Noise propagation in wave-front sensing with phase diversity

Ludovic Meynadier, Vincent Michau, Marie-Thérèse Velluet, Jean-Marc Conan, Laurent M. Mugnier, and Gérard Rousset

The phase diversity technique is studied as a wave-front sensor to be implemented with widely extended sources. The wave-front phase expanded on the Zernike polynomials is estimated from a pair of images (in focus and out of focus) by use of a maximum-likelihood approach. The propagation of the photon noise in the images on the estimated phase is derived from a theoretical analysis. The covariance matrix of the phase estimator is calculated, and the optimal distance between the observation planes that minimizes the noise propagation is determined. The phase error is inversely proportional to the number of photons in the images. The noise variance on the Zernike polynomials increases with the order of the polynomial. These results are confirmed with both numerical and experimental validations. The influence of the spectral bandwidth on the phase estimator is also studied with simulations. © 1999 Optical Society of America

OCIS codes: 010.1080, 010.7350, 100.5070, 100.3190, 100.3020, 120.5050.

1. Introduction

The wave-front sensor is a key component of adaptive (or active) optics systems. For these applications many wave-front sensing techniques have been developed and characterized.[1] However, only a few techniques can be used with widely extended sources.[2–5] Among them, phase diversity, which requires only two focal-plane images, presents some interesting characteristics. The optical setup is simple and can be part of the imaging camera. This sensor does not require any calibration, unlike Shack–Hartmann-type sensors. However, this sensor does not lead to a direct measurement of the aberrations. It requires iterative data reduction methods to estimate the phase, unlike the wave-front sensors based on a geometrical optics approximation, which provide a noniterative estimation of the phase, since the signal is the gradient or the Laplacian of the phase.

The phase diversity technique was first proposed by Gonsalves[6] to improve the quality of the images degraded by aberrations and was then applied by many authors,[7–10] particularly to solar imaging through turbulence.[11–13] Simultaneously with the derivation of the restored image, the aberrations of the optical system can also be derived as a byproduct. The phase diversity technique was, for example, successfully applied to the determination of the Hubble Space Telescope aberrations.[14–16] Some studies on the performance evaluation of phase diversity have been published,[17–19] but a modal quantitative evaluation of the performance of phase diversity, as a wave-front sensor, has not, to our knowledge, been performed. This is our objective in this paper.

A data reduction method was implemented to derive the wave front from the phase diversity data. The wave-front measurement quality was studied theoretically. The technique was tested both on numerically simulated data and on experimentally recorded images. The principle of phase diversity is presented in Section 2. In Section 3 the data reduction method and the corresponding wave-front estimation algorithm are described. When faint sources are used, the main source of wave-front error is the noise in the images. Section 4 is dedicated to the theoretical study of the noise propagation on the phase estimation. The performance evaluation of the wave-front estimation algorithm is obtained by numerical simulations in Section 5. Section 6 con-

When this research was performed, the authors were with the Office Nationale d'Études et de Recherches Aérospatiales, B.P. 72, 92322 Châtillon CEDEX, France. L. Meynadier is now with the Groupement de Recherche, Ecole Nationale Supérieure des Télécommunications, 46 rue Barraut, 75634 Paris CEDEX 13, France. J.-M. Conan's e-mail address is conan@onera.fr.

Received 11 September 1998; revised manuscript received 22 March 1999.
0003-6935/99/234967-13$15.00/0

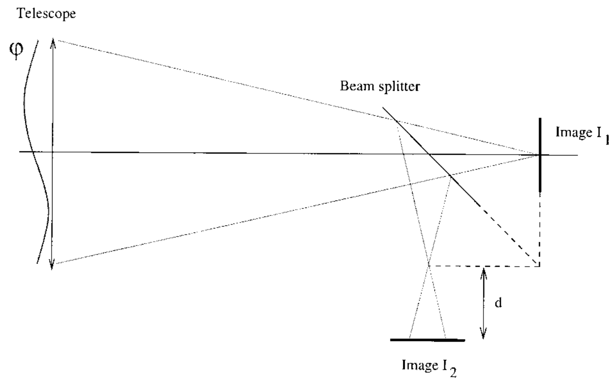


Fig. 1. Phase diversity principle.

tains the experimental validation of the behavior of the phase diversity.

2. Principle

The phase diversity principle[6] is based on the simultaneous recording of two or more quasi-monochromatic images. In the following we consider the use of only two images. The first image is recorded in the focal plane of the optical system. The second image, called the diverse image, is recorded in an out-of-focus plane. The distance between these two planes is calibrated and corresponds to a small defocus. With extended sources, the use of the additional image is required so that the solution is more likely to be unique.[6,20–22]

An implementation of the phase diversity is illustrated in Fig. 1. A beam splitter and two detector arrays placed near the focus of the telescope are used to record simultaneously the focal and the out-of-focus images.

Assuming that the light is spatially incoherent, the two recorded images $I_k (k = 1, 2)$ can be expressed as functions of the aberrated phase in the optical system pupil φ and of the intensity distribution of the source O:

$$I_k(\mathbf{r}) = O(\mathbf{r}) * S_k(\mathbf{r}), \tag{1}$$

where $*$ denotes the convolution product, S_k is the point-spread function (PSF) in the observation plane number k, and $\mathbf{r}$ is a two-dimensional vector in the image plane. For a monochromatic wave, S_k is expressed by

$$S_k(\mathbf{r}) = \left| \int_{-\infty}^{+\infty} A_k(\mathbf{x}) \exp\left(i \frac{2\pi}{\lambda F} \mathbf{r} \cdot \mathbf{x} \right) \mathrm{d}\mathbf{x} \right|^2,$$

with

$$A_1(\mathbf{x}) = \alpha(\mathbf{x}) \exp[i\varphi(\mathbf{x})],$$

$$A_2(\mathbf{x}) = A_1(\mathbf{x}) \exp[i\phi_d(\mathbf{x})],$$

$$\varphi(\mathbf{x}) = \frac{2\pi\Delta(\mathbf{x})}{\lambda}, \tag{2}$$

where $\Delta(\mathbf{x})$ is the optical path, assumed to be independent of the wavelength λ; A_1 is the complex amplitude in the pupil plane; F is the focal length; ϕ_d is the defocus phase for image I_2; $\mathbf{x}$ is a two-dimensional vector in the pupil plane; and α is the characteristic function of the pupil (1 inside, 0 outside for a binary pupil). Indeed, the intensity fluctuations in the pupil plane are neglected.

From Eqs. (1) and (2) it is clear that the relationship between the recorded images and the aberrated wave front is not linear. Furthermore, there is no analytical solution that gives the wave front from an expression that combines the two images. Similar to phase diversity applied to image restoration, we chose an iterative estimation by minimizing an error metric.

3. Maximum-Likelihood Estimation

The error metric is derived from a stochastic approach. The noise in the images is the sum of the photon noise (Poisson-distributed random variable) and the Gaussian CCD readout noise. For a bright and extended object, stationary white Gaussian noise, with a uniform variance equal to the mean number of photons/pixel, is a first approximation of photon noise. With this assumption the joint maximum-likelihood (ML) estimate of the wave front φ and of the object O is jointly determined by minimization of the following criterion:

$$E = \sum_{k,i} |I_k(\mathbf{r}_i) - O(\mathbf{r}_i)*S_k(\mathbf{r}_i)|^2. \tag{3}$$

The spatial sampling of the images ($\mathbf{r}_i$) is determined by that of the detector array. The object is estimated with the same sampling. In fact, for a low photon count or an object that does not cover the whole field of view (FOV), the approximation of stationary Gaussian noise is no longer valid and the estimator is no longer a true ML estimator but rather a least-squares estimator. Of course, it is possible to use the likelihood of the true photon noise.[8] In any case, even if the least-squares estimator is suboptimal, it still provides well-restored phases, as shown in Sections 5 and 6.

To take advantage of the discrete Fourier transforms (DFT's) in the implementation of the previous criterion, we treat the object, the PSF's, and the images as periodic arrays with a periodic cell size of $N \times N$. The criterion becomes

$$E \propto \sum_{k=1}^{2} \sum_{i=1}^{N^2} |\tilde{I}_k(\mathbf{f}_i) - \tilde{O}(\mathbf{f}_i)\tilde{S}_k(\mathbf{f}_i)|^2, \tag{4}$$

where $\tilde{X}$ is the DFT of X and $\mathbf{f}_i$ is a two-dimensional vector in the discrete spatial-frequency space.

For monochromatic simulations we consider that the images are sampled at the Shannon rate, i.e., 2 pixels per λ/D, where λ is the wavelength and D is the telescope diameter.

The estimated phase is described by use of its expansion on the Zernike polynomials.[23] Only a limited number of Zernike coefficients a_l are estimated, to as great as $l = M$. The three first coefficients a_{1-3} are not determined. The first coefficient, the piston coefficient, is a constant added to the phase and has no influence on the PSF. The others, the tilt coefficients, are not estimated, since they introduce a shift only in the image that is of no importance for widely extended sources. $(M - 3)$ Zernike coefficients are therefore estimated.

To avoid edge effects in the case of widely extended sources, the convolution in Eq. (3) is performed with an object support that is extended by a guard-band as used by Seldin.[12] In our case we chose a guard-band width equal to $N/2$. If the support of the PSF is small, the guard-band width can be reduced.

To minimize the error metric [Eq. (4)], the gradient-conjugate method[24,25] was chosen. Through this minimization we jointly estimate the sampled object O and the $(M - 3)$ Zernike coefficients of the phase φ, applying a strict positivity constraint on the sampled object, thanks to a reparametrization[26,27] (see also Appendix A). The gradients of the error metric with respect to the object and the phase estimates are presented in Appendix A.

4. Theoretical Study of the Noise Propagation

A. Analytical Approach

Fessler[28] proposed a formalism to study the propagation of the measurement noise on the set of estimated parameters $\{p_e\}$ at convergence, in the case of the resolution of an inverse problem by a ML approach. $\{p_e\}$ are the parameters estimated by minimization of the error metric E, which is a function of both the parameters $\{p\}$ and the measurements $\{m\}$:

$$\{p_e\} = \arg\min_{\{p\}} E(\{p\}, \{m\}). \tag{5}$$

The parameters to be estimated are the Zernike coefficients a_4–a_M and the object. The measurements are the pixel intensities in each image.

By use of the second-order Taylor expansion of E, the covariance matrix of the estimated parameters [Cov $\{p_e\}$] reads as[28]

$$[\mathrm{Cov}\{p_e\}] \approx [\nabla^2_{p,p}E]^{-1}[\nabla^2_{m,p}\, E] \times [\mathrm{Cov}\{m\}]^t[\nabla^2_{m,p}\, E]^t[\nabla^2_{p,p}\, E]^{-1}, \tag{6}$$

where [Cov $\{m\}$] is the covariance matrix of the measurement noise; $[\nabla^2_{p,p}E]$ is the second partial derivative matrix of the error metric with respect to parameters, also called the Hessian matrix of the error; $[\nabla^2_{m,p}\, E]$ is the second partial derivative matrix of the error with respect to m and p; and superscript t denotes transposition of the matrix that follows. To derive [Cov $\{p_e\}$], the partial derivatives in the matrices of relation (6) are computed at a specific point that corresponds to the mean measurements (i.e., without noise) and to the associated estimated parameters.

In our case (see Section 3) we have no spatial correlation of the noise; the covariance matrix [Cov $\{m\}$] is therefore diagonal. In addition, for widely extended sources, the fluctuation of the intensity in the object is small compared with the mean intensity level. So the variance of the noise, whether photon or detector noise, can be assumed to be constant and will be expressed in photoelectrons/pixel. Consequently, the covariance matrix of the noise in the images is proportional to the identity matrix. Finally, to make the computation of relation (6) tractable, we assume that the object is known and we do not use a guard band. This assumption may seem like an oversimplification; however, it is justified *a posteriori* by the fact that the theoretically estimated modal variances are found to be in good agreement with the simulations presented in Section 5. We study the noise propagation only on the estimated Zernike coefficients. Relation (6) becomes[29]

$$[\mathrm{Cov}\{a_{l,e}\}] \approx \frac{N_{\mathrm{ph}}}{N^2}[\nabla^2_{\{a_{l'}\},\{a_l\}}\, E]^{-1}[\nabla^2_{\{I_k\},\{a_l\}}\, E] \times {}^t[\nabla^2_{\{I_k\},\{a_l\}}E]^t[\nabla^2_{\{a_{l'}\},\{a_l\}}\, E]^{-1}, \tag{7}$$

where N_{ph} is the number of photons per image and $a_{l,e}$ is the estimate of the coefficient a_l.

The expressions of the two partial derivative matrices of the error are given in Appendix B. We have demonstrated that the product of $[\nabla^2_{\{Ik\},\{al\}}\, E]$ with its transpose matrix is proportional to $[\nabla^2_{\{al'\},\{al\}}\, E]$. Therefore the covariance matrix of the noise for the estimated Zernike coefficients is proportional to the inverse of the Hessian matrix of the error metric and to the noise variance in the images (N_{ph}):

$$[\mathrm{Cov}\ \{a_{l,e}\}] \approx 2\,\frac{N_{\mathrm{ph}}}{N^4}[\nabla^2_{a_{l'},a_l}\, E]^{-1}. \tag{8}$$

The expression of the Hessian matrix depends on the object power spectrum and on the complex amplitudes in the pupil for each image [see Eq. (B3)].

When the complex amplitude in the pupil tends toward the amplitude modulus distribution in the pupil (zero-order expansion of the complex amplitude), it can be shown that the covariance matrix of the noise for estimated Zernike coefficients is inversely proportional to the number of photons/image:

$$[\mathrm{Cov}\{a_{l,e}\}] \propto \frac{1}{KN_{\mathrm{ph}}}[\mathcal{M}], \tag{9}$$

where K ($K = 2$ generally) is the number of observation planes and $[\mathcal{M}]$ is a matrix independent of N_{ph}. $[\mathcal{M}]$ is a function only of the nature of the object and of experimental parameters.

Relation (9) demonstrates that the noise variance on the Zernike coefficients is inversely proportional to the total number of photons KN_{ph}, collected by the telescope in the image FOV. This behavior is similar to that of other wave-front sensors.[1]

Writing $N_{\mathrm{ph}} = N^2 n_{\mathrm{ph}}$, where n_{ph} is the average number of photons/pixel, shows that it is interesting to increase N^2, which means increase the FOV, for a given widely extended source and a given n_{ph}. This behavior is also observed in the wave-front error ex-

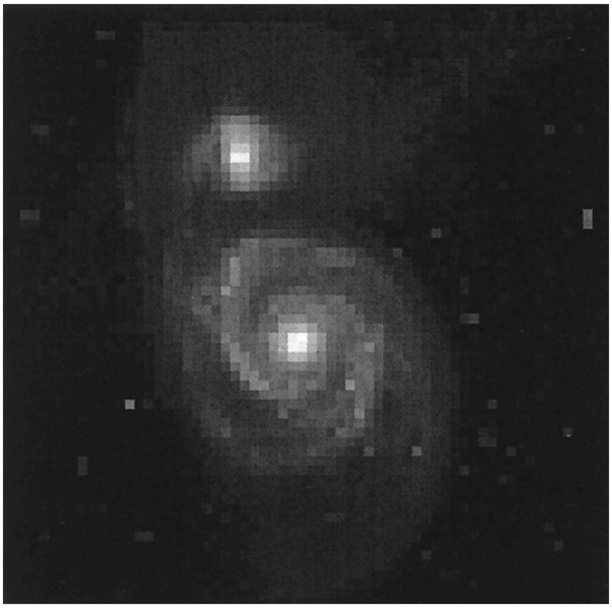


Fig. 2. Spiral galaxy object.

Table 1. Values in Radians of the Coefficients Used for the Simulation

Coefficient	Value (rad)
a_4	−0.2
a_5	0.3
a_6	−0.45
a_7	0.4
a_8	0.3
a_9	−0.25
a_{10}	0.35
a_{11}	0.2
a_{12}	0.1
a_{13}	0.05
a_{14}	−0.05
a_{15}	0.05
a_{16}	0.02
a_{17}	0.01
a_{18}	−0.01
a_{19}	−0.02
a_{20}	0.01
a_{21}	0.01

pression obtained with a Shack–Hartmann wave-front sensor used with a widely extended object.[3]

To take advantage of this gain brought on by the increase of the FOV, one must however, remain in the anisoplanatic FOV. This condition is easily verified for a space active telescope or in ophthalmology,[30] for example.

Relation (9) is an approximation. The true expression depends on the actual phase. However, it is easy to demonstrate, with a higher-order expansion of the complex amplitude, that the matrix [Cov $\{a_{l,e}\}$] higher-order terms depend only on the second-order (or more) phase terms. Consequently, the noise propagation is almost independent of the phase amplitude for small aberrations, which is the case in closed-loop adaptive (or active) optics.

B. Numerical Application

The full expression of the covariance matrix is complicated [see Eq. (B3)], and its analytical evaluation is difficult. Instead, we compute it numerically on a particular example, using relations (8) and (B3). The object is a spiral galaxy sampled by 64 × 64 pixels (see Fig. 2). This object, of limited extent, does not require a guard band. This is the only case in this paper in which a guard band is not used. The distorted phase is described by the first 21 Zernike polynomials and has a spatial standard deviation corresponding to $\lambda/7$ rms. The values of the coefficients are listed in Table 1, and a perspective view of the wave front is shown in Fig. 3. The image obtained in the focal plane is shown in Fig. 4: Fig. 4(a) represents the noiseless diffraction limited image, and Fig. 4(b) shows the aberrated image with a total photon number of 10^7. We perform the phase estimation with $M = 21$ or 36 Zernike coefficients.

After numerical computation of relation (8) we notice first that the covariance matrix [Cov $\{a_{l,e}\}$] is almost diagonal. Figure 5 presents the noise variance of each estimated Zernike coefficient (i.e., the diagonal of the covariance matrix) for $N_{\mathrm{ph}} = 10^7$ photons/image. Note that the variances are relatively low ($\sim 10^{-3}$ rad^2, considering 10^7 photons/image). For a given number of estimated polynomials M, the noise variance increases with the polynomial azimuthal and radial degrees.[23] This behavior is specific to phase diversity. It is different from the noise propagation in the other wave-front sensors. With Shack–Hartmann or curvature wave-front sensors the noise variance decreases quickly with the radial degree of the Zernike polynomials.[1] Moreover, we can see in Fig. 5 that the noise variance for a given polynomial increases slightly with M, the number of estimated coefficients. In addition, the total variance ($\Sigma_i a_i^2$) increases significantly with M. Limiting the number of reconstructed Zernike coefficients is indeed equivalent to an implicit regularization that limits the noise amplification but generates a bias.[31] Hence there is a higher noise level for a larger number of estimated Zernike coefficients.

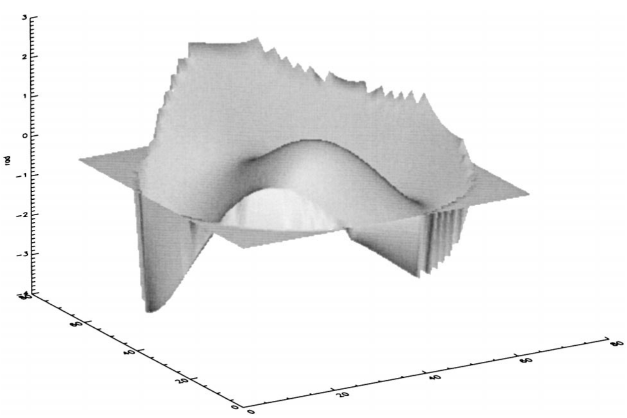
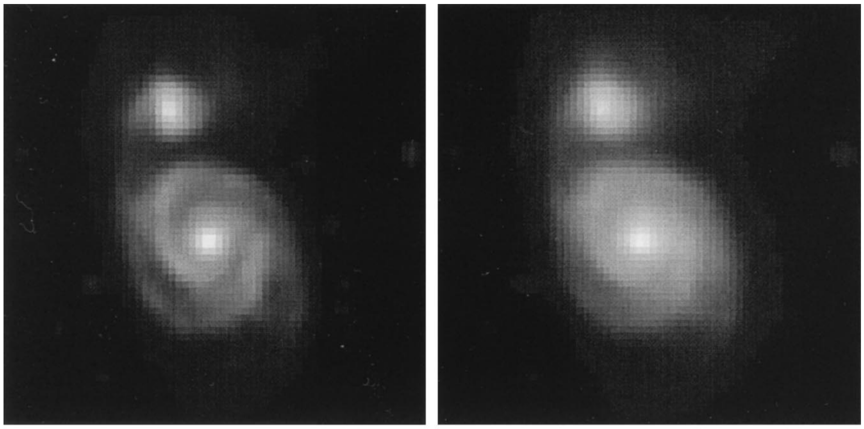


Fig. 3. Wave front used for the numerical approach of the noise propagation theory.



(a) (b)

Fig. 4. Images of the spiral galaxy in the focal plane. (a) Noiseless diffraction-limited image. (b) Aberrated image with 10^7 total photon number.

The phase diversity is, consequently, better adapted for the estimation of the low-order aberrations. This property is particularly interesting for active optic systems, i.e., the compensation for the telescope aberrations, which do not need a large number of Zernike coefficients to be estimated, but requires a great accuracy. The use of the phase diversity for estimation of high-order modes (i.e., in adaptive optics system) requires a better regularization technique than the one used in this paper. For example, *a priori* knowledge of the phase, such as the statistics of the turbulent phase, could be easily incorporated into the error metric[32,33] to regularize the problem and to preserve a good accuracy even on the high orders.

The influence of the distance between the two observation planes on the noise propagation must also be studied. Figure 6 presents the total noise variance of the first 21 estimated Zernike coefficients as a function of the amount of defocus between the two observation planes. For the first time to our knowledge we demonstrate the existence of an optimal defocus that minimizes the noise propagation on the estimated phase. The minimum is approximately equal to 2.6π rad of defocus wave-front amplitude for this particular case. The corresponding focus distance depends on the optical system focal ratio. When the defocus amplitude decreases, the difference between the focal image and the diverse one is no longer sufficient to allow for a good convergence of the phase diversity algorithm. However, when the defocus amplitude is too large, the contrast in the out-of-focus image is attenuated and this image is no longer usable.

5. Performance Estimation by Numerical Simulation

In Section 4 we developed the analytical expression of the covariance matrix of the estimated parameters. We showed that, for a small coefficient amplitude, the expression depends only on the detected flux level, on the image size, and on the number of estimated Zernike polynomials. To confirm this dependence without any assumption, below we present results

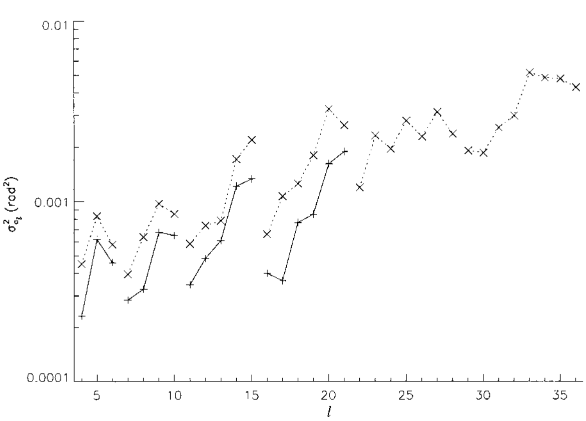


Fig. 5. Noise variance (in radians squared) of the first 21 (solid curve) and 36 (dotted curve) estimated Zernike coefficients for 10^7 photons and 64×64 pixels/image.

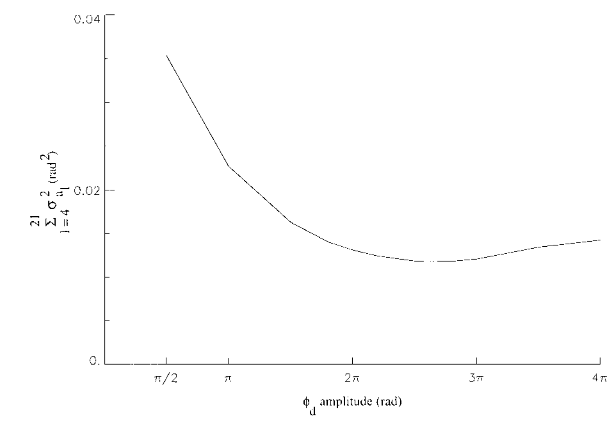


Fig. 6. The total noise variance (in radians squared) of the first 21 estimated Zernike coefficients versus the defocus wave-front amplitude for 10^7 photons and 64×64 pixels/image.

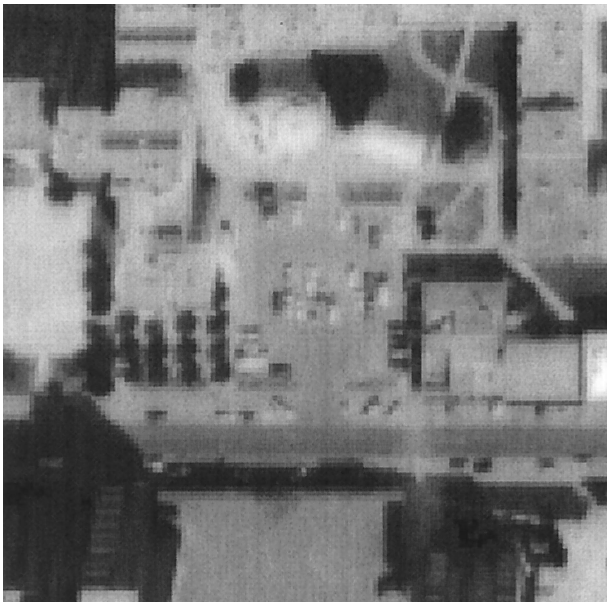


Fig. 7. Urban scene used for computing the images in the simulations.

obtained by numerical simulations, using the ML estimation presented in Section 3.

A. Image Simulation

We performed a simulation in which the diverse number K of observation planes is equal to 2. The defocus amplitude for the second observation plane is 2π rad, as proposed by many authors.[7,10,12,13,34,35] The monochromatic images are sampled according to the Shannon criterion.

The simulation of the images used to study wavefront sensing with phase diversity presents several steps: the phase and the PSF simulations, the noiseless image simulation, and the noise simulation.

The phase in the pupil, which is assumed to be an unobstructed disk, is computed as a linear combination of the first Zernike polynomials.[23] The amplitude of the distorted phase is limited to 2π rad. In practice, when this condition is met, the gradient-based minimization algorithm used does not fall within local minima.

The PSF in the focal plane is deduced from the phase in the pupil [see Eq. (2)]. To satisfy the Shannon criterion, the complex amplitude in the pupil, which is of size 64×64 pixels, is zero padded to a 128×128 pixel array then Fourier transformed and squared to obtain a 128×128 pixel PSF array. The PSF in the second plane is computed similarly by addition of ϕ_d to the phase in the pupil.

To obtain the image pair, a discrete convolution product is performed between an extended 128×128 pixel object, such as the Earth viewed from a satellite (see Fig. 7), and the PSF corresponding to each observation plane. The control area (64×64 or 32×32 pixels) is then extracted to obtain the actual images used in the deconvolution.

The phase diversity method is sensitive to image formation chromaticity [see Eq. (2)]. We simulate polychromatic images by adding monochromatic PSF's obtained at different wavelengths. To simulate these images, we assume that the Shannon criterion is satisfied at the shortest wavelength of the spectral band. Consequently, the other PSF's are oversampled. This is achieved by reduction of the size of the pupil to smaller than 64×64 pixels. Note that the complex amplitude has to be computed with a phase scaled according to the wavelength [see Eq. (2)]. Because the size of the pupil can be reduced only by an integer number of pixels, the number of possible wavelengths is limited and depends on the bandwidth (see Subsection 5.C). The source is assumed to have a flat chromatic spectrum.

Furthermore, noise can be added to the images. The whole image is normalized by the given average number of photons N_{ph}. Then the intensity distribution of the noisy image is determined from a rejection method[25] where we use the noiseless image and assume that the noise in each pixel is statistically independent and follows a Poisson distribution.

B. Noise Propagation and Bias

The distorted phase is simulated with $L = 21$ Zernike polynomials, and we take the same value for each Zernike coefficient. The estimated phase is expanded on the same polynomials to avoid aliasing effects, i.e., $M = L$. In these simulations the object is considered to be unknown and is also estimated by the algorithm.

For each studied case the noise influence on the estimated phase is obtained when we process n realizations of noisy monochromatic image pairs ($n = 50$). Therefore, for each estimated Zernike coefficient a_l, its mean value is expressed by $\langle a_l \rangle_n$ (in radians), where $\langle\ \rangle_n$ is the average on the n realizations. The variance of the Zernike coefficient is determined by ${\sigma_{a_l}}^2 = \langle {a_l}^2 \rangle_n - {\langle a_l \rangle_n}^2$ (in radians squared). For a given case of phase estimation we compute the total mean-squared error by the sum of the variance $\sigma_{a_l}^2$ and of the square of the bias. The bias is defined by $\langle a_l \rangle_n - a_{l,\mathrm{true}}$.

The standard deviation σ_{a_l} and the bias of the estimated coefficients are determined for three different sets of low-amplitude aberrations (a_4–a_L all equal to 0, 0.037, or 0.05 rad, respectively) with 32×32 pixels in the images and a total flux level of 10^7 photons/image. The different images observed in the focal plane are reproduced in Fig. 8. We observe (see Fig. 9) that the variance of the estimated Zernike coefficients is independent of the values of these low-amplitude aberrations. In addition, there is no evidence of a bias when we consider the error bars (see Fig. 10). Therefore the standard deviation dominates the bias (see Fig. 9), and the mean-squared error is almost equal to the noise variance.

An additional study was performed with a bigger part of the noisy blurred scene (64×64 pixel images). Flux levels range from 10^6 to 10^9 photons/image, and the amplitude of the coefficients a_4–a_{21} is equal to 0.05 rad. Figures 11(a) and 11(b) show images with

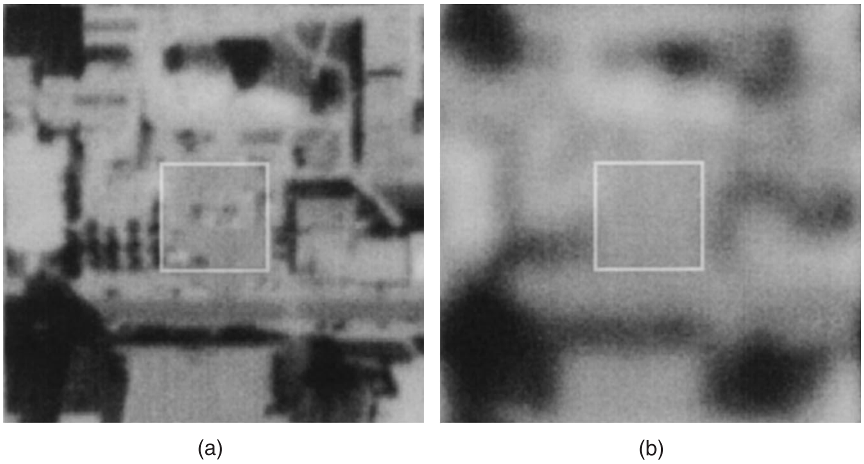


Fig. 8. Full-extent noisy blurred scenes in the focal plane with a total of 10^7 photons and with the amplitude of coefficients a_4–a_{21} equal to 0.05 rad (a) in the focal plane, (b) in the out-of-focus plane. The other aberrated images are similar, because they are of low amplitude.

10^6 and 10^9, respectively, per image. In Fig. 12 we can observe that the variance clearly decreases with the flux level and that it tends to increase with the Zernike polynomial number, even if this trend was more obvious on the theoretical variances (see Fig. 5). The decrease of the variance is inversely proportional to the number of photons in the image, as shown in Fig. 13. The results of the theoretical study are therefore confirmed.

Figure 14 represents the bias in the estimation of the Zernike coefficients for the different flux levels. As we saw in Fig. 10, the estimated value is not biased, except for low flux level. With 10^6 photons/image, the bias becomes too large for active optics applications. Indeed, the performance requirements on the standard deviation of the estimated wave front are ~0.1 rad, and some estimated terms are far from the true values (the bias dominates on the standard deviation). So there is a practical minimum flux level for the reconstruction to achieve a good estimation.

C. Image Spectral Bandwidth Effect

We assumed that the image channel was strictly monochromatic in the phase estimation algorithm already presented. For practical implementation it is necessary to determine the behavior of the data reduction process with respect to the source spectral bandwidth.[29,36]

The data reduction method was applied to polychromatic images numerically simulated with different spectral bandwidths as described in the Subsection 5.A. The aberrated optical path Δ, used for image formation, has a spatial standard deviation of 63-nm rms and is assumed to be independent of wavelength. It was described with the first 21 Zernike polynomials (see Table 1), and the estimated phase was determined with the same number of polynomials to eliminate aliasing errors of this study. The chosen image size is 64×64 pixels. The polychromatic images were generated, assuming that the spatial distribution of the object luminance was in-

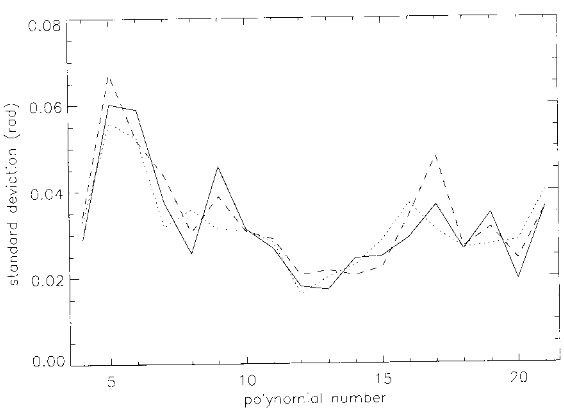


Fig. 9. Standard deviation of estimated Zernike coefficients on 50 image pairs of 32×32 pixels for different values of Zernike coefficients a_4–a_{21} used to simulate the distorted phase, dotted curve, $a_l = 0$ rad; dashed curve, $a_l = 0.037$ rad; solid curve, $a_l = 0.05$ rad; and 10^7 photons/image.

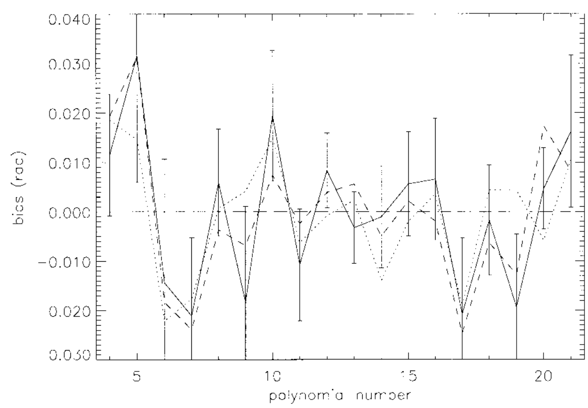
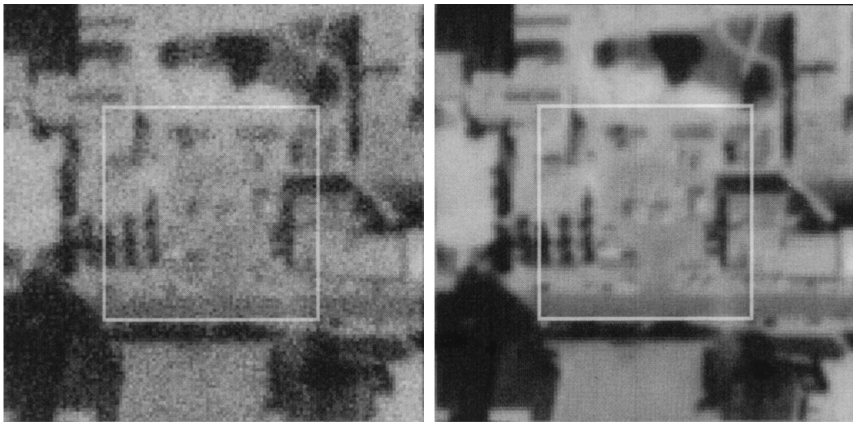


Fig. 10. Bias of estimated Zernike coefficients on 50 image pairs of 32×32 pixels for different values of Zernike coefficients a_4–a_{21} used to simulate the distorted phase, dotted curve, $a_l = 0$ rad; dashed curve, $a_l = 0.037$ rad; solid curve $a_l = 0.05$ rad; and 10^7 photons per image. For this final case the $\pm 3\sigma$ error bars are plotted.



(a) (b)

Fig. 11. Images in the focal plane for different fluxes and $a_4 - a_{21} = 0.05$ rad. The square indicates the actual recorded image taken into account for the algorithm. (a) Total flux level is 10^6 photons; (b) total flux level is 10^9 photons.

dependent of the wavelength. No noise was added to the images. Finally, the data were reduced, assuming image formation at the mean wavelength of the spectral bandwidth with the adequate oversampling, considering that the Shannon criterion is satisfied at the shortest wavelength, $\lambda_{\min} = 440$ nm.

We consider three different spectral bandwidths: 100, 200, and 300 nm. The polychromatic PSF is simulated as described in Subsection 5.A by addition of 7, 11, and 14 monochromatic PSF's, respectively, that span the spectral interval. The results are summarized in Table 2. For example, the residual phase error, i.e., the difference between the estimated and the true phases, with 300-nm spectral bandwidth noiseless images, is $\lambda/200$ rms. The same accuracy is obtained with 64×64 pixel monochromatic noisy images with 10^8 photons.

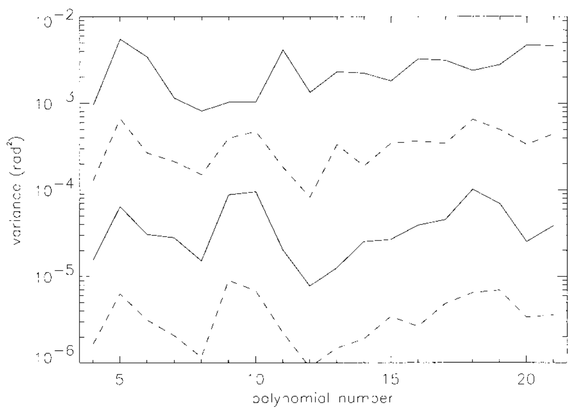


Fig. 12. Noise variance of estimated Zernike coefficients on 50 image pairs of 64×64 pixels for different total photon numbers/image. From the top, $N_{\mathrm{ph}} = 10^6$, 10^7, 10^8, and 10^9 photons, and $a_4 - a_{21} = 0.05$ rad.

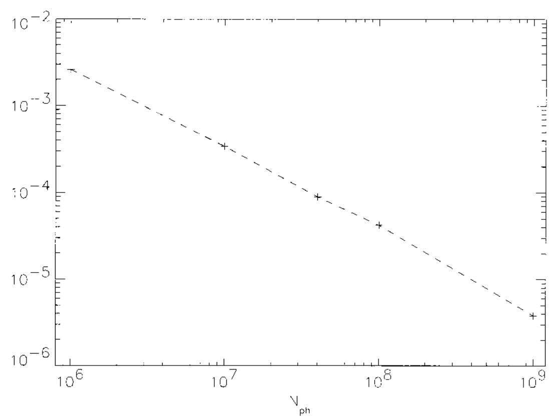


Fig. 13. Average noise variance of estimated Zernike coefficients versus the detected flux level/image.

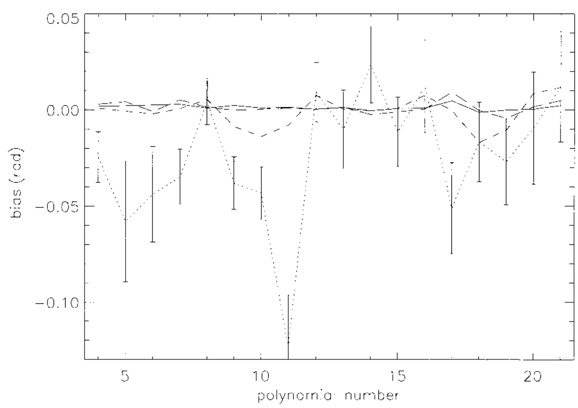


Fig. 14. Bias of estimated Zernike coefficients on 50 image pairs of 64×64 pixels for different numbers of photons/image. Solid line, $N_{\mathrm{ph}} = 10^9$; long-dashed curve, $N_{\mathrm{ph}} = 10^8$; dashed curve, $N_{\mathrm{ph}} = 10^7$; dotted curve, $N_{\mathrm{ph}} = 10^6$; and $a_4 - a_{21} = 0.05$ rad. For this final case, the $\pm 3\sigma$ error bars are plotted.

Table 2. Phase Residual Error in Function of the Spectral Bandwidth

Spectral bandwidth $\Delta\lambda$ (nm)[a]	100	200	300
Phase residual error (λ rms)	$\lambda/700$	$\lambda/350$	$\lambda/200$

[a] $\lambda_{\min} = 440$ nm.

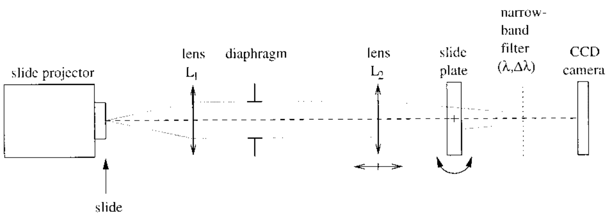


Fig. 15. Optical setup of phase diversity experiment.

This short study demonstrates that the phase diversity method can be used with quite a large spectral bandwidth in the visible spectrum. Therefore, to minimize the total error, a trade-off has to be made between errors that are due to the spectral bandwidth and those that are due to the limited number of detected photons.

6. Experimental Results

To validate the main results presented in Sections 4 and 5, an experiment was performed. The experimental setup is presented in Fig. 15. The extended source is simulated with a slide, illuminated with a slide projector placed at the focus of lens L_1. The slide is imaged on a CCD camera placed near the focal plane of lens L_2. The focal and the out-of-focus images are recorded in succession by a translation of L_2. These two images are shown in Fig. 16. The spectral bandwidth is limited by a filter (central wavelength 633 nm with 10-nm width). A parallel face plate placed on a rotational stage is installed between L_2 and the CCD camera to generate known aberrations (astigmatism mainly).

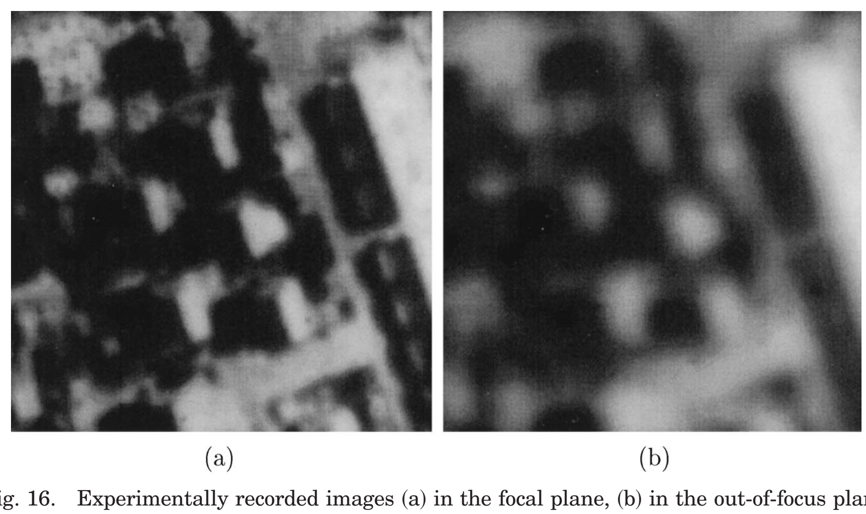
The data reduction software was modified to reduce the experimental data, with the geometric misregistration between the two observation planes taken into account.[12] Table 3 presents a comparison between the generated theoretical astigmatism coefficients and those estimated by the phase diversity technique for different angular positions θ of the plate. The differences between theoretical and estimated values are lower than $\lambda/125$ rms and validate this experimental estimation.

The measurement repeatability was then studied with 50 image pairs of 64×64 pixels (cf. Subsection 5.B). Figure 17 presents the variance of the estimated Zernike coefficients for different fluxes (5×10^5, 5×10^6, 5×10^7, and 5×10^8 photons/image). For low flux (here 5×10^5 and 5×10^6) the general behavior and the level of the variances are similar to those obtained by simulation. The fluctuations of the estimated Zernike coefficients, owing to the photon noise in the images, are predominant. As expected, the variance level is inversely proportional to the number of photons in the images. Nevertheless for higher fluxes (5×10^7 and 5×10^8) it is no longer dominated by the photon noise effect but rather by other experimental sources of fluctuation such as the fluctuations of the air refraction index along the optical path. The experimental bench was placed in a box to reduce this effect, but it is still noticeable for high flux levels. Therefore the spectrum of the Zernike coefficient variance is modified, especially for low-order aberrations.

7. Conclusion

In this paper we have presented a data reduction method for the phase diversity technique to estimate the wave front from two images of widely extended sources.

We have theoretically studied the effect of image photon noise on the phase estimation. For the first time to our knowledge we have demonstrated the increase of the noise propagation with the aberration order as opposed to other conventional wave-front sensors (Shack–Hartmann or curvature sensor) and the existence of an optimal amount of defocus between the two observation planes. We have confirmed the noise behavior of the phase diversity by numerical simulations and have noticed a bias appearing on the phase estimation at low light levels. We have also shown the capability of phase diversity to work within a large spectral bandwidth. In addi-



(a) (b)

Fig. 16. Experimentally recorded images (a) in the focal plane, (b) in the out-of-focus plane.

Table 3. Theoretical and Estimated Astigmatism Zernike Coefficients in the Function of the Angular Position θ of the Slide Plate with Respect to the Optical Axis

θ (°)	0°	15°	30°	−30°	45°
$a_{6,\mathrm{true}}$ (rad)	0	−0.056	−0.226	−0.226	−0.507
$a_{6,e}$ (rad)	0.006	−0.043	−0.197	−0.247	−0.556

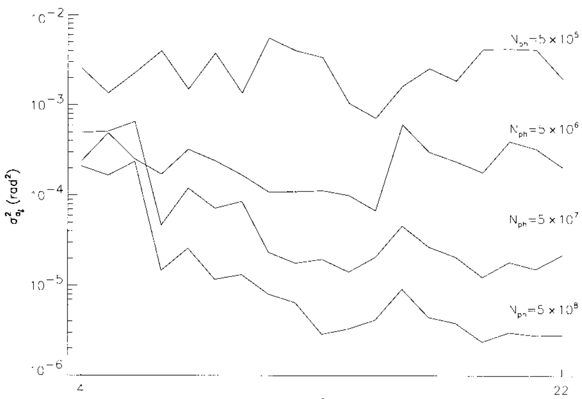


Fig. 17. Experimental noise variance of estimated Zernike coefficients on 50 image pairs of 64 × 64 pixels for different numbers of photons/image.

tion, we have verified by experiment the performance of phase diversity applied to wave-front sensing with widely extended sources.

In the field of data reduction, better phase regularization is currently under study. It should allow for reduction of the noise propagation and of the aliasing effect that is due to the high-order aberrations. We also plan to perform an experimental comparison between the phase diversity technique and the Shack–Hartmann wave-front sensor for extended objects.

Appendix A: Conjugate Gradient Minimization for an Extended Source

The mean-squared error between the image in each observation plane and the convolution product of the estimated object and the estimated PSF is given by relation (4):

$$E = \sum_i |\tilde{I}_1(\mathbf{f}_i) - \tilde{O}(\mathbf{f}_i)\tilde{S}_1(\mathbf{f}_i)|^2 + |\tilde{I}_2(\mathbf{f}_i) - \tilde{O}(\mathbf{f}_i)\tilde{S}_2(\mathbf{f}_i)|^2. \tag{A1}$$

The error is expressed in the Fourier domain to simplify the first derivative calculation (i.e., the convolution product is replaced with the multiplication).

The DFT of X is given by

$$\tilde{X}_{k,l} = \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N} X_{n,m}\exp\left[\frac{-2i\pi(nk+ml)}{N}\right], \tag{A2}$$

where $i^2 = -1$ and the inverse DFT of X is

$$X_{n,m} = \sum_{k=1}^{N}\sum_{l=1}^{N} \tilde{X}_{k,l}\exp\left[\frac{2i\pi(nk+ml)}{N}\right], \tag{A3}$$

where N^2 is the pixel number/image.

The gradient (or first partial derivative) of the error metric E with respect to the parameter X is deduced from gradients in the Fourier domain[27,37]:

$$\frac{\partial E}{\partial X} = \frac{1}{N^2}\,\mathrm{DFT}^{-1}\left(\frac{\partial E}{\partial \tilde{X}}\right), \tag{A4}$$

with the abusive notation

$$\frac{\partial E}{\partial \tilde{X}} = \frac{\partial E}{\partial \tilde{X}_R} + i\,\frac{\partial E}{\partial \tilde{X}_I}, \tag{A5}$$

where $\tilde{X}_R = \mathrm{Re}(\tilde{X})$ is the real part of $\tilde{X}$ and $\tilde{X}_I = \mathrm{Im}(\tilde{X})$, is the imaginary part of $\tilde{X}$.

The gradient of the error E with respect to the DFT of the object $\tilde{O}$ for two observation planes therefore reads as

$$\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}) = 2[\tilde{S}_1^*(\tilde{O}\tilde{S}_1 - \tilde{I}_1) + \tilde{S}_2^*(\tilde{O}\tilde{S}_2 - \tilde{I}_2)](\mathbf{f}). \tag{A6}$$

In the same way the gradient of E with respect to the DFT of the PSF in the kth observation plane $\tilde{S}_k$ is

$$\frac{\partial E}{\partial \tilde{S}_k}(\mathbf{f}) = 2[\tilde{O}^*(\tilde{O}\tilde{S}_k - \tilde{I}_k)](\mathbf{f}), \tag{A7}$$

and the gradients of E with respect to O and S_k are simply [see Eq. (A4)]

$$\frac{\partial E}{\partial O}(\mathbf{r}) = \frac{1}{N^2}\,\mathrm{DFT}^{-1}\left[\frac{\partial E}{\partial \tilde{O}}(\mathbf{f})\right],$$
$$\frac{\partial E}{\partial S_k}(\mathbf{r}) = \frac{1}{N^2}\,\mathrm{DFT}^{-1}\left[\frac{\partial E}{\partial \tilde{S}_k}(\mathbf{f})\right], \tag{A8}$$

where N is the number of pixels in the image side.

To enforce the object positivity, the object O is described as the square of a function Ω, $O(\mathbf{r}) = \Omega^2(\mathbf{r})$. The gradient of E with respect to these parameters Ω is[27]

$$\frac{\partial E}{\partial \Omega}(\mathbf{r}) = \frac{\partial E}{\partial O}\frac{\partial O}{\partial \Omega}(\mathbf{r}) = 2\Omega(\mathbf{r})\,\frac{\partial E}{\partial O}(\mathbf{r}). \tag{A9}$$

To smooth the phase, and to take into account a circular support constraint in the pupil plane, the phase is projected onto the Zernike polynomial base. We search the gradient of the error with respect to a limited number M of Zernike polynomial coefficients[12]:

$$\frac{\partial E}{\partial a_l} = \sum_{k=1}^{2}\left(\sum_{j\in\mathrm{support}}\left\{\sum_i\left[\frac{\partial E}{\partial S_k(\mathbf{r}_i)}\frac{\partial S_k}{\partial \varphi(\mathbf{x}_j)}\right]\frac{\partial \varphi}{\partial a_l}\right\}\right), \tag{A10}$$

where $\varphi(\mathbf{x}) = \sum_{l=1}^{M} a_l Z_l(\mathbf{x})$ and a_l (in radians) is the coefficient of the lth Zernike polynomial Z_l. Other advantages include an implicit PSF positivity as is seen in Eq. (2) and a small number of parameters to be optimized (only a limited number of coefficients instead of a great number of points). The gradient of S_k with respect to a_l is[27]

$$\frac{\partial S_k(\mathbf{r}_i)}{\partial a_l} = -2\,\mathrm{Im}\{\tilde{A}_k^*(\mathbf{r}_i)[\tilde{A}_k(\mathbf{r}_i) * \tilde{Z}_l(\mathbf{r}_i)]\}, \tag{A11}$$

where $\mathrm{Im}(X)$ is the imaginary part of the complex variable X and X^* is the complex conjugate of X.

For convenience in the following developments we define I' as

$$I'(\mathbf{r}) = I(\mathbf{r})\Pi(\mathbf{r}), \tag{A12}$$

where I is the full image of an infinite extended source [see Eq. (1)] and Π is defined by

$$\Pi(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \in \text{FOV} \\ 0 & \text{otherwise} \end{cases}. \tag{A13}$$

So the error [see relation (4)] becomes

$$E = \sum_i |\tilde{I}_1'(\mathbf{f}_i) - [\tilde{O}(\mathbf{f}_i)\tilde{S}_1(\mathbf{f}_i)] * \tilde{\Pi}(\mathbf{f}_i)|^2 + |\tilde{I}_2'(\mathbf{f}_i) - [\tilde{O}(\mathbf{f}_i)\tilde{S}_2(\mathbf{f}_i)] * \tilde{\Pi}(\mathbf{f}_i)|^2. \tag{A14}$$

To simplify the following equations, the discrete convolution product is noted:

$$C_k(\mathbf{f}_i) = \sum_j \tilde{O}(\mathbf{f}_j)\tilde{S}_k(\mathbf{f}_j)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_j). \tag{A15}$$

The gradient of E with respect to $\tilde{O}$ is defined by [see Eq. (A4)]

$$\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}_i) = \left[\frac{\partial}{\partial \tilde{O}_R(\mathbf{f}_i)} + i\frac{\partial}{\partial \tilde{O}_I(\mathbf{f}_i)}\right]E, \tag{A16}$$

$$\frac{\partial}{\partial \tilde{O}_R(\mathbf{f}_0)}E = \sum_k\sum_i\left(C_k\frac{\partial}{\partial \tilde{O}_R}C_k^* + C_k^*\frac{\partial}{\partial \tilde{O}_R}C_k - \tilde{I}'\frac{\partial}{\partial \tilde{O}_R}C_k^* - \tilde{I}'^*\frac{\partial}{\partial \tilde{O}_R}C_k\right)(\mathbf{f}_i),$$

$$\frac{\partial}{\partial \tilde{O}_I(\mathbf{f}_0)}E = \sum_k\sum_i\left(C_k\frac{\partial}{\partial \tilde{O}_I}C_k^* + C_k^*\frac{\partial}{\partial \tilde{O}_I}C_k - \tilde{I}'\frac{\partial}{\partial \tilde{O}_I}C_k^* - \tilde{I}'^*\frac{\partial}{\partial \tilde{O}_I}C_k\right)(\mathbf{f}_i). \tag{A17}$$

Then

$$\frac{\partial}{\partial \tilde{O}_R(\mathbf{f}_0)}C_k(\mathbf{f}_i) = \frac{\partial}{\partial \tilde{O}_R(\mathbf{f}_0)}\sum_j \tilde{O}(\mathbf{f}_j)\tilde{S}_k(\mathbf{f}_j)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_j) = \tilde{S}_k(\mathbf{f}_0)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_0), \tag{A18}$$

$$\frac{\partial}{\partial \tilde{O}_R(\mathbf{f}_0)}C_k^*(\mathbf{f}_i) = \tilde{S}_k^*(\mathbf{f}_0)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_0), \tag{A19}$$

$$\frac{\partial}{\partial \tilde{O}_I(\mathbf{f}_0)}C_k(\mathbf{f}_i) = i\tilde{S}_k(\mathbf{f}_0)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_0), \tag{A20}$$

$$\frac{\partial}{\partial \tilde{O}_I(\mathbf{f}_0)}C_k^*(\mathbf{f}_i) = -i\tilde{S}_k^*(\mathbf{f}_0)\tilde{\Pi}(\mathbf{f}_i - \mathbf{f}_0). \tag{A21}$$

So Eq. (A16) becomes

$$\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}_0) = 2\sum_k \tilde{S}_k^*(\mathbf{f}_0)\sum_i [C_k(\mathbf{f}_i) - \tilde{I}'(\mathbf{f}_i)]\tilde{\Pi}^*(\mathbf{f}_i - \mathbf{f}_0),$$

$$\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}_0) = 2(\tilde{S}_1^*(\mathbf{f}_0)\{[\tilde{O}(\mathbf{f}_0)\tilde{S}_1(\mathbf{f}_0)] * \tilde{\Pi}(\mathbf{f}_0) - \tilde{I}'(\mathbf{f}_0)\} * \tilde{\Pi}^*(-\mathbf{f}_0) + \tilde{S}_2^*(\mathbf{f}_0) \times \{[\tilde{O}(\mathbf{f}_0)\tilde{S}_2(\mathbf{f}_0)] * \tilde{\Pi}(\mathbf{f}_0) - \tilde{I}'(\mathbf{f}_0)\} * \tilde{\Pi}^*(-\mathbf{f}_0)).$$

With the DFT properties[38] and with Π a real function,

$$\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}_0) = 2(\tilde{S}_1^*(\mathbf{f}_0)\{[\tilde{O}(\mathbf{f}_0)\tilde{S}_1(\mathbf{f}_0)] * \tilde{\Pi}(\mathbf{f}_0) - \tilde{I}'(\mathbf{f}_0)\} * \tilde{\Pi}(\mathbf{f}_0) + \tilde{S}_2^*(\mathbf{f}_0)\{[\tilde{O}(\mathbf{f}_0)\tilde{S}_2(\mathbf{f}_0)] * \tilde{\Pi}(\mathbf{f}_0) - \tilde{I}'(\mathbf{f}_0)\} * \tilde{\Pi}(\mathbf{f}_0)). \tag{A22}$$

In the same way the gradient of E with respect to $\tilde{S}_k$ is given by

$$\frac{\partial E}{\partial \tilde{S}_k}(\mathbf{f}_0) = 2\tilde{O}^*(\mathbf{f}_0)\{[\tilde{O}(\mathbf{f}_0)\tilde{S}_k(\mathbf{f}_0)] * \tilde{\Pi}(\mathbf{f}_0) - \tilde{I}'(\mathbf{f}_0)\} * \tilde{\Pi}(\mathbf{f}_0), \tag{A23}$$

$$\frac{\partial E}{\partial O}(\mathbf{r}_0) = \frac{4}{N'^2}\mathrm{DFT}^{-1}\left[\frac{\partial E}{\partial \tilde{O}}(\mathbf{f}_0)\right],$$

$$\frac{\partial E}{\partial S_k}(\mathbf{r}_0) = \frac{4}{N'^2}\mathrm{DFT}^{-1}\left[\frac{\partial E}{\partial \tilde{S}_k}(\mathbf{f}_0)\right], \tag{A24}$$

with $N' = 2N$, the number of pixels of a double support side.

The new expression of the PSF gradient with a guard band can be used in Eq. (A10) to obtain the gradient with respect to the Zernike coefficients.

Appendix B: Second Partial Derivatives of the Error Metric for Noise Propagation

The second partial derivatives of the error metric with respect to the images and the Zernike coefficients are deduced from equations (A7) and (A11). In addition, the determination of the noise propagation requires only the computation of the second partial derivatives of the error metric with the mean of the measurements and the corresponding estimated parameters (see Section 4). The difference between a noise-free image and the corresponding estimation is equal to zero, such as its Fourier transform:

$$\langle \tilde{I}_k \rangle - \tilde{O}\tilde{S}_{k,e} = 0. \tag{B1}$$

Equation (A7) is also equal to zero. Therefore, in the computation of the derivative of $\partial E/\partial a_l$ with respect to Zernike coefficients, only the derivative of the DFT of the PSF with respect to $a_{l'}$ is not null:

$$[\nabla^2_{a_{l'},a_l} E] = \frac{\partial^2 E}{\partial a_l \partial a_l'} = 2\,\mathrm{Re}\left(\sum_{k=1}^{2}\sum_{h=1}^{N^2} |\tilde{O}(\mathbf{f}_h)|^2 \frac{\partial \tilde{S}^*_{k,e}}{\partial a_l}\frac{\partial \tilde{S}_{k,e}}{\partial a_l'}\right), \tag{B2}$$

and its full expression when accounting for equation (A12) is

$$\frac{\partial^2 E}{\partial a_l \partial a_l'} = 8\sum_{k=1}^{2}\sum_{h=1}^{N^2} |\tilde{O}(\mathbf{f}_h)|^2 (\mathrm{DFT}\{\mathrm{Im}[\tilde{A}^*_{k,e} \times (\tilde{A}_{k,e} * \tilde{Z}_l)]\})^* \mathrm{DFT}\{\mathrm{Im}[\tilde{A}^*_{k,e}(\tilde{A}_{k,e} * \tilde{Z}_{l'})]\}. \tag{B3}$$

The derivative of $\partial E/\partial a_l$ with respect to the images consists simply of deriving the DFT of the image:

$$\frac{\partial \tilde{I}_k(\mathbf{f}_h)}{\partial I_k(\mathbf{r}_i)} = \frac{1}{N^2}\exp\left(-i\,\frac{2\pi}{N}\,\mathbf{f}_h\mathbf{r}_i\right). \tag{B4}$$

Therefore the second derivative of the error with respect to the Zernike coefficients and the images is

$$[\nabla^2_{I_k,a_l} E] = \frac{\partial^2 E}{\partial a_l \partial I_k(\mathbf{r}_i)} = -\frac{2}{N^2}\left[O(\mathbf{r}_i) * \frac{\partial S_{k,e}(\mathbf{r}_i)}{\partial a_l}\right], \tag{B5}$$

and its full expression when accounting for Eq. (A11) is given by

$$\frac{\partial^2 E}{\partial a_l \partial I_k(\mathbf{r}_i)} = \frac{4}{N^2}\sum_{k=1}^{2}\mathrm{Im}(\{[\tilde{Z}_l(\mathbf{r}_i) * \tilde{A}_{k,e}(\mathbf{r}_i)] \times \tilde{A}^*_{k,e}(\mathbf{r}_i)\} * O(\mathbf{r}_i)). \tag{B6}$$

The product of the matrix $[\nabla_{I_k(\mathbf{r}),a_l}{}^2\ E]$ with its transpose matrix is equal to

$$[\nabla^2_{I_k,a_l} E]^t[\nabla^2_{I_k,a_{l'}} E] = \frac{4}{N^4}\sum_{k=1}^{2}\sum_i \mathrm{Re}\left\{\left[O(\mathbf{r}_i) * \frac{\partial S_{k,e}(\mathbf{r}_i)}{\partial a_l}\right] \times \left[O(\mathbf{r}_i) * \frac{\partial S_{k,e}(\mathbf{r}_i)}{\partial a_{l'}}\right]\right\}. \tag{B7}$$

It is easy to demonstrate with the Parseval theorem that Eq. (B7) is similar to Eq. (B2). In fact we can write

$$[\nabla^2_{I_k,a_l} E]^t[\nabla^2_{I_k,a_{l'}} E] = \frac{2}{N^2}[\nabla^2_{a_{l'},a_l} E]. \tag{B8}$$

The authors thank C. Aime and H. Lanteri of the Université de Nice—Sophia Antipolis and E. Thiébaut of the Université Claude Bernard Lyon I for fruitful discussions as well as F. Mendez for his contribution to the experimental part. Special thanks are extended to both anonymous referees, whose suggestions greatly improved the clarity of this paper.

References

1. G. Rousset, "Wavefront sensing," in *Adaptive Optics for Astronomy*, D. M. Alloin and J.-M. Mariotti, eds., Vol. 423 of NATO Asi Series C: Mathematical and Physical Sciences (Kluwer Academic, Dordrecht, The Netherlands, 1994).
2. O. Von der Lühe, "Wavefront error measurement technique using extended, incoherent light sources," Opt. Eng. **278,** 1078–1087 (1988).
3. V. Michau, G. Rousset, and J.-C. Fontanella, "Wavefront sensing from extended sources," in *Real Time and Post Facto Solar Image Correction* (National Solar Observatory, Sunspot, N.M., 1992), pp. 124–128.
4. R. Kupke, F. Roddier, and D. L. Mickey, "Curvature-based wavefront sensor for use on extended patterns," in *Adaptive Optics in Astronomy*, M. A. Ealey and F. Merkle, eds., Proc. SPIE **2201,** 519–527 (1994).
5. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," Appl. Opt. **33,** 6533–6546 (1994).
6. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21,** 829–832 (1982).
7. R. G. Paxman and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," J. Opt. Soc. Am. A **5,** 914–923 (1988).
8. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9,** 1072–1085 (1992).
9. N. Miller and A. Ling, "Imaging with phase diversity: experimental results," in *Digital Image Recovery and Synthesis II*, P. S. Idell, ed., Proc. SPIE **2029,** 227–238 (1993).
10. R. A. Carreras, S. R. Restaino, G. D. Love, G. L. Tarr, and J. S. Fender, "Phase diversity experimental results: deconvolution of ν Scorpii," Opt. Commun. **130,** 13–19 (1996).
11. M. G. Löfdahl and G. B. Scharmer, "Application of phase-diversity to solar images," in *Image Reconstruction and Restoration*, T. J. Schulz and D. L. Snyder, eds., Proc. SPIE **2302,** 254–267 (1994).
12. J. H. Seldin and R. G. Paxman, "Phase-diverse speckle reconstruction of solar data," in *Image Reconstruction and Restoration*, T. J. Schulz and D. L. Snyder, eds., Proc. SPIE **2302,** 268–280 (1994).
13. R. G. Paxman, J. H. Seldin, M. G. Löfdahl, G. B. Scharmer, and C. U. Keller, "Evaluation of phase-diversity techniques for solar-image restoration," Astrophys. J. **466,** 1087–1099 (1996).
14. C. Roddier and F. Roddier, "Reconstruction of the Hubble Space Telescope mirror figure from out-of-focus stellar images," in *Space Astronomical Telescopes and Instruments*, P. Y. Bely and J. B. Breckinridge, eds., Proc. SPIE **1494,** 78–84 (1991).
15. J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, "Hubble Space Telescope characterized by using phase-retrieval algorithms," Appl. Opt. **32,** 1747–1767 (1993).
16. C. Roddier and F. Roddier, "Combined approach to the Hubble Space Telescope wave-front distortion analysis," Appl. Opt. **32,** 2992–3008 (1993).
17. B. L. Ellerbroek, B. J. Thelen, D. J. Lee, D. A. Carrara, and R. G. Paxman, "Comparison of Shack–Hartmann wavefront sensing and phase-diverse phase retrieval," in *Adaptive Optics and Applications*, R. K. Tyson and R. Q. Fugate, eds., Proc. SPIE **3126,** 307–320 (1997).
18. D. J. Lee, B. Welsh, and M. Roggemann, "Cramer–Rao evaluation of phase diversity aberration sensing," in *Image Reconstruction and Restoration II*, T. J. Schulz, ed., Proc. SPIE **3170,** 161–172 (1997).
19. J. R. Fienup, B. J. Thelen, R. G. Paxman, and D. A. Carrara, "Comparison of phase-diversity and curvature wavefront sensing," in *Adaptive Optical Systems Technologies*, D. Bonaccini and R. K. Tyson, eds., Proc. SPIE **3353,** 930–940 (1998).
20. J. R. Fienup and C. C. Wackerman, "Phase retrieval stagnation problems and solutions," J. Opt. Soc. Am. **3,** 1897–1907 (1986).

21. R. A. Gonsalves, "Phase retrieval by differential intensity measurements," J. Opt. Soc. Am. A **4,** 166–170 (1987).
22. J. H. Seldin and J. R. Fienup, "Numerical investigation of the uniqueness of phase retrieval," J. Opt. Soc. Am. A **7,** 412–427 (1990).
23. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66,** 207–211 (1976).
24. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21,** 2758–2769 (1982).
25. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, in *Numerical Recipes in C*, 2nd ed. (Cambridge U. Press, New York, 1992).
26. Y. Biraud, "A new approach for increasing the resolving power by data processing," Astron. Astrophys. **1,** 124–127 (1969).
27. E. Thiébaut and J. M. Conan, "Strict *a priori* constraints for maximum-likelihood blind deconvolution," J. Opt. Soc. Am. A **12,** 485–492 (1995).
28. J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography," IEEE Trans. Image Process. **5,** 493–506 (1996).
29. L. Meynadier, "Analyse de surface d'onde pour le contrôle actif d'un télescope spatial," Ph.D. dissertation (Université de Nice-Sophia Antipolis, Nice, France, 1997).
30. D. R. Williams, J. Liang, and D. T. Miller, "Adaptive optics for the human eye," in *Adaptive Optics,* Vol. 13 of 1996 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1996), pp. 145–147.
31. G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," IEEE Trans. Acoust. Speech Signal Process. **37,** 2024–2036 (1989).
32. B. J. Thelen, R. G. Paxman, J. H. Seldin, and D. R. Rice, "Bayes estimation of dynamic and fixed aberrations, and object from phase-diverse speckle data," in *Digital Image Recovery and Synthesis III*, P. S. Idell and T. J. Schulz, eds., Proc. SPIE **2827,** 32–42 (1996).
33. L. M. Mugnier, J.-M. Conan, V. Michau, and G. Rousset, "Imagerie à travers la turbulence par déconvolution myope multitrame," in *Seizième Colloque sur le Traitement du Signal et des Images GRETSI* (Groupe de Recherche et d'Études sur le Traitement du Signal et des Images, Grenoble, France, 1997), pp. 567–570.
34. R. G. Paxman, B. J. Thelen, and J. H. Seldin, "Phase-diversity correction of turbulence-induced space-variant blur," Opt. Lett. **19,** 1231–1233 (1994).
35. N. Baba, H. Tomita, and N. Miura, "Iterative reconstruction method in phase-diversity imaging," Appl. Opt. **33,** 4428–4433 (1994).
36. L. Meynadier, V. Michau, and M. T. Velluet, "Wavefront sensing by phase diversity," in *THEMIS Forum* (Observatoire de Paris, Paris, 1996), pp. 205–210.
37. R. G. Lane, "Blind deconvolution of speckle images," J. Opt. Soc. Am. A **9,** 1508–1514 (1992).
38. R. N. Bracewell, *The Fourier Transform and its Applications*, 2nd revised ed. (McGraw-Hill, New York, 1986).