_{\text{illum.drift}})) \le \delta_{\text{illum.drift}}^2\left(\frac{1}{\nu^2} + \frac{1}{4}\right)\frac{1}{3}N(N-1)(2N-1)F$	$u_N^2(\varphi, \delta_{\text{illum.drift}})$ $= \delta_{\text{illum.drift}}^2\left\{\frac{1}{\nu}\frac{\cos(\varphi - \frac{\pi}{N})}{\sin(\frac{\pi}{N})} + \frac{\cos(2\varphi - \frac{2\pi}{N})}{2\sin(\frac{\pi}{N})}\right\}^2$

[a]Units are (GL) for Gray levels, (%) for dimensionless ratios, and (cnts) for electron counts.

Figure 82: Table containing the necessary expressions for the analytical calculation of the uncertainties originating from individual error sources. [50]

93

B Edge Detection

Visibility γ, also called "Data Modulation" [3], is a measure of the quality of the datapoints of the phase map. It is generally defined as the ratio between I'' and I', or as an analytical expression in terms of the recorded interferograms, where 1 corresponds to the highest visibility/data quality and 0 to the lowest one. The wrapped phase map calculated from interferograms is often "contaminated". That is, datapoints outside of the pupil may have values different than 0. This may be due to stray light or reflections, for example. However, the quality of the data in these locations is usually much lower compared to the quality over the pupil, hence, the calculated visibility can be used to filter them out. This can be seen in Fig.(83-86).
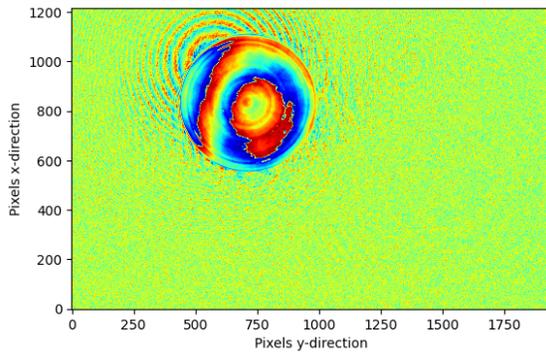

Figure 83: Contaminated phase map as calculated directly from the Schwieder-Hariharan PSA.
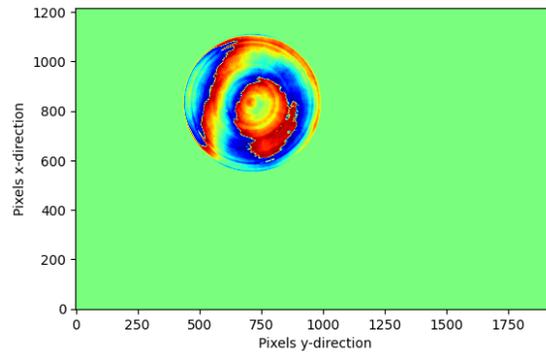

Figure 84: Decontaminated phase map using the visibility γ and >0.1 as the threshold condition.
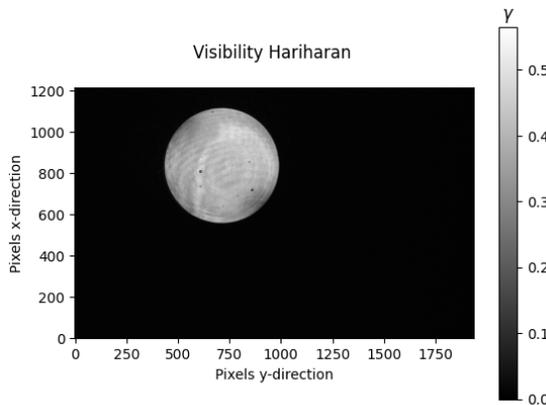

Figure 85: Calculated visibility from the analytical expression for Schwieder-Hariharan PSA [3].
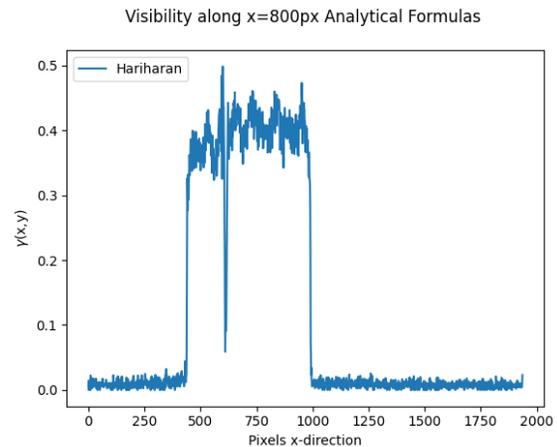

Figure 86: Cross-section of the calculated visibility. The sharp drop in γ at about x=600px corresponds to a dust particle.

After successful filtering, a simple path-following algorithm is used to find the location and di-

ameter of the pupil and crop the original phase map accordingly. This can be seen in Fig.(87) In this report we have only considered circular pupils. In reality the pupil may not necessarily be exactly circular and may have a more complex shape.
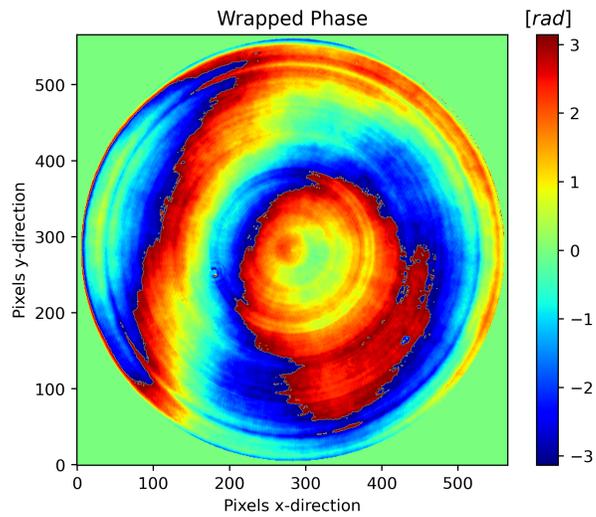

Figure 87: Example of using edge detection to crop the wrapped phase map. The phase map is identical to the one shown in Fig.(84).