Curvature sensors: noise and its propagation

Aglae Kellerer

Institute for Astronomy,
640 N. A’ohoku Place, HI, Hilo 96720, USA
kellerer@ifa.hawaii.edu

The signal measured with a curvature sensor is here analyzed. In the outset, we derive the required minimum number of sensing elements at the pupil edges, in dependence on the total number of sensing elements. The distribution of the sensor signal is further characterized in terms of its mean, variance, kurtosis and skewness. It is established that while the approximation in terms of a gaussian distribution is correct down to fairly low photon numbers, much higher numbers are required to obtain meaningful sensor measurements for small wavefront distortions. Finally, we indicate a closed expression for the error propagation factor and for the photon-noise induced Strehl loss. © 2010 Optical Society of America

OCIS codes: 110.0115, 110.1080

1. Introduction

In an adaptive optics system, a wavefront sensor measures phase distortions in real-time. These measurements are used to update the shape of a deformable mirror located upstream in the optical beam. One such wavefront sensor was developed by F. Roddier and presented for the first time in 1988 [1]. In a curvature sensor (C-S), two defocused images are recorded, one on each side of the nominal focal plane. The wavefront distortion is then inferred from the difference in intensity between the two images. F. Roddier and his group at the Institute of Astronomy, Hawaii, implemented a low-order adaptive-optics (AO) system based on a curvature WFS with 13 actuators [2]. In the following years many other curvature based AO systems were successfully implemented [3], [4], [5], [6].

However, as adaptive optic systems evolve towards increasingly large numbers of sensing elements and actuators, curvature sensors tend to be discarded in favor of Shack-Hartmann (S-H) sensors. This is mainly based on the assertion that the effect of measurement errors on the corrected phase increases steeply with the number of sensing elements in C-S. This
steep, linear increase of the error propagation term was established by N. Roddier through simulations [7]. The author further reported on concurring analytic calculations by J.B. Shellan for more than 10 000 actuators. These calculations are however not publicly available. High-order curvature systems have been simulated by e.g. O. Lai et al. [8] and Q. Yang et al. [9]. Both authors conclude in general favor of high-order C-S systems. Yang et al. conclude: “No degradation of the phase transfer function with increasing actuator number is observed, implying that still higher actuator number systems are possible with no significant modification of the basic concept and design”. We feel that, in addition to these results, it is necessary to analyze the effect of measurement noise on the performance of curvature systems, and in particular on the performance of high-order systems.

Sec. 2 of this article considers the special role of the sensing elements that lie on the pupil edge. The minimum number of such edge elements is derived in dependence on the total number of sensing elements. This result is then used in section 4 where C-S with up to 225 elements are simulated. Sec. 3 analyzes the measurement noise in curvature sensors: the quantity measured by curvature sensors is \( v = (x - y)/(x + y) \), where the photon numbers \( x, y \) follow a Poisson distribution. Can the distribution of \( v \) be approximated by a Gaussian and is the 68\% -confidence interval well approximated by \( \bar{v} \pm \sigma_v \)? Sec. 3 further derives the number of photons needed to obtain meaningful measurements of \( v \). Sec. 4 studies the effect of measurement noise on the phase reconstruction. We agree with N. Roddier’s result, that were obtained for a fixed size of the sensing elements and a fixed defocus length. However, for a constant pupil surface – hence a decreasing sensing-element size – and with an adjusted extrafocal distance, the error propagation factor is found to be almost independent of the number of sensing elements. Sec. 4 concludes with a closed expression for the error propagation factor and an estimate of the Strehl decrease due to photon noise.

2. Minimum number of edge sensors

2.A. Signal measured by a curvature sensor

The output signal of a curvature wavefront sensor is the intensity difference \( v = (x - y)/(x + y) \), where \( x \) and \( y \) are the photon numbers measured on either side of the focal plane. In closed-loop AO observations, \( v \) tends towards 0 and can therefore be expressed as a function of the wavefront phase, \( \phi \): [10]

\[
v(\vec{r}) = \frac{-\lambda f (f - l)}{2 \pi l} \left[ \nabla (P(\frac{f}{l} \vec{r})) \cdot \nabla (\phi(\frac{f}{l} \vec{r})) + P(\frac{f}{l} \vec{r}) \cdot \nabla^2 \phi(\frac{f}{l} \vec{r}) \right]
\]

\( \vec{r} \) is the position vector in the detector plane. \( f \) is the focal length, \( l \) the distance between the detector and the focal plane and \( f/l \vec{r} \) is thus the position vector in the pupil plane. \( \lambda \) is the wavelength and \( P \) – the pupil transmission function – equals 1 inside and 0 outside the pupil. Eq. 1 makes use of the Fresnel approximation: \( D^2/((f - l)\lambda) \geq 1 \), where \( D \) is
the pupil diameter. Eq. 1 further assumes the size of the diffraction patterns to be negligible compared to the characteristic scales of intensity fluctuations in the detector plane. This sets a condition on the extra focal distance: \(\lambda f^2 << r_0^2 l\), where \(r_0\) is the Fried parameter.

In regions where the pupil transmission is uniform, the measured signal is the laplacian of the wavefront phase. On the pupil edge – i.a. along the central obstruction and along the telescope spiders – curvature sensors measure curvature and radial tilt. Phase aberrations with a laplacian equal to zero are sensed solely at the pupil edges.

2.B. Zero-laplacian Zernike modes

The phase of the distorted wavefront is expressed in terms of a series of Zernike polynomials:

\[
\phi = \sum_{n \geq 0} \sum_{m=-n}^{n} a_{nm}^m Z_{nm}^m
\]  

(2)

\(n\) and \(m\) are integers and \(n - |m|\) is even. \(a_{nm}^m\) is the coefficient of the Zernike polynomial \(Z_{nm}^m\). In atmospherically distorted wavefronts, the lowest order Zernike modes have the highest variance and need to be corrected by the AO system first. The first order Zernike mode, piston, is an exception: it corresponds to a constant phase change over the whole pupil and does not affect the image quality. Table 1 lists the laplacian of the fifteen lowest Zernike polynomials. Eight out of fourteen polynomials (piston excluded) have zero laplacian, and their contribution to the distorted wavefront is therefore only measured at the pupil edges.

In an ideal AO system \(K + 1\) ‘sensor-actuator’ pairs (i.e. \(K + 1\) sub-apertures) are required to sense and correct \(K\) modes of the wavefront phase. It is assumed that the lowest Zernike modes are corrected first (piston excluded). \(K\) is the number of corrected modes and \(n\) the radial index of the highest, corrected mode. Since there are \(p + 1\) Zernike polynomials for each radial index \(p\), \(K\) and \(n\) are related by the following inequality:

\[
\frac{n(n+1)}{2} - 1 < K \leq \frac{(n+1)(n+2)}{2} - 1
\]  

(3)

The Zernike modes for which the absolute value of the meridional index, \(m\), equals the radial index, \(n\), have zero laplacian. There are, consequently, two zero-laplacian polynomials per radial index, hence in order to correct the first \(K\) Zernike polynomials, a curvature sensor needs at least \(N_e = 2n\) sensors at the aperture edge, where \(n\) is determined by Eq. 3:

\[
N_e = 2 \cdot \left\lfloor \frac{1}{2} \left( \sqrt{8K + 9} - 1 \right) \right\rfloor
\]  

(4)

\(\lfloor \rfloor\) is the floor function. Figure 1 traces the minimum number of edge sensors in dependence on the number of modes to be corrected. Instead of increasing the number of edge sensors, it is also possible to smoothen the pupil edges. This distributes the radial tilt signal over more sensors, at the expense of smaller signal in each individual sensor.
Table 1. Laplacian of Zernike polynomials: $\nabla^2 Z(r, \theta) = \frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial Z}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 Z}{\partial \theta^2}$

| n  | m | 0  | 1  | 2  | 3  | 4  |
|----|---|----|----|----|----|----|
| 0  | 0 | $\nabla^2 Z_0^0 = 0$ | \begin{align*} 
\nabla^2 Z_1^{-1} &= 0 \\
\nabla^2 Z_1^1 &= 0 \\
\text{tip & tilt} 
\end{align*} | \begin{align*} 
\nabla^2 Z_2^{-2} &= 0 \\
\nabla^2 Z_2^2 &= 0 \\
\text{astigmatism (3rd order)} 
\end{align*} | \begin{align*} 
\nabla^2 Z_3^{-3} &= 0 \\
\nabla^2 Z_3^3 &= 0 \\
\text{trefoil} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-4} &= 0 \\
\nabla^2 Z_4^4 &= 0 \\
\text{ashtray} 
\end{align*} | 
| 1  | 1 | $\nabla^2 Z_1^0 = 8\sqrt{3}$ | \begin{align*} 
\nabla^2 Z_1^{-1} &= 0 \\
\nabla^2 Z_1^1 &= 0 \\
\text{defocus} 
\end{align*} | \begin{align*} 
\nabla^2 Z_2^{-2} &= 0 \\
\nabla^2 Z_2^2 &= 0 \\
\text{astigmatism (3rd order)} 
\end{align*} | \begin{align*} 
\nabla^2 Z_3^{-3} &= 0 \\
\nabla^2 Z_3^3 &= 0 \\
\text{trefoil} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-4} &= 0 \\
\nabla^2 Z_4^4 &= 0 \\
\text{ashtray} 
\end{align*} | 
| 2  | 2 | $\nabla^2 Z_2^0 = 24\sqrt{8}r \sin(\theta)$ | \begin{align*} 
\nabla^2 Z_2^{-2} &= 0 \\
\nabla^2 Z_2^2 &= 0 \\
\text{astigmatism (3rd order)} 
\end{align*} | \begin{align*} 
\nabla^2 Z_3^{-3} &= 0 \\
\nabla^2 Z_3^3 &= 0 \\
\text{trefoil} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-4} &= 0 \\
\nabla^2 Z_4^4 &= 0 \\
\text{ashtray} 
\end{align*} | 
| 3  | 3 | $\nabla^2 Z_3^{-1} = 24\sqrt{8}r \sin(\theta)$ | \begin{align*} 
\nabla^2 Z_3^{-2} &= 0 \\
\nabla^2 Z_3^2 &= 0 \\
\text{astigmatism (3rd order)} 
\end{align*} | \begin{align*} 
\nabla^2 Z_3^{-3} &= 0 \\
\nabla^2 Z_3^3 &= 0 \\
\text{trefoil} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-4} &= 0 \\
\nabla^2 Z_4^4 &= 0 \\
\text{ashtray} 
\end{align*} | 
| 4  | 4 | $\nabla^2 Z_4^0 = 24\sqrt{5}(4r^2 - 1)$ | \begin{align*} 
\nabla^2 Z_4^{-2} &= 120\sqrt{10}r^2 \cos(2\theta) \\
\nabla^2 Z_4^2 &= 120\sqrt{10}r^2 \sin(2\theta) \\
\text{astigmatism (5th order)} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-3} &= 0 \\
\nabla^2 Z_4^3 &= 0 \\
\text{trefoil} 
\end{align*} | \begin{align*} 
\nabla^2 Z_4^{-4} &= 0 \\
\nabla^2 Z_4^4 &= 0 \\
\text{ashtray} 
\end{align*} | 

Fig. 1. Minimum number of sensing elements at the aperture edge required to sense the $N$ lowest Zernike modes.
3. Uncertainty of the sensor measurement

In this section, a single sub-aperture of the sensor is considered. Be \( x \) and \( y \) the photon numbers measured on either side of the focal plane. Be \( X \) and \( Y \) their expected values, i.e. the number of photons that one would measure in average under frozen atmospheric conditions and repeated measurements. The difference between \( x, y \) and \( X, Y \) is taken to be solely due to photon noise. This is a valid approximation if the sky background signal is low and if the detector readout noise and dark current are negligible. Existing curvature sensors use avalanche photo-diodes where the readout noise and dark current are indeed negligible.

The variable \( v = (x - y)/(x + y) \) serves as estimate of the deviation \( V = (X - Y)/(X + Y) \). \( x \) and \( y \) are independent Poisson distributed random variables with mean value \( X \) and \( Y \).

To judge the precision of the estimate the parameters of the distribution of \( v \) need to be known, i.e. the mean, the variance, the skewness and the kurtosis. These are closely related to the moments or, more specifically, the cumulants, \( \kappa_i \), of the distribution. In closed-loop AO observations, \( V \) tends towards 0. Furthermore the photon numbers, \( Z = X + Y \), are usually large. The parameters of the distribution of \( v \) are then, as will be seen, simple: the mean and the skewness are zero and both the variance and the excess kurtosis are \( 1/Z \). It is nevertheless of interest to derive the exact solution for the more general case where \( V \) is not zero, and \( Z \) is not very large.

3.A. Non-central moments of \( x/z \)

It is convenient to deal first with the variable \( x/z \), rather than \( v = 2x/z - 1 \). The cumulants, \( \kappa_i \) of \( v \) are subsequently expressed in terms of the cumulants, \( \kappa_i \), of \( x/z \):

\[
\kappa_1 = 2k_1 - 1 \quad (5)
\]
\[
\kappa_2 = 4k_2 \quad (6)
\]
\[
\kappa_3 = 8k_3 \quad (7)
\]
\[
\kappa_4 = 16k_4 \quad (8)
\]

The joint distribution of \( x \) and \( z \) equals the product of the Poisson probability, \( p(z; Z) \), for the variable \( z \) with mean value \( Z \), and the binomial probability, \( b(x; z, p) \), that out of the \( z \) events \( x \) are of the type that occurs with relative probability \( p = X/Z \):

\[
f(x, z) = p(z, Z) \cdot b(x; z, p) \quad (9)
\]

The outcome \( z = 0 \) is excluded since \( v \) is then undefined. The probabilities \( p(z; Z) \) are thus normalized to the positive integers \( z \), i.e. are increased by the factor \( 1/(1 - \exp(-z)) \) relative to the familiar Poisson probabilities.
The \(i\)-th moment, \(u_i = \langle (x/z)^i \rangle\) can be written in the form:

\[
\langle (x/z)^i \rangle = \sum_{z>0} \left( \sum_{x=0}^{z} x^i b(x; z, p) \right) z^{-i} \frac{p(z, Z)}{1 - e^{-z}}
\]

(10)

\[
= \sum_{z>0} \mu_i(z, p) z^{-i} \frac{p(z, Z)}{1 - e^{-z}}
\]

(11)

where \(\mu_i(z, p)\) are the moments about zero of the binomial distribution with mean value \(z\) and parameter \(p\). Using the familiar equations for these moments one obtains the expressions:

\[
\mu_1/z = p
\]

(12)

\[
\mu_2/z^2 = (p - p^2)/z + p^2
\]

(13)

\[
\mu_3/z^3 = (p - 3p^2 + 2p^3)/z + 3(p^2 - p^3)/z + p^3
\]

(14)

\[
\mu_4/z^4 = (p - 7p^2 + 12p^3 - 6p^4)/z^3 + (7p^2 - 18p^3 + 11p^4)/z^2 + 6(p^3 - p^4)/z + p^4
\]

(15)

The moments \(u_i = \langle (x/z)^i \rangle\) of the variable \(x/z\) are obtained by summation over \(z\) according to Eq. 11. The problem is thus reduced to the determination of the Poisson mean values \(< z^{-i} >\). Since there are no closed expressions, one needs to evaluate those mean values numerically, which is conveniently done in term of the ratios:

\[
\rho_i = < z^{-i} > /Z^{-i}
\]

(16)

\[
= Z^i \sum_{z>0} z^{-i} \frac{p(z, Z)}{1 - e^{-z}}
\]

(17)

\[
= Z^i \sum_{z>0} z^{-i} \frac{e^{-Z} \cdot Z^z}{z!} \cdot \frac{1}{1 - e^{-z}}
\]

(18)

The \(\rho_i\) converge to unity at large \(Z\).

3.B. Transition to the cumulants

The cumulants of the distribution of \(x/z\) stand in a general relation to the moments:

\[
k_1 = u_1
\]

(19)

\[
k_2 = u_2 - u_1^2
\]

(20)

\[
k_3 = u_3 - 3u_2u_1 + 2u_1^3
\]

(21)

\[
k_4 = u_4 - 4u_3u_1 + 6u_2u_1^2 - 3u_1^4 - u_3 + 3u_2u_1 - 2u_1^3
\]

(22)
Using Eqs. 13-18 and switching in terms of Eqs. 6-8 to the cumulants of \( v \) one obtains:

\[
\begin{align*}
\kappa_1 &= 2p - 1 \\
\kappa_2 &= 4(p - p^2)\rho_1/Z \\
\kappa_3 &= 8(p - 3p^2 + 2p^3)\rho_2/Z^2 \\
\kappa_4 &= 16(p - 7p^2 + 12p^3 - 6p^4)Z^2 + 16(3p^2 - 6p^3 + 3p^4)(\rho_2 - \rho_1^2)/Z^2
\end{align*}
\] (24)

Instead of the parameter \( p = X/Z \), i.e. the mean value of \( x/z \), one wishes to use the parameter \( V = (X - Y)/(X + Y) \):

\[
\begin{align*}
\kappa_1 &= V \\
\kappa_2 &= (1 - V^2)\rho_1/Z \\
\kappa_3 &= -2V(1 - V^2)\rho_2/Z^2 \\
\kappa_4 &= (-2 + 8V^2 - 6V^4)\rho_3/Z^3 + 3(1 - V^2)^2(\rho_2 - \rho_1^2)/Z^2
\end{align*}
\] (28)

The commonly used parameters of the distribution of \( v \), i.e. the mean value \( <v> = \kappa_1 \), the variance, \( \sigma_v^2 = \kappa_2 \), the skewness, \( \text{sk}_v = \kappa_3/\kappa_2^{1.5} \) and the (excess) kurtosis, \( \text{kur}_v = \kappa_4/\kappa_2^2 \) are, therefore:

\[
\begin{align*}
<v> &= V \\
\sigma_v^2 &= (1 - V^2)\rho_1/Z \\
\text{sk}_v &= -\frac{2V}{\sqrt{1 - V^2}}\frac{1}{\rho_1^{1.5}\sqrt{Z}} \\
\text{kur}_v &= \left(\frac{4V^2}{1 - V^2} - 2\right)\rho_3\frac{1}{\rho_1^2\sqrt{Z}} + 3\left(\frac{\rho_2}{\rho_1} - 1\right)
\end{align*}
\] (32)

For large \( Z \), \( \rho_1 \) tends to unity, and, as can be shown, \( (\rho_2/\rho_1^2 - 1) \) converges to \( 1/Z \). The formulae then reduce to:

\[
\begin{align*}
\sigma_v^2 &= (1 - V^2)/Z \\
\text{sk}_v &= -\frac{2V}{\sqrt{1 - V^2}}\frac{1}{\sqrt{Z}} \\
\text{kur}_v &= \left(\frac{4V^2}{1 - V^2} + 1\right)\frac{1}{Z}
\end{align*}
\] (36)

As stated in the outset, the parameters become very simple for small \( V \), i.e. for nearly equal light intensities in the two images, and if \( Z \) is large. The mean and skewness of \( v \) are then zero and the variance and the (excess) kurtosis are \( 1/Z \).

3.C. Visualisation

Fig. 2 gives the standard deviation and the excess kurtosis of the measured \( v \) for the case of equal light intensity of the two AO images, i.e. for \( V = 0 \). It shows that the kurtosis is close
to zero even for fairly low photon numbers, $Z$. The approximation in terms of the Gaussian distribution is, thus, justified and the standard error range in terms of the 68%-confidence range for a measured value $v$ can be set equal to $v \pm \sigma$ while the 95%-confidence range is $v \pm 2\sigma$. Since $\sigma$ still has the substantial value 0.1 for $Z = 100$, one requires much higher photon numbers to obtain meaningful measurements at small values of $V$. Even at the large photon number $Z = 10^4$ the standard error of $v$ is 0.01, i.e. one needs a measured value $v \geq 0.02$ to reject the null-hypothesis of equal light intensity in the two images ($V = 0$) on the 95%-confidence level. This is not trivial, since the measured $v$-values tend to be small in curvature-sensor systems for AO.

Fig. 2. The standard deviation and the kurtosis of $v$ in dependence on $Z$ for $V = 0$: The approximation in terms of the Gaussian distribution is justified down to fairly low photon values (right panel), but much higher photon numbers are required to obtain meaningful measurements at small values of $V$, i.e. for small wavefront distortions (left panel).

This being a critical issue, it is useful to consider the distribution parameters not just for $V = 0$. While fairly low values of $V$ are important in practice, it is instructive to take a broader view and consider the full range of possible intensity differences. Fig. 3 gives for $Z = 100$ the standard error, the skewness and the kurtosis in dependence on $V$. As seen in Eqs. 37-38, the shape of these dependencies remains the same for higher photon numbers, $Z$, while the absolute values of the standard deviation and the skewness decrease proportionally to $1/\sqrt{Z}$ and the kurtosis decreases proportionally to $1/Z$. At the large photon numbers $Z$ that are required the Gaussian approximation remains thus valid inspite of the increase of the
absolute value of the skewness and the kurtosis at large intensity differences. The decrease of the standard error is helpful in principle, but is not substantial at moderate values of $V$.

Fig. 3. The standard deviation (solid line), the skewness (dashed) and the kurtosis (dotted) of $v$ in dependence on $V$ for $Z = 100$. The decrease of the standard error is not substantial in the [0,0.2] range representative of closed-loop adaptive optics observations. At the large photon numbers that are thus required, the Gaussian approximation remains valid despite the increase of the skewness and kurtosis.

Fig. 4 illustrates the need for large photon numbers directly. It gives for any measured value $v$ and for different photon numbers the 68%-confidence range of the estimated value $V$.

4. **Error propagation**

4.A. **Definition and context**

An error on the sensor measurement translates into an error on the reconstructed or corrected wavefront: Inside a correction loop, the error on the sensor measurement propagates on the voltages through the *command matrix*, onto the mirror surface through the actuator *influence functions* and finally to the phase as the wavefront reflects on the mirror. If the sensor is not part of an AO control-loop, the error propagates through a phase reconstruction algorithm. The error propagation factor equals the phase variance due to statistically independent
The need for large photon numbers: percent standard error in dependence on the measured value $v$ for photon numbers $Z = X + Y = 100$ (solid line), $10^3$ (dashed), $10^4$ (dotted) and $10^5$ (dashed-dotted).

Random measurements of unit variance. It is defined by the following relation.

$$G = \frac{\sigma_\phi^2}{\sigma_v^2}$$  \hspace{1cm} (39)

where $\sigma_v^2$ is the variance of the sensor measurements and $\sigma_\phi^2$ the induced variance of the corrected or reconstructed phase.

The error propagation factor in Shack-Hartmann sensors has been analyzed by Hudgin [11] and Fried [12]. Both, Hudgin’s and Fried’s wavefront-reconstruction algorithms lead to a logarithmic increase of the error propagation term with the number of sensing elements: $G \propto \ln(N)$. Both authors assume that the size of the sampling elements is constant with changing $N$.

4.B. Curvature sensors

Simulations by N. Roddier for actuator numbers in the 5-100 range have shown that the error propagation term in curvature sensors increases linearly with $N$ [7]. The behavior of the error propagation term is here analyzed. It is found – in agreement with N. Roddier’s simulations – that $G$ increases linearly with the number of sub-apertures, if the pupil size increases with $N$ and if the defocus distance is kept fixed. The behavior of the error propagation term for a given pupil surface and an increasing defocus distance is then examined and a closed expression for the error propagation factor is indicated.
4.B.1. Phase reconstruction

To obtain an expression for the error propagation factor in curvature sensors, we adapt Hudgin’s phase-reconstruction algorithm. Let us briefly recall the principle of this algorithm [11]: A S-H sensor determines the phase difference, e.g. \( d\phi_x(l, n) = \phi(l+1, n) - \phi(l, n) \) and \( d\phi_y(l, n) = \phi(l, n+1) - \phi(l, n) \). For a set of noiseless slope measurements, the phase value at a grid point \((i, k)\) relates to the phase values of its four nearest neighbors:

\[
\phi(l, n) = \frac{\phi(l-1, n) + d\phi_x(l-1, n) + \phi(l+1, n) - d\phi_x(l, n) + \\
\phi(l, n-1) + d\phi_y(l, n-1) + \phi(l, n+1) - d\phi_y(l, n)}{4}
\]

This relation only applies to inner grid points and Hudgin did consider an infinitely large pupil. For corner (resp. border) points we modify this relation to use only the two (resp. three) nearest neighbors. The denominator in Eq. 40 is accordingly changed to 2 (resp. 3).

For slope values with measurement errors there is no perfect fit that fulfills this condition, but the best fit minimizes the mean squared deviations from the equality. Hudgin referred, accordingly, to a reconstruction method where one starts with an arbitrary phase estimate, \( \phi(i, k) \). In each iteration cycle the phase estimates on each grid point are replaced by the average of the phase values of its four closest neighbors, each adjusted by the fixed values of the phase differences:

\[
\phi(l, n) \leftarrow \frac{\phi(l-1, n) + \phi(l+1, n) + \phi(l, n-1) + \phi(l, n+1)}{4} + d^2\phi(l, n)
\]

The new value \( \phi(l, n) \) is the one that best fits the 4 latest phase estimates for the nearest neighbors and the values of the associated slopes. Eq. 41 suggests a simplification: The sum of the 4 phase-difference terms in the equation is an estimate of the second-order phase-difference at grid point \((l, n)\):

\[
d^2\phi(l, n) = \frac{(d\phi_x(l-1, n) - d\phi_x(l, n) + d\phi_y(l, n-1) - d\phi_y(l, n))/4}{4}
\]

And Eq. 41 can, accordingly, be written in the simplified form:

\[
\phi(l, n) \leftarrow \frac{\phi(l-1, n) + \phi(l+1, n) + \phi(l, n-1) + \phi(l, n+1)}{4} + d^2\phi(l, n)
\]

with the modification that has been noted for the border and corner points.

For a set of \( N \) sensor measurements \( v(l, n) \) \((l, n = 1..\sqrt{N})\) the second-order phase-differences are obtained through:

\[
d^2\phi(l, n) = \nabla^2 \phi(l, n) \cdot a^2 / 4 = -\frac{\pi la^2}{2 \lambda f (f-l)} \cdot v(l, n)
\]

where \( a \) is the distance between two adjacent grid points. The \( N \) phase values are obtained via the iteration method described by Eq. 43.
4.C. Sensing elements of constant surface and fixed extra-focal distance

The error propagation factor is computed for the parameter values listed in the first column of Table 2: a fixed sub-aperture size \( a \) and extra-focal distance \( l \). In agreement with N. Roddier’s [7], \( G \) increases linearly with the number of sensing elements:

\[
G = (2.20 \pm 0.05) \times 10^3 \cdot N - (2.03 \pm 0.05) \times 10^4 
\]

(45)

Table 2. Parameter values for the computation of the error propagation term.

| Parameter                           | Sec. 4.C | Sec. 4.D | Sec. 4.E |
|-------------------------------------|----------|----------|----------|
| Pupil shape:                        | square   | square   | circular |
| Pupil size, \( D \):                | \( a \cdot \sqrt{N} \) | 8 m      | 3 m      |
| Sub-aperture size, \( a \):         | 1 m      | \( D/\sqrt{N} \) | \( D/\sqrt{N} \) |
| Extra-focal distance, \( l \):      | 0.5 m    | 0.05 \( \sqrt{N} \) [m] | 0.8 \( \sqrt{N} \) [m] |
| Wavelength, \( \lambda \):          | 0.7 \( \mu m \) | 0.7 \( \mu m \) | 0.7 \( \mu m \) |
| focal length, \( f \):              | 120      | 120      | 180      |
| Number of sensing elements, \( N \):| 25..225  | 25..225  | 25..225  |

4.D. Fixed pupil surface and adjusted extra-focal distance

On a given telescope and for a constant pupil surface, the sub-aperture size decreases with the number of sensing elements: \( a = D/\sqrt{N} \). The measurements would eventually get blurred by diffraction if the extra-focal distance was kept constant, because the diffraction pattern has a fixed size in the image plane: \( \lambda f/r_0 \), where \( r_0 \) is the Fried parameter. The extra-focal distance therefore needs to be increased, \( l \propto \sqrt{N} \), so that the size of the sensing elements in the imaging plane, \( a l/f \), stay constant.

The error propagation factor is thus computed for the parameter values listed in the second column of Table 2: a constant pupil surface and an adjusted extra-focal distance. As seen on Fig. 5, \( G \) is then almost independent of the number of sensing elements. A general expression for the error propagation factor is deduced from these numerical results, using Eq. 44:

\[
G = \frac{\sigma^2}{\sigma^2_{\phi}} = G_0 \cdot \left[ (0.85 \pm 0.05) - \frac{(6.0 \pm 0.5)}{N} \right] 
\]

(46)

\[
G_0 = \left( \frac{l}{\sqrt{N}} \right)^2 \cdot \left( \frac{D}{f} \right)^4 \cdot \frac{1}{\lambda^2} 
\]

(47)
This equation translates that the error propagation factor increases linearly with $N$ for a unit sampling size and a constant extra-focal distance (Eq. 45), and that it is almost independent of $N$ when the pupil surface is constant and the extra-focal distance is adjusted (see Fig. 5).

![Error propagation term, $G$, in dependence on the number of sub-apertures.](image)

**Fig. 5.** Error propagation term, $G$, in dependence on the number of sub-apertures. **Full line:** in terms of the iterative method for the parameters values listed in the second column of Table 2 (square pupil). **Dotted line:** in terms of the adaptive-optics simulation package *yao* for the parameter values listed in the third column of Table 2 (circular pupil). $G_0$ is defined by Eq. 47.

4.E. **Comparison against an adaptive-optics simulation package**

In this section, the values obtained in sec. 4.C and 4.D are checked against those obtained with the adaptive-optics simulation package *Yao* [13]. In sections 4.C and 4.D the error propagation factor has been calculated on square grids, while with *Yao* the pupil is circular. The sampling geometry is inspired by Roddier et al. [2]: the sub-apertures are distributed along concentric rings and cover equal surfaces. The outer elements have the highest error propagation factors and it is therefore desirable to minimize their number\(^1\): we choose the number of outer elements to be above the minimum required number (Eq. 4) by 10-20%. Figure 6 shows two of the eleven generated sensor configurations.

The procedure adopted to calculate the error propagation factor under *Yao* is the following:

\(^1\)Without any sensing element on the pupil edge, the error propagation factor is reduced by a factor of about 5.
– given \(N\) sub-apertures, \(Yao\) computes \(N\) influence functions: \(N\) wavefronts resulting from unit voltages applied to actuators \(i = 1\ldots N\).

– the command matrix, which converts a set of \(N\) sensor measurements into \(N\) voltage values, is determined from its inverse, the interaction matrix.

– for statistically independent random measurements of standard deviation \(\sigma_v = 0.1\), the set of corresponding voltage values is obtained through the command matrix. The wavefront is the sum of the \(N\) influence functions weighted by the \(N\) voltages. It is expressed in radians.

– \(\sigma^2_\phi(i)\) is the phase variance of the \(i\)-th wavefront realization. The error propagation factor is determined from the following equation:

\[
G = \frac{1}{1000} \sum_{i=1}^{1000} \frac{\sigma^2_\phi(i)}{\sigma^2_v} \quad (48)
\]

As seen on Fig. 5, the error propagation factors obtained with \(Yao\) are in good agreement with the values obtained in terms of the iterative method (sec. 4.C and 4.D). The fact that they do not coincide is probably due to the iterative method and \(Yao\) considering square and circular pupils respectively. A general expression for the error propagation factor on a circular pupil is derived:

\[
G = \frac{\sigma^2_\phi}{\sigma^2_v} = \left(\frac{l}{\sqrt{N}}\right)^2 \cdot \left(\frac{D}{f}\right)^4 \cdot \frac{1}{\lambda^2} \cdot \left[\left(0.64 \pm 0.04\right) - \frac{(2.7 \pm 0.3)}{N}\right] \quad (49)
\]

Fig. 6. Configurations of two of the simulated sensors: the wavefront is sampled over 25 and 225 elements. Dashed line: Pupil edge.
The decrease in Strehl due to photon noise is then given by the following relation:

\[
\frac{S_1}{S_0} = \exp(-\sigma^2_\phi) = \exp\left(-G \cdot \frac{N}{Z}\right) \tag{50}
\]

\[
\frac{S_1}{S_0} = \exp\left(-\frac{G_0}{Z} \cdot [(0.64 \pm 0.04) N - (2.7 \pm 0.3)]\right) \tag{51}
\]

where \(S_1\) and \(S_0\) are the Strehl ratios with and in absence of photon noise. \(G_0\) is defined by Eq. 47, \(N\) is the number of sensing elements and \(Z\) the total number of photons collected by the sensor during the acquisition of the two extra-focal images. Figure 7 traces the Strehl loss due to photon noise for an 8 m telescope, in dependence on the number of sensing elements and on the photon number.

Fig. 7. Decrease of the instantaneous Strehl ratio due to photon noise in dependence on the number of sensing elements \(N\) and on the average number of photons, \(Z\), received by the sensor during the acquisition of the two extra-focal images. Other parameter values are: \(l = 0.05\sqrt{N} \text{ [m]}\), \(D = 8 \text{ m}\), \(f = 120 \text{ m}\) and \(\lambda = 0.7 \mu\text{m}\).

5. Conclusion

In this article, we have derived – for any given curvature sensor – a minimum required number of sub-apertures on the pupil-edge, in dependence on its total sub-aperture number.

We further examined the distribution of the sensor signal in terms of its mean, variance, skewness and kurtosis. The approximation in terms of a Gaussian distribution is found to be correct, even for fairly low photon numbers in the 50-100 range (mean photon number per
sub-aperture and iteration). Yet, one requires much higher photon numbers (in the $10^3 - 10^4$ range) to obtain meaningful measurements at small signal values, i.e. at small wavefront phase distortions.

Finally, and in accordance with N. Roddier, the error propagation term has been shown to increase linearly with $N$ – the number of sub-apertures – if the sampling-surface stays constant. That is, if the pupil-surface likewise increases linearly with $N$. For a given telescope, however, and a fixed pupil-surface, the error propagation term is almost independent of $N$. This has been derived by an iterative method inspired by Hudgin’s phase-reconstruction algorithm, and has been validated with an adaptive optics simulations software. A closed expression for the error propagation factor has been suggested.

Acknowledgments

I thank Mark Chun and Christ Ftaclas for many helpful discussions, and I am grateful for the referees’ constructive suggestions. Financial support from the United States Air Force Office of Scientific Research is likewise acknowledged.

References

1. F. Roddier, “Curvature sensing and compensation: a new concept in adaptive optics”, Applied Optics, vol. 27, i. 7, p.1223-1225 (1988)
2. F. Roddier, M. Northcott, J.E. Graves, “A simple low-order adaptive optics system for near-infrared applications”, Astronomical Society of the Pacific, Publications (ISSN 0004-6280), vol. 103, p. 131-149 (1991)
3. J. E. Graves, M. J. Northcott, F.J. Roddier, C.A. Roddier, L.M. Close, “First light for Hokupa’a: 36-element curvature AO system at UH”, in Adaptive Optical Systems Technologies, D. Bonaccini and R.K. Tyson, eds. Proc. SPIE, vol. 3353, p. 34-43 (1998)
4. M. Chun, D. Toomey, Z. Wahhaj, B. Biller, E. Artigau, T. Hayward, M. Liu, L. Close, M. Hartung, F. Rigaut, C. Ftaclas, “Performance of the near-infrared coronagraphic imager on Gemini-South”, Adaptive Optics Systems. Edited by N. Hubin, C. Max, P. Wizinowich, Proc. SPIE, vol. 7015, p. 70151V-70151V-9 (2008)
5. R. Arsenault, A. Jaime, H. Bonnet, J. Brynnel, B. Delabre, R. Donaldson, C. Dupuy, E. Fedrigo, J. Farinato, N. Hubin, L. Ivanescu, M. Kasper, J. Paufique, S. Rossi, S. Tordo, S. Stroebele, J.L. Lizon, P. Gigan, F. Delplancke, A. Silber, M. Quattri, R. Reiss, “MACAO-VLTI: An Adaptive Optics system for the ESO VLT interferometer”, Adaptive Optical System Technologies II. Edited by P. L. Wizinowich, D. Bonaccini. Proceedings of the SPIE, vol. 4839, p. 174-185 (2003)
6. M. Watanabe, S. Oya, Y. Hayano, H. Takami, M. Hattori, Y. Minowa, Y. Saito, M. Ito, N. Murakami, M. Iye, O. Guyon, S. Colley, M. Eldred, T. Golota, M. Dinkins, “Im-
plementation of 188-element Curvature-based Wavefront Sensor and Calibration Source Unit for the Subaru LGSAO System”, Adaptive Optics Systems. Edited by N. Hubin, C. Max, P. Wizinowich, Proc. SPIE, vol. 7015, p. 701564-701564-8 (2008)

7. N. Roddier, “Curvature sensing for adaptive optics: a computer simulation”, thesis for the degree of master of science submitted to the University of Arizona (1989)

8. O. Lai, T. V. Craven-Bartle, “Simulations of high order curvature adaptive optics”, Conference Information: Conference on High-Contrast Imaging for Exo-Planet Detection, high contrast imaging for exo-planet detection, SPIE proc, vol. 4860, p. 334-342 (2003)

9. Q. Yang, C. Ftaclas, M. Chun, “Wavefront correction with high-order curvature adaptive optics systems”, JOSA A, vol. 23, I. 6, p. 1375-1381 (2006)

10. F. Roddier, “Curvature sensing: a diffraction theory”, NOAO advanced development program, n. 87-3, (1987)

11. R. Hudgin, “Wave-front reconstruction for compensated imaging”, JOSA, vol. 67, p. 375-378 (1977)

12. D. Fried, “Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements”, JOSA, vol. 67, p. 370-375 (1997)

13. www.maumae.net/yao/