Extended analysis of curvature sensing

Marcos A. van Dam and Richard G. Lane

Department of Electrical and Computer Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

Received September 4, 2001; revised manuscript received February 15, 2002; accepted February 25, 2002

Curvature sensors are used in adaptive optics to measure the wave-front aberrations. In practice, their performance is limited by their nonlinear behavior, which we characterize by solving simultaneously the irradiance transport equation and the accompanying wave-front transport equation. We show how the presence of nonlinear geometric terms limits the accuracy of the sensor and how diffraction effects limit the spatial resolution. The effect of photon noise on the sensor is also quantified. © 2002 Optical Society of America

OCIS codes: 010.1080, 010.1330, 010.7350.

1. INTRODUCTION

An important problem in astronomical imaging is to estimate the aberrations of an incoming wave front. This wave-front estimate can then be used to drive an adaptive optics system or to deconvolve a distorted image to improve the image of the object of interest. There are several different wave-front sensor designs and of these, curvature sensors are the basis of a number of the more successful adaptive optics systems.[1] Despite this, there have been few serious attempts to characterize their nonlinear behavior or noise performance.[2,3] Most analyses have relied instead on a discrete approximation to the irradiance transport equation to predict the relationship between the wave-front phase and the corresponding intensity of the propagated wave front. This approximation does not necessarily hold under the normal operating conditions of a curvature sensor.

Consider an aberrated wave front, $W(x, y)$, propagating in the z direction. The wave front experiences a change in both its intensity and its phase. The instantaneous change in intensity I, $I_z = \partial I/\partial z$, is governed by the well-known irradiance transport equation (ITE)[4]:

$$I_z = -I\nabla^2 W - \nabla I \cdot \nabla W, \tag{1}$$

where $\nabla W = W_x\hat{\mathbf{x}} + W_y\hat{\mathbf{y}}$, $\nabla I = I_x\hat{\mathbf{x}} + I_y\hat{\mathbf{y}}$, and $\nabla^2 W = W_{xx} + W_{yy}$. We denote the unit vectors in the x and y direction as $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, respectively. The ITE was originally derived for the case of monochromatic light, and Streibl extended the results to the case of an incoherent source.[5]

An accurate analysis must also include the effects of propagation on the wave-front phase, which changes in a manner that depends on the wavelength λ. The wave-front transport equation (WTE) is given by[4]

$$W_z = 1 - \frac{1}{2}|\nabla W|^2 + \frac{\lambda^2}{16\pi^2 I}\nabla^2 I - \frac{\lambda^2}{32\pi^2 I^2}|\nabla I|^2, \tag{2}$$

where $|\nabla W|^2 = W_x^2 + W_y^2$. We derive the equations governing curvature sensing using both the ITE and the WTE, since as the wave front propagates, its curvature changes and consequently this causes changes in intensity. We also use the WTE to show that diffraction effects limit the spatial resolution of the curvature measurement. Finally, we derive the relationship between the number of photons and the error of the curvature measurement.

2. INTRODUCTION TO CURVATURE SENSING

Practical curvature sensors were initially proposed in 1988 by Roddier.[6] They reconstruct the wave front from estimates of the radial tilt at the edge of the aperture and of the curvature measured within the aperture. Figure 1 shows the setup initially proposed. A lens of focal length f is placed at the aperture. Two intensity images, $I_1(\mathbf{x})$ and $I_2(\mathbf{x})$, corresponding to blurred images of the aperture, are recorded at a distance l from focus on either side of the focal plane. There is an additional lens of focal length $f/2$ at the focal plane to ensure that the two images are of the same scale. Conventional curvature-sensing analysis assumes a small effective propagation distance, $z = f(f - l)/l$, allowing a finite-difference approximation to the irradiance transport equation.[6] In geometric optics, the intensities of images $I_1(\mathbf{x})$ and $I_2(\mathbf{x})$ are conventionally assumed to be related by[7,8]

$$\frac{I_1(\mathbf{x}) - I_2(-\mathbf{x})}{I_1(\mathbf{x}) + I_2(-\mathbf{x})} = \frac{f(f-l)}{l}\left[\frac{\partial W}{\partial n}\left(\frac{f\mathbf{x}}{l}\right)\delta_c - P\nabla^2 W\left(\frac{f\mathbf{x}}{l}\right)\right], \tag{3}$$

where P is the transmission function of the aperture, δ_c is a Dirac distribution around the aperture edge, and $\partial W/\partial n$ is the outward wave-front derivative perpendicular to the aperture edge. Equation (3) indicates that the signal consists of two terms: the curvature measurement due to the Laplacian of the wave front and an edge signal proportional to the slope of the wave front at the edge of the aperture. The wave front is reconstructed by solving the Poisson equation, using the edge signal to provide Neumann boundary conditions.

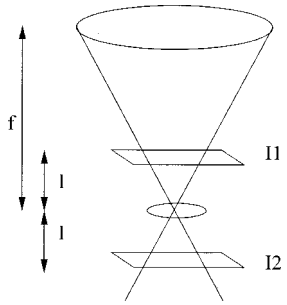

Fig. 1. Standard curvature-sensing arrangement comprising two intensity measurements taken on planes at distance l from the focal plane.

3. NONLINEAR EFFECTS

A. Wave-Front Propagation

Although, traditionally, analyses of the curvature sensor have ignored the effects of change in the wave front, it is clear that a change in the wave front must ultimately be reflected as a change in the intensity distribution. Unlike the ITE, the WTE is wavelength dependent, with the term

$$W_z^{\text{diff}} = \frac{\lambda^2}{16\pi^2 I}\nabla^2 I - \frac{\lambda^2}{32\pi^2 I^2}|\nabla I|^2 \tag{4}$$

encompassing the effect of diffraction on the wave front. Even when diffraction is not considered, the importance of incorporating the wave-front change given by the WTE in the analysis of the curvature sensor can be seen by integrating both sides of the ITE:

$$\int \frac{\mathrm{d}I}{I} = -\int \nabla^2 W \mathrm{d}z - \int \frac{\nabla I \cdot \nabla W}{I}\mathrm{d}z. \tag{5}$$

The wavelength-independent component of the WTE, which represents the change in the wave front described by geometric optics, is

$$W_z = 1 - \frac{1}{2}|\nabla W|^2. \tag{6}$$

With Eq. (6), the first-order Taylor series expansion in z for the wave front is

$$W(z) = W(0) + zW_z(0) + O(z^2) \approx W + z - \frac{z}{2}|\nabla W|^2, \tag{7}$$

where $O(z^2)$ denotes terms of order z^2 or higher. For compactness of notation, we drop the argument of W, I, and all their derivatives when they are evaluated at $z = 0$. With this approximation to the wave front inside the ITE, the first term on the right-hand side of Eq. (5) can be written as

$$\begin{aligned}\int \nabla^2 W \mathrm{d}z &= z\nabla^2 W - \frac{z^2}{4}\nabla^2|\nabla W|^2 + O(z^3)\\ &\approx z\nabla^2 W - \frac{z^2}{2}(W_{xx}^2 + W_{yy}^2 + 2W_{xy}^2\\ &\quad + W_x W_{xxx} + W_x W_{xyy} + W_y W_{xxy}\\ &\quad + W_y W_{yyy})\\ &= z\nabla^2 W - \frac{z^2}{2}[(\nabla^2 W)^2 - 2W_{xx}W_{yy} + 2W_{xy}^2\\ &\quad + W_x W_{xxx} + W_x W_{xyy} + W_y W_{xxy}\\ &\quad + W_y W_{yyy}]\\ &= zH - \frac{z^2}{2}(H^2 - 2K + T).\end{aligned} \tag{8}$$

Here, the mean curvature $\nabla^2 W$ is denoted by H, the Gaussian curvature $W_{xx}W_{yy} - W_{xy}^2$ by K, and the term $W_x W_{xxx} + W_x W_{xyy} + W_y W_{xxy} + W_y W_{yyy}$ by T. Any terms of order z^3 or higher are truncated, because their effect is smaller than the z^2 terms in the region where the curvature sensor is operated. Conventionally, the right-hand side is approximated by just the first term, zH, yielding a curvature-sensing signal proportional to the mean curvature. However, the presence of terms in z^2 renders the relationship between the change in intensity and the mean curvature nonlinear, an effect that becomes more severe as the propagation distance increases. This effect limits the range of z over which the curvature sensor can be operated.

The nonlinearity caused by the Gaussian curvature term K was encountered by Milman *et al.* using a geometric optics approach.[2] The T term has not previously been noted and occurs because a local tilt displaces the wave front, so that the measured intensity corresponds to the curvature at a point in the aperture shifted by $z\nabla W$ relative to the measured point. Figure 2 illustrates this source of the T term.

The effect of any variation in intensity of the propagating wave on the curvature-sensing signal is affected by

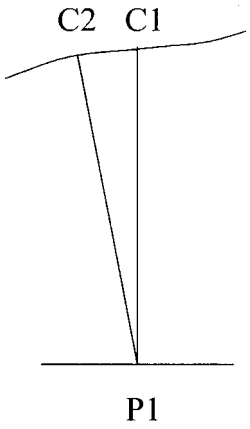

Fig. 2. Effect of the slope on the curvature measurement. The intensity variations measured at P1 are due to the curvature at C2, not at C1, because of the displacement caused by the local tilt in the wave front.

the local wave-front slope. Another identical term proportional to T arises from the second term on the right-hand side of Eq. (5). To show this formally, we assume that the turbulence is weak so that the intensity is constant at the aperture, which implies that at $z = 0$, $\nabla I = 0$ away from the edges. However, as the wave front propagates, the intensity starts to vary. The Taylor series expansion of ∇I is

$$\begin{aligned}\nabla I(z) &= \nabla I + z\nabla\frac{\partial I}{\partial z} + O(z^2)\\ &= -z\nabla(I\nabla^2 W)\\ &= -zI\nabla(\nabla^2 W)\\ &= -zI[(W_{xxx} + W_{xyy})\hat{\mathbf{x}} + (W_{xxy} + W_{yyy})\hat{\mathbf{y}}]. \quad (9)\end{aligned}$$

This result was obtained by substituting the right-hand side of the ITE for I_z and recalling the assumption that $\nabla I(0) = 0$. The second term on the right-hand side of Eq. (5) takes the value

$$\int \frac{\nabla I \cdot \nabla W}{I}\mathrm{d}z = -\frac{z^2}{2}T + O(z^3) \quad (10)$$

within the aperture. Overall, the solution to Eq. (5) is thus

$$\ln(I(z)/I_0) = -zH + \frac{z^2}{2}(H^2 - 2K + 2T) + O(z^3), \quad (11)$$

where I_0 is the constant intensity within the aperture at $z = 0$. The logarithmic change in intensity with z was first noted by Streibl.[5] Using the power series expansion $\ln(1 + a) = a - a^2/2 + O(a^3)$, we obtain

$$\ln(I(z)/I_0) = -\ln(1 + zH + z^2K - z^2T) + O(z^3). \quad (12)$$

Equation (12) can be approximated as

$$I(z) \approx \frac{I_0}{1 + zH + z^2(K - T)}. \quad (13)$$

In curvature sensing, two intensities I_1 and I_2 are measured at distances $\pm z$. The signal S is usually defined as

$$S = \frac{I_1 - I_2}{I_1 + I_2}, \quad (14)$$

which leads on substitution of Eq. (13) to

$$S = \frac{-zH + O(z^3)}{1 + z^2(K - T)}. \quad (15)$$

If we assume that the intensity at any point $I(x, y)$ has a polynomial variation of intensity with z, any dependence of $I(z)$ on the even-order powers of z, such as z^0 (i.e., variations in I_0) and z^2, causes the same change in intensity at both I_1 and I_2. Since the even-order terms affect the curvature signal only by way of the denominator of Eq. (14), this results in a sensor relatively insensitive to intensity fluctuations and scintillation in the aperture. Indeed, if the curvature is perfectly compensated, the signal is equal to zero and the z^2 terms have no effect. However, when this is not the case, any residual K or T terms distort the signal.

The odd-order power of z terms, e.g., $-zH$, result in equal and opposite changes in intensity at the two measurement planes, and this is the basis of the curvature-sensing signal. However, the z^3 terms also contribute to the curvature signal seen by the sensor, constraining the accuracy of the curvature estimate. The effect of diffraction terms in z^3 is investigated in Subsection 4.B, but the computation of all z^3 terms is beyond the scope of this paper.

An alternative to Eq. (14), used to reduce the error when the photon count is very low, is to replace the denominator by the average number of photons, $2I_0$.[9] This gives an estimate that is biased toward zero when there are fewer photons than $2I_0$, thereby reducing the variance of the curvature estimates. In this case,

$$\begin{aligned}S_\star &= \frac{I_1 - I_2}{2I_0}\\ &= \frac{-zH + O(z^3)}{(1 + zH + z^2(K - T))(1 - zH + z^2(K - T))}\\ &= \frac{-zH + O(z^3)}{1 + z^2(2K - 2T - H^2)}. \quad (16)\end{aligned}$$

By direct comparison with Eq. (14), we can see that Eq. (16) contains an additional nonlinear term in the denominator, $-z^2H^2$, and that the magnitude of the $K - T$ term is doubled. As a consequence, using $S_\star$ instead of S means that the measured signal deviates more from the ideal linear model.

B. Expected Values of *H*, *K*, and *T*

The magnitude of the H, K, and T terms determines the modeling error in the curvature sensor. As an example, we consider the problem of estimating the curvature over a circular subregion of the aperture, with the radius of this subregion being α. The values of the H, K, and T terms can be found by expanding the phase aberration over this region, ϕ, as an infinite sum of orthonormal Zernike polynomials:[10]

$$\phi = \sum_{m=1}^{\infty} d_m Z_m, \quad (17)$$

where d_m is the coefficient of the mth Zernike polynomial, Z_m, given by

$$d_m = \pi^{-1}\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\phi Z_m \mathrm{d}x\,\mathrm{d}y. \quad (18)$$

The phase aberration is related to the wave front by

$$\phi = \frac{2\pi}{\lambda}W. \quad (19)$$

We scale the coordinates by α so that the region considered is a circle of unit radius. The value of the H, K, and T terms over the region is found by differentiating the Zernike polynomials and then integrating over the circle. For example, H is given by

$$H = k^{-1}\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\left[\frac{\partial^2\phi}{\partial(\alpha x)^2} + \frac{\partial^2\phi}{\partial(\alpha y)^2}\right]\mathrm{d}x\,\mathrm{d}y, \quad (20)$$

where $k = 2\pi/\lambda$. The calculations of the analytic value of K and T become more and more complicated as the number of Zernike polynomials increases. Evaluating over the first ten Zernike polynomials, we obtain

$$H = \sqrt{192} d_4 \alpha^{-2} k^{-1},$$

$$K = (48d_4^2 - 24d_5^2 - 24d_6^2 + 144d_7^2 + 144d_8^2 - 144d_9^2 - 144d_{10}^2)\alpha^{-4}k^{-2},$$

$$T = (192d_7^2 + 192d_8^2 + \sqrt{18432} d_2 d_8 + \sqrt{18432} d_3 d_7)\alpha^{-4}k^{-2}. \quad (21)$$

The expansion of $(K - T)^2$ up to the third radial order (up to Z_{10}) already contains 45 terms. In addition, the integrals of Eq. (20) do not necessarily have a simple analytic solution, and as a consequence we truncated the polynomial expansion at the third radial order. If we assume Kolmogorov turbulence,[11] we can use the covariance matrix of the d_i terms to evaluate Eq. (20). For example, $E[d_4^2] = 0.0232(2\alpha/r_0)^{5/3}$. Table 1 shows analytic results that have been confirmed by Monte Carlo simulations. Using the first three radial orders of Zernike polynomials only, the root-mean-squared (RMS) value of the desired signal is

$$\sqrt{E[S^2]} = z\sqrt{E[H^2]} = z\sqrt{14.1}\alpha^{-7/6} r_0^{-5/6} k^{-1}. \quad (22)$$

The RMS value of the $z^2(K - T)$ term is

$$z^2\sqrt{E[(K - T)^2]} = z^2\sqrt{1150}\alpha^{-7/3} r_0^{-5/3} k^{-2}. \quad (23)$$

When either α or r_0 decreases, the signal increases by $\alpha^{-7/6} r_0^{-5/6}$ and the error term increases by $\alpha^{-7/3} r_0^{-5/3}$, indicating that the modeling error is more severe when the detector pixels are small and the turbulence is strong. There is also a bias in the signal as $E[K - T]$ is greater than zero. The effect of the bias is to underestimate the curvature.

Care must be taken in interpreting this analysis since the variance of the mean curvature over a pixel appears to increase without bound as the radial order increases. The contribution of the fourth and the sixth radial degrees to the variance of H, due to Z_{11} and Z_{22} respectively, is 21.9 and 61.4 $\alpha^{-7/3} r_0^{-5/3} k^{-2}$. In addition, the wave-front slope at the edge of the aperture is also theoretically unbounded for ideal Kolmogorov turbulence.[12] However, the influence of the high-order aberrations is limited in practice by diffraction effects; this is discussed in Subsection 4.B. As a consequence, it is reasonable to truncate the series to evaluate the behavior of the H, K, and T terms.

Table 1. Relative Weight of the *H*, *K*, and *T* Terms in the Curvature-Sensing Signal

Radial Order	$E[H]$	$\times \alpha^{-7/3} r_0^{-5/3} k^{-2}$			$\times \alpha^{-14/3} r_0^{-10/3} k^{-4}$
		$E[K]$	$E[T]$	$E[H^2]$	$E[(K - T)^2]$
1 and 2	0	0	0	14.1	37.5
1, 2, and 3	0	0	−4.62	14.1	115.0

A feature of the curvature sensor in an adaptive optics system is that there is a strong relationship between the curvature measurements and the driving signal to the deformable mirror.[6] In practical curvature sensors, the geometries of the mirror segmentation and the detector are the same.[9] In closed-loop operation, one can ideally expect to attain perfect compensation of the tip, tilt, and defocus terms within the pixel. Thus the expected value of H is reduced, allowing the sensor to be operated with a larger effective propagation distance z. However, modes corresponding to Z_5 or higher within the region of the aperture subtended by the detector element cannot be detected or compensated. This means that the nonlinearities of the sensor still limit the maximum propagation distance in closed-loop systems.

Using the parameters for the Gemini telescope,[9] we deduce that the nonlinearities become very significant for effective propagation distances of $z > 200$ km and $z > 500$ km in the open-loop and closed-loop cases respectively. In actual fact, the system is operated at $z = 900$ km.[9] The reasons that this distance is greater in practice are first that the effect of the z^3 terms is reduced by diffraction. As explained in Subsection 4.B, diffraction attenuates the wave-front aberrations as the wave propagates, reducing the effect of nonlinear terms. Second, although the nonlinearities may be significant, this does not necessarily preclude successful operation of the loop.

C. Edge Signal

For completeness, we compute the second-order terms for the effect of a wave-front slope at the edge of the aperture. Although the signal in this region is dominated by diffraction, it is still instructive to consider the geometric optics behavior. To simplify the analysis, we perform the analysis at an edge of the aperture parallel to the y axis. The result is then converted to polar coordinates, which are more convenient for circular apertures.

At the edge of the aperture, Eq. (5) gives us the term

$$\int \mathrm{d}I = -\int \nabla I \cdot \nabla W \mathrm{d}z. \quad (24)$$

Let the edge of the aperture be located at $x = R$ and parallel to the y axis. The intensity at $z = 0$ is

$$I(x, y, 0) = I_0 - I_0 U(x - R), \quad (25)$$

where $U(x)$ denotes the Heaviside function, which takes a value of 1 for $x > 0$ and is zero valued elsewhere. Its derivative is $\delta(x)$, the Dirac delta function. Hence

$$I_x(x, y, 0) = -I_0 \delta(x - R), \quad (26)$$

and, assuming that the edge is displaced linearly as $zW_x(R, y, 0)$,

$$I_x(x, y, z) = -I_0 \delta(x - R - zW_x(R, y, 0)). \quad (27)$$

For compactness of notation, we now drop the argument of W and its derivatives when it is evaluated at $(R, y, 0)$. Evaluating to first order, we find that the slope term of the ITE gives

$$\int \mathrm{d}I = I_0 \int_0^z \delta(x - R - zW_x) W_x \mathrm{d}z$$
$$= I_0[1 - U(x - R - zW_x)]_0^z$$
$$= I_0[U(x - R) - U(x - R - zW_x)]. \quad (28)$$

Adding this quantity to Eq. (25) gives

$$I(x, y, z) = I_0 - I_0 U(x - R - zW_x). \quad (29)$$

We now evaluate ∇I as

$$\nabla I(x, y, z) = -I_0 \delta(x - R - zW_x)\hat{\mathbf{x}} + I_0 z W_{xy} \delta(x - R - zW_x)\hat{\mathbf{y}}. \quad (30)$$

The gradient of the wave front at the edge changes as the wave propagates, and we calculate this using the WTE. Since both ∇I and $\nabla^2 I$ are very large owing to the discontinuity at the edge of the aperture, we are not justified in neglecting the diffraction terms, which are significant, but we do so to observe the geometric optics behavior of the system. Differentiating the wavelength-independent terms of the WTE given in Eq. (6) with respect to x and y gives

$$\frac{\partial(\nabla W)}{\partial z} = \nabla W_z$$
$$= -\frac{1}{2}\nabla|\nabla W|^2$$
$$= -\frac{1}{2}\left[\frac{\partial(W_x^2 + W_y^2)}{\partial x}\hat{\mathbf{x}} + \frac{\partial(W_x^2 + W_y^2)}{\partial y}\hat{\mathbf{y}}\right]$$
$$= -(W_x W_{xx} + W_y W_{xy})\hat{\mathbf{x}} - (W_x W_{xy} + W_y W_{yy})\hat{\mathbf{y}}. \quad (31)$$

The wave-front slope at the edge of the propagated aperture, $W_x(R + zW_x, y, z)$, is found by taking the first-order Taylor series expansion around the point $(R, y, 0)$ in both x and z and the relationship $W_{xz} = -(W_x W_{xx} + W_y W_{xy})$, obtained by differentiating Eq. (6) with respect to x:

$$W_x(R + zW_x, y, z) = W_x + zW_x W_{xx} + zW_{xz} + O(z^2)$$
$$= W_x - zW_y W_{xy}. \quad (32)$$

Similarly, if we use $W_{yz} = -(W_x W_{xy} + W_y W_{yy})$, the gradient tangential to the edge is

$$W_y(R + zW_x, y, z) = W_y + zW_x W_{xy} + zW_{yz} + O(z^2)$$
$$= W_y - zW_y W_{yy}. \quad (33)$$

We now use Eqs. (30), (32), and (33) to evaluate

$$\int \mathrm{d}I = -\int_0^z \nabla I \cdot \nabla W \mathrm{d}z$$
$$= I_0 \int_0^z [W_x - zW_y W_{xy} - zW_{xy}(W_y - zW_y W_{yy})] \times \delta(x - R - zW_x)\mathrm{d}z. \quad (34)$$

Truncating the terms of z^2 or higher gives

$$\int \mathrm{d}I = I_0 \int_0^z (W_x - 2zW_y W_{xy})\delta(x - R - zW_x)\mathrm{d}z. \quad (35)$$

Finally, using the relationship

$$\int z\delta(x - az)\mathrm{d}z = -\frac{x}{a^2}U(x - az), \quad (36)$$

we obtain

$$\int \mathrm{d}I = I_0\left[1 - 2(x - R)\frac{W_y W_{xy}}{W_x^2}\right] \times [U(x - R) - U(x - R - zW_x)], \quad (37)$$

$$I(z) = I_0\left\{1 - U(x - R - zW_x) - 2(x - R)\frac{W_y W_{xy}}{W_x^2} \times [U(x - R) - U(x - R - zW_x)]\right\}. \quad (38)$$

Equation (38) states that the edge signal also depends on the tangential gradient of the wave front. The additional term is proportional to $W_y W_{xy}$ and is analogous to the T term when W_x, rather than $W_{xx} + W_{yy}$, gives rise to the signal.

If we let $x = \rho\cos(\theta)$ and $y = \rho\sin(\theta)$, this result can easily be extended to polar coordinates:

$$I(z) = I_0\left\{1 - U(\rho - R - zW_\rho) - \frac{2(\rho - R)}{R^2}\frac{W_0 W_{\rho 0}}{W_\rho^2} \times [U(\rho - R) - U(\rho - R - zW_\rho)]\right\}. \quad (39)$$

D. Comparison with Fresnel Diffraction by Simulation

The accuracy of Eq. (13) was verified by comparison with the intensity distribution obtained by Fresnel propagation. Figure 3 plots the intensity distribution obtained by using Fresnel diffraction for a 1-m linear aperture. The wavelength is 0.1 μm, the propagation distance is 10 km, and the aberration is $W = 100x^3\ \mu$m. Superimposed on this curve are the linear approximation, $I = I_0(1 - zH)$, and the nonlinear approximation of Eq. (13). It can be seen that away from the edges, the nonlinear curve closely resembles the Fresnel distribution

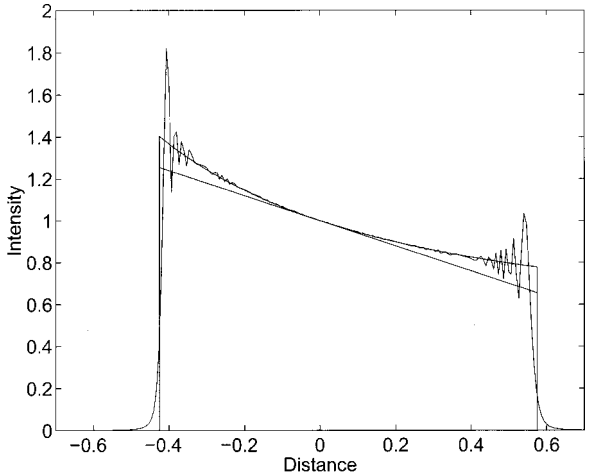

Fig. 3. Comparison of Fresnel (irregular curve), nonlinear (smooth curve) and linear (straight line) intensity curves for a cubic aberration. Here λ=0.1 μm, z = 10 km, and $W = 100x^3\ \mu$m.

and is much better than the linear approximation. Although the linear and the nonlinear intensity distributions differ significantly, the curvature-sensing signal corresponding to the linear and the nonlinear distributions differs by less than 1%. This reduction in sensitivity is directly attributable to the symmetric distribution of the detection planes around the focal plane in the curvature sensor.

4. DIFFRACTION

The irradiance transport equation suggests that the spacing of the curvature measurements is limited only by the size of the detector pixels. This is true only for an infinitesimally small propagation distance. As the wave propagates and the intensity begins to fluctuate, the wave spreads as a result of diffraction.[13]

A. Diffraction Due to the Finite Aperture Size

The main effects of the discontinuity are that the intensity changes more smoothly between 0 and I_0 at the edge, and ringing in the intensity is observed near the edge of the aperture (Fig. 3). These effects are not well modeled by the WTE because the first and second derivatives of the intensity are infinite at the discontinuity, but they can be computed instead by using the Fresnel diffraction integral.[13]

B. Diffraction Due to Aberrations

There are also diffraction effects resulting from the phase distortion of the wave front. The aberrated wave front gives rise to scintillation, which, in turn, changes the wave front. Conventionally, the diffraction has been defined in terms of a blur angle, λ/r_0.[6] However, this quantity was derived by considering a long-exposure image at the focal plane.[14] This is not the case in curvature sensing, which considers a short-exposure image in the Fresnel diffraction regime.

The blurring effect is not predicted by geometric optics and is a consequence of the diffraction of the wave front. We analyze this effect by using $I(z) = I - zI\nabla^2 W$, the first-order approximation to the ITE that describes the variation in I across the aperture:

$$\frac{\nabla^2 I}{I} = \frac{\nabla^2(I - zI\nabla^2 W)}{I} = -z(W_{xxxx} + 2W_{xxyy} + W_{yyyy}), \tag{40}$$

$$\left(\frac{|\nabla I|}{I}\right)^2 = z^2[(W_{xxx} + W_{xyy})^2 + (W_{xxy} + W_{yyy})^2]. \tag{41}$$

We now substitute the relationships of Eqs. (40) and (41) into the wavelength-dependent component of the WTE, Eq. (4), to obtain an expression that describes how diffraction affects the wave front. The diffraction term with the lowest order in z consists of

$$W_z^{\text{diff}} = \frac{-z\lambda^2}{16\pi^2}(W_{xxxx} + 2W_{xxyy} + W_{yyyy}) \tag{42}$$

and changes the wave front according to

$$W(z) = W(0) - \frac{z^2\lambda^2}{32\pi^2} \times (W_{xxxx} + 2W_{xxyy} + W_{yyyy}) + O(z^3). \tag{43}$$

Referring to Eq. (5), this change in the wave front is reflected in the intensity according to

$$\ln(I(z)/I_0) = \int_0^z \nabla^2\left[\frac{z^2\lambda^2}{32\pi^2}(W_{xxxx} + 2W_{xxyy} + W_{yyyy})\right]dz = \frac{z^3\lambda^2}{96\pi^2}(W_{xxxxxx} + 3W_{xxxxyy} + 3W_{xxyyyy} + W_{yyyyyy}) + O(z^4). \tag{44}$$

Denoting

$$D = \frac{\lambda^2}{96\pi^2}(W_{xxxxxx} + 3W_{xxxxyy} + 3W_{xxyyyy} + W_{yyyyyy}), \tag{45}$$

we can write

$$S = \frac{-zH + z^3 D + O(z^3)}{1 + z^2(K - T)}. \tag{46}$$

We assume that the diffraction blurring is a significant limitation on the curvature sensor when it cancels half the signal, i.e., $z^3 D = zH/2$. To show how this condition defines a spatial resolution, consider a one-dimensional sine-wave aberration,

$$W(x) = \sin\left(\frac{2\pi}{a}x\right), \tag{47}$$

where a is the wavelength of the aberration. The condition states that

$$z^3\frac{\lambda^2}{96\pi^2}\left(\frac{2\pi}{a}\right)^6 = \frac{1}{2}z\left(\frac{2\pi}{a}\right)^2. \tag{48}$$

Rearranging, we obtain the variation

$$a = 3^{-1/4}\sqrt{\pi\lambda z} \approx \sqrt{\lambda z}, \tag{49}$$

which confirms that it is the Fresnel length, $\sqrt{\lambda z}$, that determines the resolution of the intensity fluctuations.[13] Its significance is that any wave-front perturbations of scale smaller than the Fresnel length are blurred and invisible to the detector. It is interesting to note that the resolution is proportional to $\sqrt{z}$ and not z, implying that the diffraction effects due to the atmosphere are less severe than previously thought. It is apparent that any wave-front perturbation of scale $\sqrt{\lambda z}$ or smaller cannot be measured and compensated, and increasing the propagation distance increases the scale over which the sensor is blind. This is more accurate than a previous analysis that assumes that in closed-loop operation, the blur angle is that imposed by diffraction, λ/D, where D is the diameter of the aperture.[15]

5. SIGNAL-TO-NOISE RATIO

We assume that the intensities I_1 and I_2 are normally distributed with the mean and the variance equal to $\overline{I_1}$ and

$\overline{I_2}$, respectively. This is approximately true if the photon arrival is a Poisson process and the average number of photons is large. We denote the noise on I_1 and I_2 as n_1 and n_2, respectively, and assume that it is uncorrelated. Using the approximation that the mean of a ratio is approximately equal to the ratio of the means, which holds if the signal is much greater than the noise, the expected value of S, $\bar{S}$, is

$$\bar{S} = E\left[\frac{\overline{I_1} + n_1 - \overline{I_2} - n_2}{\overline{I_1} + n_1 + \overline{I_2} + n_2}\right] \approx \frac{\overline{I_1} - \overline{I_2}}{\overline{I_1} + \overline{I_2}}, \tag{50}$$

where S is the conventional curvature signal of Eq. (14). Using this approximation to compute the variance of S gives

$$\begin{aligned}
\mathrm{Var}(S) &= E\left[\left(\frac{I_1 - I_2}{I_1 + I_2} - \frac{\overline{I_1} - \overline{I_2}}{\overline{I_1} + \overline{I_2}}\right)^2\right] \\
&= E\left[\left(\frac{2(I_1\overline{I_2} - \overline{I_1}I_2)}{(I_1 + I_2)(\overline{I_1} + \overline{I_2})}\right)^2\right] \\
&= 4E\left[\frac{n_1^2\overline{I_2}^2 + 2n_1n_2\overline{I_1}\,\overline{I_2} + n_2^2\overline{I_1}^2}{(\overline{I_1} + n_1 + \overline{I_2} + n_2)^2(\overline{I_1} + \overline{I_2})^2}\right] \\
&\approx 4\frac{\overline{I_1}^2\overline{I_2} + \overline{I_2}^2\overline{I_1}}{(\overline{I_1} + \overline{I_2})^2(\overline{I_1} + \overline{I_2})^2} \\
&= 4\frac{\overline{I_1}\,\overline{I_2}}{(\overline{I_1} + \overline{I_2})^3} = \frac{1 - \bar{S}^2}{N}.
\end{aligned} \tag{51}$$

$N = \overline{I_1} + \overline{I_2}$ is the total expected number of photons and is weakly dependent on the signal owing to the nonlinearities in the denominator of Eq. (14). Using the linear relationship $S = -z\nabla^2 W$, we obtain the variance of $\nabla^2 W$,

$$\mathrm{Var}(\nabla^2 W) = z^{-2}\ \mathrm{Var}(S). \tag{52}$$

If we know the expected value of N *a priori* and use the signal defined in Eq. (16), then it is easy to show that the variance of $S_\star$ is

$$\mathrm{Var}(S_\star) = E\left[\left(\frac{I_1 - I_2}{\overline{I_1} + \overline{I_2}} - \frac{\overline{I_1} - \overline{I_2}}{\overline{I_1} + \overline{I_2}}\right)^2\right] = \frac{1}{N}. \tag{53}$$

Equation (53) is usually quoted to be the variance of S but is less accurate than the expression of Eq. (51) when the conventional signal is used. The difference is a decrease in variance proportional to the square of the expected signal.

Rigaut *et al.* state that at low photon counts, when the statistics of photon arrival cannot be assumed to be Gaussian, it is better to use Eq. (16).[9] We simulated the error in estimating S as a function of S and N. The arrival of photons was Poisson distributed with means $\overline{I_1} = N(1 + S)/2$ and $\overline{I_2} = N(1 - S)/2$. The variance of the error, $E[(S - \bar{S})^2]$, is plotted in Figs. 4 and 5. It can be seen that the curves exhibit the behavior predicted by Eqs. (51) and (53) for $N = 1000$. When there are only ten photons, the error is higher than that implied by Eq. (51). The results demonstrate that it is better to use the conventional signal of Eq. (14) except when the signal is weak and the photon count is very low. This is despite the signal-dependent error introduced by substituting $2I_0$ for $I_1 + I_2$, given in Eq. (16).

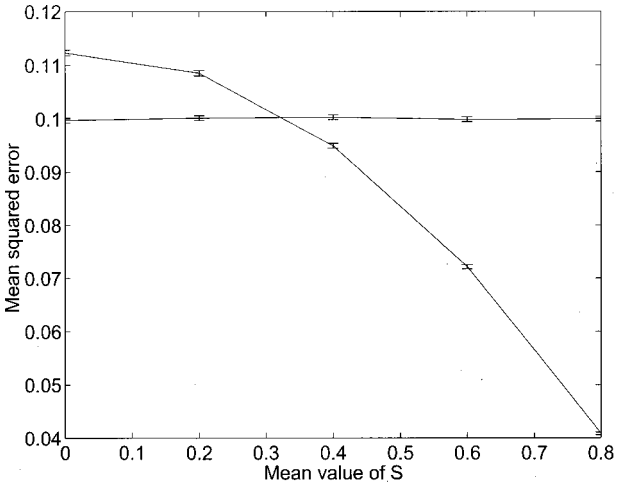

Fig. 4. Variance of the error in the estimate of S (parabola) and $S_\star$ (straight line) for ten photons.

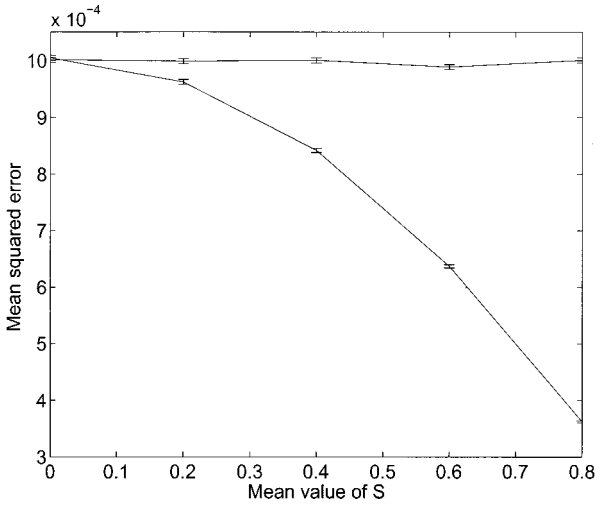

Fig. 5. Variance of the error in the estimate of S (parabola) and $S_\star$ (straight line) for 1000 photons.

6. CONCLUSION

As the wave front propagates, the intensity and the wave front change. This leads to a nonlinear relationship between the intensity measurement and the curvature at the aperture. These nonlinear effects become more severe as the propagation distance increases and consequently impose a limit on the propagation distance. In addition, diffraction effects limit the spatial resolution of the measurements to $\sqrt{\lambda z}$; any smaller-scale aberrations are difficult to detect. We also find that the variance of the signal is equal to $(1 - S^2)/N$. Replacing the denominator of the curvature signal by the average number of photons can reduce the effects of the photon noise at a low photon-noise and signal level but incurs a bigger error owing to additional nonlinearities.

ACKNOWLEDGMENTS

The authors thank the New Zealand government for financial aid in the form of a Bright Futures scholarship and a Marsden Fund research grant. We are also grateful to the Observatoire de Lyon, France, for their support of this research.

R. G. Lane may be reached at R.Lane @elec.canterbury.ac.nz.

REFERENCES

1. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford, New York, 1998), pp. 165–168, 377–394.
2. M. Milman, D. Redding, and L. Needels, "Analysis of curvature sensing for large-aperture adaptive optics systems," J. Opt. Soc. Am. A **13**, 1226–1238 (1996).
3. D. C. Johnston, B. L. Ellerbroek, and S. M. Pompea, "Curvature sensing analysis," in *Adaptive Optics in Astronomy*, M. A. Ealey and F. Merkle, eds., Proc. SPIE **2201**, 528–538 (1994).
4. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am. **73**, 1434–1441 (1983).
5. N. Streibl, "Phase imaging by the transport equation of intensity," Opt. Commun. **49**, 6–10 (1984).
6. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," Appl. Opt. **27**, 1223–1225 (1988).
7. G. Rousset, "Wave-front sensors," in *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, Cambridge, UK, 1999), pp. 91–130.
8. F. Roddier, "Wavefront sensing and the irradiance transport equation," Appl. Opt. **29**, 1402–1403 (1990).
9. F. Rigaut, B. L. Ellerbroek, and M. J. Northcott, "Comparison of curvature-based and Shack–Hartmann-based adaptive optics for the Gemini telescope," Appl. Opt. **36**, 2856–2868 (1997).
10. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
11. M. C. Roggemann and B. Welsh, *Imaging through Turbulence* (CRC, Boca Raton, Fla., 1996), p. 98.
12. M. A. van Dam and R. G. Lane, "Tip/tilt estimation from defocused images," J. Opt. Soc. Am. A **19**, 745–752 (2002).
13. J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996), pp. 63–75.
14. D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," J. Opt. Soc. Am. **56**, 1372–1379 (1966).
15. F. Roddier, "Error propagation in a closed-loop adaptive optics system: a comparison between Shack–Hartmann and curvature wave-front sensors," Opt. Commun. **113**, 357–359 (1995).