Analysis of wavefront sensing techniques for extended scene imaging

Justin M. Knight
James C. Wyant College of Optical Sciences, University of Arizona
Michael Hart
James C. Wyant College of Optical Sciences, University of Arizona
HartSCI LLC

ABSTRACT

A design goal for Earth observing satellites is to launch a low-weight, large aperture primary mirror which provides high-resolution images of extended scenes of interest. In remote sensing, this translates to achieving a small ground sampling distance over a possibly large bandwidth on the imaging detector in the presence of noise sources as well as error inducing environmental factors such as repeated exposure to the sun in low-earth orbit, causing thermal fluctuations which distort the primary mirror shape and misalign optical elements present in the satellite – ultimately manifesting as wavefront aberrations across the telescope pupil. A practiced method of accomplishing this goal is to equip such a satellite with an adaptive optics (AO) system to provide a means of measuring and compensating for wavefront errors through the combination of a wavefront sensor and a corrector element such as a fast-steering or deformable mirror. A standard technique employed for wavefront sensing is to reconstruct the wavefront across the telescope pupil from wavefront slope estimates calculated using Shack-Hartmann wavefront sensor (SHWFS) sub-aperture information; in extended scenes, this is usually performed via subimage correlations with reference images. While popular, Shack-Hartmann wavefront sensors are subject to performance limitations, many of which are addressed by developing correlation algorithms to overcome practical issues such as thresholding useless information across subapertures, for example, from cloudy scenes, poor wavefront estimates from both low signal-to-noise ratio across subapertures and using subimage boundary information, and so on. While other extended scene wavefront sensing techniques such as broadband phase diversity and plenoptic sensors have been studied with respect to adaptive correction in remote sensing, even being compared to the SHWFS directly, we seek to compare a suite of wavefront sensing techniques with respect to their performance in the limit of error and noise sources – focusing on the latter in this study. To this end, we consider wavefront sensing techniques previously developed for use in traditional nighttime astronomy, solar astronomy, or optical metrology such as the aforementioned SHWFS for extended scenes, hereafter referred to as the Shack-Hartmann correlation tracker, and optical differentiation wavefront sensor. Common error sources under our consideration include the previously detailed thermal stresses in orbit which cause telescope pupil aberrations, as well as other various environmental exposures which result in a degradation of specular efficiency from the surfaces of optical elements, producing stray light, and detector pixel loss from radiation damage. Meanwhile noise sources include various forms of detector noise, of which the most fundamental to light-matter interactions is photon noise; as such, it is our primary focus. We present a photon-noise sensitivity parameter in the presence of a telescope pupil phase aberration for various wavefront sensor techniques; such a parameter demonstrates how well the Shack-Hartmann correlation tracker performs against less well-characterized wavefront sensing techniques in the realm of remote sensing under ideal imaging conditions. This allows us to consider the viability of wavefront sensing techniques for an AO system based on physics-imposed limitations before extending our analysis to include practical ones.

1. INTRODUCTION

Ideally, Earth observation satellites performing remote sensing are launched into space with lightweight, large aperture primary mirrors which are impervious to changes in their environment that are responsible for wavefront aberrations; such telescope systems are, in theory, capable of achieving sub-meter ground sampling distances over a large bandwidth on their noiseless detectors and in doing so, produce high-resolution image quality of extended scenes of interest. For this work, we assume these ideals can be approached by equipping the satellite with an optical compensatory...
mechanism, here referred to as an adaptive optics (AO) system, with the express purpose of sensing and compensating for wavefront aberrations in the presence of noise brought about by the various realities of imaging from space. We assume the wavefront sensing is performed using a wavefront sensor (WFS) such as a Shack-Hartmann WFS, and the subsequent compensation of the estimated wavefront distortion via an optical element such as a deformable mirror. To begin, we review the physical mechanisms of inducing wavefront aberration from space, followed by a description of noise sources which serve as the fundamental physical limitations to compensating for wavefront errors in a remote sensing satellite image.

1.1 Wavefront error sources

The comprehensive review in [4] on the precise origin of thermal and mechanical issues of deployable space optics provides a backdrop for the similar issues which can plague more conventional remote sensing satellites; here we outline the wavefront error sources which bear repeating for our intended analysis, as they are usually considered the most damaging to image quality. Namely, thermal and ground-testing issues:

- Lightweight materials used to cast large primary mirrors deform under several mechanisms, including the difference between ground-testing with the effect of gravity and orbit in space [7, 12, 24], as well as non-uniform thermal contraction [2] from, for example, exposure to sunlight during parts of their orbit such as satellites in LEO. Both deformation mechanisms manifest as some combination of, typically, low-order aberrations such as tip, tilt, defocus, spherical aberration, coma, etc.

- Environmental effects such as thermal stress or spacecraft vibrations impacting the telescope system, causing a cascade of optical element misalignments with respect to their ground-tested configuration [19, 24]. The standard for characterizing optical misalignments is using Zernike polynomials or the like to describe the aberrated wavefront.

There are other sources of error which should be taken into consideration, as they ultimately degrade image quality as well. These include atmospheric effects like scintillation [24] and stray light and radiation damage on detector pixel arrays. No matter the case, each of these error sources is categorized as practical, meaning these are tied to the actual operation of the satellite in space and as such, can be addressed in some way by design choices and real-time operation of an AO system.

1.2 Noise sources

Noise sources are present in any real imaging system. To include the most pertinent noise sources for remote sensing, we follow the model paradigm introduced by the the Image Based Sensor Model (IBSM) [13] as well as noise sources included in the work of estimating sparse aperture telescope segment phases and wavefront information via broadband phase diversity [5]. Moreover the work of identifying the limits of adaptive optics [9] for high-contrast imaging, which is ultimately determined as the ability to accurately extract wavefront information for compensation, provides a groundwork for which noise sources to more carefully consider when selecting a wavefront sensing technique.

The following noise sources to consider are:

- photon noise from the scene, which may be both from a target as well as its background
- detector noise sources, including dark current, readout noise, quantization noise from an analog-to-digital converter
- possibly other electronics noise sources, such as from the signal chain [5].

Other possible sources of error worth considering are the effects of aliasing of high spatial frequency content at the wavefront sensor, chromaticity effects in the measurement of a wavefront, and non-common path errors between sensing and imaging wavelengths.
1.3 Focus of this work
We adopt a scenario in which the performance of an AO system equipped satellite is in a photon-noise limited regime. This means there are wavefront errors of the form described in Section 1.1 which dominate the image quality of such a system, but the noise in the wavefront sensor detector plane is characterized as photon, or shot noise. This is the prevailing physical limitation in optics, as it is an inescapable result owed to the quantum nature of light. Thus, for any wavefront sensing techniques used in extended scene imaging, which we now describe, the information recorded of the wavefront is with respect to a signal in the presence of photon noise.

2. WAVEFRONT SENSING TECHNIQUES FOR EXTENDED SCENE IMAGING

From publicly available literature, the SHWFS performing correlation measurements between subapertures is the workhorse technique for estimating wavefront information in extended scene imaging; this technique is referred to as the Shack-Hartmann correlation tracker [6]. The SH correlation tracker is popular for a combination of reasons, principal among them being the ease of extension from extracting wavefront information from point sources to the analogous procedure with respect to extended scenes, as well as it being a well established and characterized method of wavefront estimation in fields such as astronomy. A recent publication [24] details the classes of correlation algorithms developed for over more than two decades. An entire account of the technical development of these wavefront estimation algorithms is unnecessary presently; however, we list all such algorithms to note the level of research into one particular wavefront sensing technique:

- spatial domain correlation [15, 16]
- spatial-frequency domain correlation [15, 22, 23]
- iterative cross-correlation [27–30]
- phase correlation [31]
- maximum likelihood estimation [8].

The algorithms developed and analyzed in [24] can serve as a baseline for Shack-Hartmann wavefront sensing performance, with the most well-known algorithm studied being spatial frequency domain correlation [22]. A more apt comparison to state-of-the-art algorithms which deal with more practical issues comes from reproducing the algorithm specifically developed in [24]. In any case, investigating the impact of noise on these algorithms is the main consideration, as it ultimately limits the performance of the SH correlation tracker.

2.1 Alternative wavefront sensing techniques
Previously studied wavefront sensing techniques in the context of remote sensing are considered for their performance in estimating wavefront information. These include wavefront sensing strategies which have been directly compared in either simulation or experiment, to the SHWFS/SH correlation tracker. In addition, techniques utilized in solar and astronomical adaptive optics, and optical metrology are considered. The entire list is as follows:

- the plenoptic wavefront sensor [11]
- phase diversity, i.e. focal plane wavefront sensing (FPWFS) [1, 5, 14, 17, 18, 20]
- the optical differentiation wavefront sensor (ODWFS), referred to as the pyramid wavefront sensor (PyWFS) for extended scenes [25, 26]
- the curvature wavefront sensor (CWFS) [9, 10]
- the differential optical transfer function (dOTF) [3].

This list is not necessarily exhaustive, so other wavefront sensing techniques may be considered as well on the basis of their merit for estimating wavefront information in an extended scene, or ability to do so in some sort of procedure where a light beacon is made available to image, for example. A thorough review of each of these wavefront sensing techniques is beyond the scope of this work, as it is only necessary after their sensitivity to photon noise has been qualitatively assessed.
3. UNDERSTANDING PHOTON NOISE SENSITIVITY: THE $\beta_p$ PARAMETER

Reference [9] develops a sensitivity parameter with respect to the photon-limited performance regime of wavefront sensing techniques. The goal of this analysis is to quantify the degree to which wavefront sensing accuracy, and therefore the quality of information available in the image plane, is dependent on spatial frequency information transmitted by the telescope and its subsequent optical subsystems. In particular, the interest for the field of high-contrast imaging in astronomy is imaging very faint astronomical objects extremely close to orders of magnitude brighter ones by using a combination of extreme adaptive optics correction and devices called coronagraphs which operate on the principle of canceling or modulating the diffraction patterns from on-axis star light; assuming exquisite wavefront correction, this produces the necessary conditions for direct imaging of said faint objects of interest, which are typically exoplanets. To facilitate the understanding of the effects of the physics of the scene being imaged, the analysis is conducted with respect to a sinusoidal wavefront representation. A derivation of the parameter for phase-only aberrations, known as $\beta_p$, can be found in Appendix A of [9]. Here we offer a similar derivation with an emphasis on understanding why this is an important parameter for precisely capturing wavefront sensing performance under ideal imaging conditions.

3.1 Probabilistic description of a residual sinusoidal phase wavefront error

The wavefront error from a sinusoidal phase distortion can be represented in a two-dimensional coordinate system by:

$$x_0 = \frac{2\pi h}{\lambda} \cos \theta$$

$$y_0 = \frac{2\pi h}{\lambda} \sin \theta$$

where $\theta$ is the spatial frequency under consideration.

A wavefront sensor produces a set of $N$ measurements $I = I_0, I_1, \ldots, I_{N-1}$, which is often information from imaging point sources such as centroids calculated from a lenslet array of a SHWFS, but any method of encoding phase information in a set of intensity measurements is sufficient to proceed with the following analysis. In any case, this measurement set is used to compute $(x, y)$, the measured values of $(x_0, y_0)$. Because of photon noise, $x \neq x_0$ and $y \neq y_0$.

After correction by the adaptive optics system, usually from a deformable mirror, the residual sine wave phase aberration at the spatial frequency $\theta$ has an amplitude $A_{\text{res}}$ (in radians) of:

$$A_{\text{res}} = \sqrt{(x-x_0)^2 + (y-y_0)^2}$$

This residual phase error may be computed by calculating the expectation value of $A_{\text{res}}^2$ given a probability density function $P(x, y)$:

$$E(A_{\text{res}}^2) = \Sigma = \sqrt{\int \int_{x,y} A_{\text{res}}^2 P(x,y) dx dy}$$

$P(x, y)$ is a function of the set of wavefront sensor measurements $I$. These measurements are considered independent from one another, as in the case of a Shack-Hartmann or pyramid WFS. Then the probability density function can be written as a product of probability density functions for each wavefront sensor measurement:

$$P(x, y) = \prod_{k=0}^{N-1} P_k(x, y)$$

Furthermore, each measurement $I_k$ is assumed to follow a normal distribution; this is a common approach when the underlying distribution is unknown, but many measurements can be made. Finally, in a closed-loop system where
residual errors ideally only become smaller with each loop iteration, each \( I_k \) is assumed to be well approximated by a first order Taylor series expansion in \( x \) and \( y \):

\[
I_k(x, y) \approx I_k(x_0, y_0) + \frac{\partial I_k}{\partial x} X + \frac{\partial I_k}{\partial y} Y
\]

where the partial derivatives are evaluated at \( x = x_0 \) and \( y = y_0 \), and \( X = x - x_0, Y = y - y_0 \). Then for a particular WFS measurement \( I_k \), the probability density function \( P_k(x, y) \) can be written as:

\[
P_k(x, y) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left[-\frac{(I_k(x, y) - \mu_k)^2}{2\sigma_k^2}\right]
\]

where \( \mu_k \) and \( \sigma_k^2 \) are the mean and the variance respectively of the \( I_k \) measurement. Applying the first order Taylor series expansion to \( P_k(x, y) \) yields:

\[
P_k(x, y) \approx \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left[-\frac{(I_k(x_0, y_0) + \frac{\partial I_k}{\partial x} X + \frac{\partial I_k}{\partial y} Y - \mu_k)^2}{2\sigma_k^2}\right]
\]

\( I_k(x_0, y_0) \) is the intensity measurement for the exact wavefront error, while \( \mu_k \) is the expected value of \( I_k \). In closed-loop, \( I_k(x_0, y_0) - \mu_k \) should be closed to zero, as the average measurement should be close to the actual measurement for a well-corrected system, and is therefore negligible.

Then

\[
P_k(x, y) \approx \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left[-\frac{\left(\frac{\partial I_k}{\partial x} X + \frac{\partial I_k}{\partial y} Y\right)^2}{2\sigma_k^2}\right]
\]

### 3.2 Manipulating the probability density function into a bivariate form

We note that in the original derivation the following relations are defined,

\[
k_x = \frac{\frac{\partial I_k}{\partial x}}{\sqrt{\left(\frac{\partial I_k}{\partial x}\right)^2 + \left(\frac{\partial I_k}{\partial y}\right)^2}}
\]

\[
k_y = \frac{\frac{\partial I_k}{\partial y}}{\sqrt{\left(\frac{\partial I_k}{\partial x}\right)^2 + \left(\frac{\partial I_k}{\partial y}\right)^2}}
\]

\[
\sigma_k = \frac{\sigma_k}{\sqrt{\left(\frac{\partial I_k}{\partial x}\right)^2 + \left(\frac{\partial I_k}{\partial y}\right)^2}}
\]

so that the probability density function is expressed by:

\[
P_k(x, y) \approx \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left[-\frac{(k_x X + k_y Y)^2}{2\sigma_k^2}\right]
\]

At this point, the normalization coefficient for each \( P_k(x, y) \) is suppressed, so the entire probability density function \( P(x, y) \) is
\[ P(x, y) = \prod_{k=0}^{N-1} P_k(x, y) \]
\[ \approx \prod_{k=0}^{N-1} \exp \left[ - \frac{(k_x X + k_y Y)^2}{2\sigma_k^2} \right] \]
\[ = \exp \left\{ - \frac{1}{2} \left[ \frac{(\partial h_0/\partial x + \partial h_{N-1}/\partial y)^2}{\sigma_{I_0}^2} + \cdots + \frac{(\partial h_{N-1}/\partial x + \partial h_{N-1}/\partial y)^2}{\sigma_{I_{N-1}}^2} \right] \right\} \]

Expanding the squared terms and collecting like terms results in the exponent argument with three distinct quadratic variables produces three distinct coefficient sums,

\[
\left( \frac{(\partial h_0/\partial x)}{\sigma_{I_0}^2} \right)^2 X^2 + \cdots + \left( \frac{(\partial h_{N-1}/\partial x)}{\sigma_{I_{N-1}}^2} \right)^2 X^2
\]
\[ + \left( \frac{(\partial h_0/\partial y)}{\sigma_{I_0}^2} \right)^2 Y^2 + \cdots + \left( \frac{(\partial h_{N-1}/\partial y)}{\sigma_{I_{N-1}}^2} \right)^2 Y^2
\]
\[ = \left( \frac{X^2}{\alpha_1} + \frac{XY}{\alpha_2} + \frac{Y^2}{\alpha_3} \right)
\]

Or written more compactly,

\[
\frac{1}{\alpha_1} = \sum_{k=0}^{N-1} \frac{k_x^2}{2\sigma_k^2}
\]
\[
\frac{1}{\alpha_2} = \sum_{k=0}^{N-1} \frac{k_x k_y}{\sigma_k^2}
\]
\[
\frac{1}{\alpha_3} = \sum_{k=0}^{N-1} \frac{k_y^2}{2\sigma_k^2}
\]

so that the entire probability density function is expressed as

\[ P(x, y) \approx \exp \left[ - \left( \frac{X^2}{\alpha_1} + \frac{XY}{\alpha_2} + \frac{Y^2}{\alpha_3} \right) \right] \] (8)

Briefly, a bivariate normal distribution is a normal distribution of two independent random variables. For zero-mean random variables \((x_1, x_2)\), this is formally written as

\[
f(x_1, x_2) = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}} \exp \left[ - \frac{1}{2(1-\rho^2)} \left( \frac{x_1^2}{\sigma_{x_1}^2} - \frac{2\rho x_1 x_2}{\sigma_{x_1} \sigma_{x_2}} + \frac{x_2^2}{\sigma_{x_2}^2} \right) \right]
\] (9)
where $\rho$ is the correlation coefficient, and $\sigma_{x_1}^2$, $\sigma_{x_2}^2$ are the variances of the variables $(x_1, x_2)$ respectively.

Comparing the arguments of equations (8) and (9), we see that $P(x, y)$ is a bivariate normal distribution with the following relationships between coefficients:

$$
\sigma_{x_1} = \sqrt{\frac{\alpha_1}{2(1 - \rho^2)}} \\
\sigma_{x_2} = \sqrt{\frac{\alpha_3}{2(1 - \rho^2)}} \\
\rho = \sqrt{\frac{\alpha_1 \alpha_3}{2 \alpha_2}}
$$

For consistency, let $x_1 = X$ and $x_2 = Y$ so that $\sigma_{x_1} = \sigma_x$ and $\sigma_{x_2} = \sigma_y$. Further suppressing the normalization coefficients yields the current form of the probability density function:

$$
P(x, y) \approx \exp \left[ -\frac{1}{2(1 - \rho^2)} \left( \frac{X^2}{\sigma_x^2} - \frac{2 \rho XY}{\sigma_x \sigma_y} + \frac{Y^2}{\sigma_y^2} \right) \right] \quad (10)
$$

The goal is to integrate this expression for $P(x, y)$ in order to evaluate $\Sigma^2$ via equation (2). A linear transformation of variables produces a more familiar integrand with which to work. Define

$$
X = \sigma_x X' \\
Y = \sigma_y \left( \rho X' + \sqrt{(1 - \rho^2)} Y' \right)
$$

Substituting these relations into equation (10), we find the polynomial terms of the exponential reduce by expanding and combining like terms:

$$
\frac{(\sigma_x X')^2}{\sigma_x^2} - \frac{2 \rho \sigma_x X'}{\sigma_x \sigma_y} \left( \rho X' + \sqrt{(1 - \rho^2)} Y' \right) + \frac{\sigma_y^2 \left( \rho X' + \sqrt{(1 - \rho^2)} Y' \right)^2}{\sigma_y^2} = (1 - \rho^2) X'^2 + (1 - \rho^2) Y'^2,
$$

so that the entire expression becomes

$$
P(x, y) \approx \exp \left[ -\frac{1}{2(1 - \rho^2)} \left( (1 - \rho^2) X'^2 + (1 - \rho^2) Y'^2 \right) \right] \quad (11)
$$

$$
= \exp \left[ -\frac{X'^2 + Y'^2}{2} \right] \quad (12)
$$

### 3.3 Evaluating the integral expression

With this change of variables of the probability density function to $P(x', y')$, the final substitutions are to the rest of the $(x, y)$ dependencies in equation (2), namely $A_{res}$ and the differential variables $(dx, dy)$. 

---

Copyright © 2021 Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS) – www.amostech.com
It follows that the differential values \((dx, dy) = (\sigma_x dx', \sigma_y \sqrt{1 - \rho^2} dy')\), which upon substitution cancel the normalization factor (formally, the Jacobian) preceding the exponential function in equation (8), so that upon the final substitution for \(A_{\text{res}}^2\), the square of the integral takes the form of:

\[
\Sigma^2 = \frac{1}{2\pi} \iint_{X', Y'} \left[ (\sigma_x X')^2 + \left( \sigma_y \left[ \rho X' + \sqrt{1 - \rho^2} Y' \right] \right)^2 \right] P(x', y') dx' dy' \\
= \frac{1}{2\pi} \iint_{X', Y'} \sigma_x^2 X'^2 P(x', y') dx' dy' + \frac{1}{2\pi} \iint_{X', Y'} \sigma_y^2 Y'^2 P(x', y') dx' dy' \\
+ \frac{1}{2\pi} \iint_{X', Y'} \rho^2 \sigma_y^2 (X'^2 - Y'^2) P(x', y') dx' dy' \\
+ \frac{1}{2\pi} \iint_{X', Y'} 2\sigma_y^2 \rho \sqrt{1 - \rho^2} X' Y' P(x', y') dx' dy' \\
= \Sigma_0 + \Sigma_1 + \Sigma_2 + \Sigma_3
\]

Evaluating these integrals requires limits of integration; determining the limits is a matter of understanding which values are admitted for \((X, Y) = (x - x_0, y - y_0)\), and thereby \((X', Y')\). Recalling that \((x, y)\) and \((x_0, y_0)\) are mismatched sinusoidal phase aberrations, it is clear that any range of value is possible for the resulting difference between the two sets of aberrations, as both the height and spatial frequency are subject to inequality from photon noise. Then it is true that we can integrate over the entire range of possible values \((X, Y)\), i.e. \((-\infty, \infty)\). Furthermore since the change of variables is a linear transformation, we integrate over the same range with respect to \((X', Y')\).

Let us write out and evaluate each separate integral \(\Sigma_i, i = 0, \ldots, 3\) with the expression for \(P(x', y')\) from equation (12):

\[
\Sigma_0 \approx \frac{\sigma_x^2}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X'^2 \exp \left[ -\frac{X'^2 + Y'^2}{2} \right] dx' dy' = \sigma_x^2 \\
\Sigma_1 \approx \frac{\sigma_y^2}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Y'^2 \exp \left[ -\frac{X'^2 + Y'^2}{2} \right] dx' dy' = \sigma_y^2 \\
\Sigma_2 \approx \rho^2 \sigma_y^2 \int_{-\infty}^{\infty} (X'^2 - Y'^2) \exp \left[ -\frac{X'^2 + Y'^2}{2} \right] dx' dy' = 0 \\
\Sigma_3 \approx \frac{\sigma_y^2 \rho \sqrt{1 - \rho^2}}{\pi} \int_{-\infty}^{\infty} (X' Y') \exp \left[ -\frac{X'^2 + Y'^2}{2} \right] dx' dy' = 0
\]

where we have evaluated each integral using integration by parts, symmetry about the limits of integration, and the result of evaluating a Gaussian integrand over the range of \((-\infty, \infty)\), namely

\[
\int_{-\infty}^{\infty} e^{-(x^2 + y^2)} dx dy = \pi
\]

In summary, the residual error from a sinusoidal wavefront aberration as estimated by a set of wavefront sensor measurements is

\[
\Sigma \approx \sqrt{\sigma_x^2 + \sigma_y^2}
\]

For most wavefront sensors, \(\Sigma\) is a function of spatial frequency \(\theta\), in which case the maximum residual error over all spatial frequencies is considered to represent the quantitative measurement of sensitivity. Since the measurement errors from wavefront sensors are assumed to be produced by photon noise, \(\Sigma\) is inversely proportional to the square root of the number of photons \(N_{ph}\) incident on the WFS detector, with equality provided by a constant of proportionality, \(\beta_p\):
The parameter $\beta_p$ represents the sensitivity of the wavefront sensor to photon noise as a function of spatial frequency.

4. AN EXAMPLE ANALYTICAL DERIVATION OF $\beta_p$

As is customary in published research, details of derivations are omitted in favor of brevity and to focus on using results to arrive at further reaching conclusions of analysis. However, the prospect of building on previous work is more elusive in such a paradigm. Therefore, being able to independently derive previously published results is an important stepping stone for moving research forward in new directions of interest; specifically, expanding $\beta_p$ to include previously unconsidered wavefront sensors, and with respect to entirely different wavefront representations more descriptive of an imaging system and its operating environment of interest. Thus, it is in our interest to determine how we may derive $\beta_p$ for the WFSs considered in [9].

We begin with the expression for the residual sinusoidal phase error, although in practice any representation of a wavefront phase error is admissible since the probability density function used to derive $\Sigma$ is constructed with respect to a non-zero difference in value due to photon noise, and is not dependent on the form of the wavefront information directly. In other words, writing wavefront information in the form of sinusoids, i.e. Fourier modes, or some other polynomial expansion such as Zernike modes, is not a necessary condition for the derivation of $\beta_p$.

First, we rewrite the expression for $\Sigma^2$ in terms of the $\alpha_i$ and $\rho$ parameters:

$$
\Sigma^2 \approx \sigma_x^2 + \sigma_y^2 = \frac{\alpha_1 + \alpha_3}{2(1 - \rho^2)}
$$

Furthermore, write $\rho$ in terms of the $\alpha_i$ parameters:

$$
1 - \rho^2 = 1 - \frac{\alpha_1 \alpha_3}{4 \alpha_2^2} = \frac{4 \alpha_2^2 - \alpha_1 \alpha_3}{4 \alpha_2^2}
$$

so that the expression for $\Sigma^2$ becomes

$$
\Sigma^2 \approx \frac{2 \alpha_2^2 (\alpha_1 + \alpha_3)}{4 \alpha_2^2 - \alpha_1 \alpha_3} = \frac{2 \alpha_1 \alpha_2 \alpha_3 \left( \frac{1}{\alpha_1} + \frac{1}{\alpha_3} \right)}{\alpha_1 \alpha_2^2 \alpha_3 \left( \frac{4}{\alpha_1 \alpha_3} - \frac{1}{\alpha_2^2} \right)} = \frac{2 \left( \frac{1}{\alpha_1} + \frac{1}{\alpha_3} \right)}{\left( \frac{4}{\alpha_1 \alpha_3} - \frac{1}{\alpha_2^2} \right)}
$$

(20)

4.1 The form of intensity measurement sensitivities

Often times in [9] the intensity expressions are formed in terms explicitly from the variables of $(x_0, y_0) = \left( \frac{2\pi h}{\lambda} \cos \theta, \frac{2\pi h}{\lambda} \sin \theta \right)$. Therefore, it is useful to express the intensity measurement sensitivities with respect to these variables as well.

Copyright © 2021 Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS) – www.amostech.com
This is accomplished by writing an intensity measurement $I(x, y)$ as $I(x(h, \theta), y(h, \theta))$ and using the chain rule for computing a partial derivative with respect to each independent variable:

$$\begin{bmatrix}
\frac{\partial I}{\partial h} \\
\frac{\partial I}{\partial \theta}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial x}{\partial h} & \frac{\partial y}{\partial h} \\
\frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta}
\end{bmatrix} \begin{bmatrix}
\frac{\partial I}{\partial x} \\
\frac{\partial I}{\partial y}
\end{bmatrix}$$

Or, in short form

$$I_{h\theta} = J_I I_{xy}$$

This $2 \times 2$ Jacobian matrix $J_I$ can be inverted to find $I_{h\theta}$ in terms of $I_{xy}$:

$$J_I^{-1} = \frac{1}{\det J_I} \begin{bmatrix}
\frac{\partial y}{\partial \theta} & -\frac{\partial y}{\partial h} \\
-\frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial h}
\end{bmatrix}$$

where $\det J_I$ is

$$\det J_I = \frac{\partial x}{\partial h} \frac{\partial y}{\partial \theta} - \frac{\partial y}{\partial h} \frac{\partial x}{\partial \theta} = \frac{4\pi^2 h}{\lambda^2}$$

Using this result the expression for $I_{xy}$ is

$$I_{xy} = J_I^{-1} I_{h\theta}$$

$$= \frac{\lambda^2}{4\pi^2 h} \begin{bmatrix}
\frac{\partial y}{\partial \theta} & -\frac{\partial y}{\partial h} \\
-\frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial h}
\end{bmatrix} \begin{bmatrix}
\frac{\partial I}{\partial x} \\
\frac{\partial I}{\partial y}
\end{bmatrix}$$

$$\therefore \begin{bmatrix}
\frac{\partial I}{\partial x} \\
\frac{\partial I}{\partial y}
\end{bmatrix} = \frac{\lambda}{2\pi} \begin{bmatrix}
\cos \theta & \frac{-\sin \theta}{h} \\
\sin \theta & \frac{\cos \theta}{h}
\end{bmatrix} \begin{bmatrix}
\frac{\partial I}{\partial h} \\
\frac{\partial I}{\partial \theta}
\end{bmatrix}$$

4.2 $\beta_p$ for the curvature wavefront sensor

According to [9], a curvature wavefront sensor works by converting phase changes into intensity variations via a Fresnel propagation. In this work it is written as the propagation of the complex amplitude resulting from a pupil plane phase aberration,

$$W(u, 0) = 1 + \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta)$$

to some distance $z$: 
\[ W(u, z) = 1 + \sin(d\phi) \left( \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta) + i \cos(d\phi) \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta) \right) \]
\[ = 1 + ie^{-id\phi} \left( \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta) \right) \]

The corresponding intensity is

\[ I_W(u, z) = W(u, z) W^*(u, z) \]
\[ = \left[ 1 + ie^{-id\phi} \left( \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta) \right) \right] \left[ 1 - ie^{id\phi} \left( \frac{2\pi h}{\lambda} \sin(2\pi fu + \theta) \right) \right] \]
\[ = 1 + \frac{4\pi h}{\lambda} \sin(d\phi) \sin(2\pi fu + \theta) + \frac{4\pi^2 h^2}{\lambda^2} \sin^2(2\pi fu + \theta) \]
\[ \approx 1 + \frac{4\pi h}{\lambda} \sin(d\phi) \sin(2\pi fu + \theta) \]

The last line assumes that since \( h \ll \lambda \), the squared term in the expression is close to zero and is safely ignored to define a linear relationship of intensity to phase.

From here it is assumed that the \( N \) intensity measurements are taken per spatial period in the pupil plane, with the \( k \)-th measurement expressed in analogous sine phase aberration arguments \( (A, \phi) \):

\[ I_k = \frac{N_{ph}}{N} \left[ 1 + \frac{4\pi A \sin(d\phi)}{\lambda} \sin \left( \frac{2\pi k}{N} + \phi \right) \right] \]

and \( \sigma_k = \left( \frac{N_{ph}}{N} \right)^{1/2} \) representing the uncertainty in the \( k \)-th measurement from photon noise. Taking a look at the sine term corresponding to the \( fu \) argument, we see that choosing such a sampling scheme permits us to look at linear combinations of \( (x_0, y_0) = \left( \frac{2\pi A \cos \phi}{\lambda}, \frac{2\pi A \sin \phi}{\lambda} \right) \), thus allowing us to compute \( \beta_p \) quite readily. We begin by computing the partial derivatives of the \( k \)-th intensity measurement of \( N \) measurements total with respect to the variables \( (A, \phi) \).

\[ \frac{\partial I_k}{\partial A} = \frac{N_{ph}}{N} \left[ \frac{4\pi \sin(d\phi)}{\lambda} \sin \left( \frac{2\pi k}{N} + \phi \right) \right] \]
\[ \frac{\partial I_k}{\partial \phi} = \frac{N_{ph}}{N} \left[ \frac{4\pi A \sin(d\phi)}{\lambda} \cos \left( \frac{2\pi k}{N} + \phi \right) \right] \]

Therefore the partial derivatives with respect to \( (x_0, y_0) \) are

\[ \frac{\partial I_k}{\partial x} = \frac{2N_{ph}}{N} \sin(d\phi) \left( \cos \phi \sin \phi_k - \sin \phi \cos \phi_k \right) \]
\[ = -\frac{2N_{ph}}{N} \sin(d\phi) \sin \left( \frac{2\pi k}{N} \right) \]

\[ \left( \frac{\partial I_k}{\partial x} \right)^2 = \frac{4N_{ph}^2}{N^2} \sin^2(d\phi) \sin^2 \left( \frac{2\pi k}{N} \right) \]

and
\[
\frac{\partial I_k}{\partial y} = \frac{2N_{ph}}{N} \sin(d\phi) (\sin \phi \sin \phi_k + \cos \phi \cos \phi_k)
\]

\[
= \frac{2N_{ph}}{N} \sin(d\phi) \cos \left(\frac{2\pi k}{N}\right)
\]

\[
\therefore \quad \left(\frac{\partial I_k}{\partial y}\right)^2 = \frac{4N_{ph}^2}{N^2} \sin^2(d\phi) \cos^2\left(\frac{2\pi k}{N}\right),
\]

where \(\phi_k = \left(\frac{2\pi k}{N} + \phi\right)\).

To solve for \(\beta_p\), we start with the numerator expansion of equation (20):

\[
2 \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_3}\right) = \sum_{k=0}^{N-1} \frac{1}{\sigma_k^2} \left(\frac{\partial I_k}{\partial x}\right)^2 + \sum_{k=0}^{N-1} \frac{1}{\sigma_k^2} \left(\frac{\partial I_k}{\partial y}\right)^2
\]

\[
= \frac{4N_{ph} \sin^2(d\phi)}{N} \left[\sum_{k=0}^{N-1} \sin^2\left(\frac{2\pi k}{N}\right) + \sum_{k=0}^{N-1} \cos^2\left(\frac{2\pi k}{N}\right)\right]
\]

\[
= \frac{4N_{ph} \sin^2(d\phi)}{N} \left(\frac{N}{2} + \frac{N}{2}\right)
\]

\[
= 4N_{ph} \sin^2(d\phi)
\]

Computing the denominator proceeds as follows:

\[
\frac{4}{\alpha_1 \alpha_3 \alpha_2} - \frac{1}{\alpha_2^2} = \left[\sum_{k=0}^{N-1} \frac{1}{\sigma_k^2} \left(\frac{\partial I_k}{\partial x}\right)^2\right]^2 - \left[\sum_{k=0}^{N-1} \frac{1}{\sigma_k^2} \left(\frac{\partial I_k}{\partial y}\right)^2\right]^2
\]

\[
= \left(\frac{4N_{ph}^2 \sin^2(d\phi)}{N^2}\right)^2 \left(\frac{N}{N_{ph}}\right)^2 \left[\sum_{k=0}^{N-1} \sin^2\left(\frac{2\pi k}{N}\right) \sum_{k=0}^{N-1} \cos^2\left(\frac{2\pi k}{N}\right)\right]
\]

\[
- \left(\frac{4N_{ph}^2 \sin^2(d\phi)}{N^2}\right)^2 \left(\frac{N}{N_{ph}}\right)^2 \left[\sum_{k=0}^{N-1} \sin\left(\frac{2\pi k}{N}\right) \cos\left(\frac{2\pi k}{N}\right)\right]^2
\]

\[
= \left(\frac{4N_{ph} \sin^2(d\phi)}{N}\right)^2 \left[\frac{N^2}{4} - 0\right]
\]

\[
= 4N_{ph}^2 \sin^4(d\phi)
\]

Computing the ratio of the two quantities gives us \(\Sigma^2\):

\[
\Sigma^2 \approx \frac{4N_{ph} \sin^2(d\phi)}{4N_{ph}^2 \sin^4(d\phi)} = \frac{1}{N_{ph} \sin^2(d\phi)}
\]

so that \(\beta_p\) is computed as

\[
\beta_p = \Sigma \sqrt{N_{ph}} \approx \frac{1}{\sin(d\phi)} = \sec(d\phi)
\]
Finally, \(d\phi\) is defined as \(\pi f^2 z \lambda\), where \(f = \frac{2\pi}{\lambda}\) relates spatial frequency with an angular sky coordinate at the imaging wavelength \(\lambda\), and in the case of the curvature WFS, the distance \(z\) is the conjugate altitude \(\delta z\). Then \(\beta_p\) becomes

\[
\beta_p \approx \sec \left( \frac{\pi \delta z \lambda \alpha^2}{\lambda_i^2} \right)
\] (22)

Equation (22) is the same as equation (35) in [9] where it is written as \(\sin^{-1}\) instead of secant.

To summarize, we have explicitly written a prescription for deriving \(\beta_p\) for wavefront sensors assuming a closed form description of the \(k\)-th intensity measurement and corresponding variance from a wavefront sensor of interest. In [9], this form of analysis is applicable for the curvature wavefront sensor and the Mach-Zehnder interferometer. For the rest of the techniques considered, numerical simulations are often required as well, since the corresponding closed form descriptions are intractable, or at least much more difficult to work with compared to their simulation representations.

### 5. CONSIDERATIONS FOR \(\beta_p\) FOR FUTURE WORK

The evaluation of \(\beta_p\) for the maximal residual phase error \(\Sigma_{\text{max}}\) is given in Table 1 with respect to intensity measurements for point sources in the presence of sinusoidal phase aberrations across the pupil plane. In the final column, we also review if the wavefront sensing technique has been considered, to our knowledge, for use in extended scene imaging. Furthermore, in [9], the ideal wavefront sensing performance is achieved by the Zernike WFS at any spatial frequency with a value of 1.

<table>
<thead>
<tr>
<th>WFS Technique</th>
<th>(\beta_p) for (\Sigma_{\text{max}})</th>
<th>Use in extended scene imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHWFS</td>
<td>(&gt;2)</td>
<td>Yes</td>
</tr>
<tr>
<td>PyWFS - Fixed</td>
<td>(\sqrt{2})</td>
<td>No</td>
</tr>
<tr>
<td>PyWFS - Modulated</td>
<td>(\geq 2)</td>
<td>No</td>
</tr>
<tr>
<td>CWFS</td>
<td>(\geq 1)</td>
<td>?</td>
</tr>
<tr>
<td>FPWFS</td>
<td>(\sqrt{2})</td>
<td>Yes</td>
</tr>
<tr>
<td>Mach-Zehnder</td>
<td>2</td>
<td>?</td>
</tr>
<tr>
<td>Zernike WFS</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>ODWFS</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Plenoptic WFS</td>
<td>-</td>
<td>Possibly</td>
</tr>
<tr>
<td>dOTF</td>
<td>-</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1: Photon-noise limited sensitivity parameter \(\beta_p\) for WFS techniques in [9] and beyond.

The immediate next steps are to simulate wavefront sensing techniques and verify the remaining forms of \(\beta_p\) can be derived. This validation step is the primer to extending analysis of \(\beta_p\) to the physics of extended scene imaging, i.e. using the SH correlation tracker and ODWFS, as well as considering the other techniques with respect to wavefront errors representative of the error sources discussed in Section 1.1. We will use a python-based simulation suite developed for high-contrast imaging called hcipy [21] which has models for the SHWFS and PyWFS, as well as wavefront reconstruction algorithms, relevant polynomial basis functions such as Zernike and KL modes, and other common tools for AO closed-loop simulations.

The crucial missing piece thus far is to tie \(\beta_p\) directly to image quality. In [9], \(\beta_p\) is related to the image quality metric of contrast, which is a natural choice of parameter to consider in the field of high-contrast imaging. Analogously we can derive a relationship between \(\beta_p\) and some form of image quality metric to be determined.

### ACKNOWLEDGEMENTS

This research is supported by an appointment to the Intelligence Community Postdoctoral Research Fellowship Program at the University of Arizona, administered by Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Office of the Director of National Intelligence.
REFERENCES


