Third AO4ELT Conference - Adaptive Optics for Extremely Large Telescopes
Florence, Italy. May 2013
ISBN: 978-88-908876-0-4
DOI: 10.12839/AO4ELT3.13243

REAL-TIME WAVEFRONT RECONSTRUCTION FROM INTENSITY MEASUREMENTS

Carlas S. Smith, Raluca Marinică[a], and Michel Verhaegen

Delft University of Technology, Delft Center for Systems and Control, Mekelweg 2, 2628 CD Delft, The Netherlands

Abstract. We propose an efficient approximation to the nonlinear phase diversity method for wavefront reconstruction from intensity measurements. The new method, iterative linear phase diversity (ILPD), assumes that the residual phase aberration is small and makes use of a first order Taylor expansion of the point spread function (PSF) performed for an arbitrary (large) diversity in order to optimize the phase retrieval. For static aberrations, ILPD makes use of two images collected at each iteration of the algorithm. In each step, the residual phase aberrations are estimated by solving a linear least squares problem, followed by the use of a deformable mirror to correct for the aberrations. A further contribution of the paper is the extension of the static ILPD to the case of dynamic wavefront reconstruction for which a computationally efficient H_2 controller is presented.

1 Introduction

Phase diversity (PD) [1] estimates wavefront aberrations using nonlinear optimization techniques from multiple images of the same unknown scene acquired simultaneously, which contain additional user introduced aberrations (diversities). To be able to uniquely estimate wavefront aberrations more than one image is needed [2]. Due to the high computational complexity and possible convergence to local optima [2], nonlinear PD has a limited usage in real-time correction algorithms [1] and different ideas have been presented to decrease its complexity [3,4]. Recently, in [5], it was shown that using a second order expansion of the generalized pupil function (GPF), wavefront retrieval algorithms give more accurate results than using the Born approximation [3,6]. The common idea in decreasing the computational complexity is the approximation of the PSF based on the assumption that the total aberration is small [5,3,7]. However, as has been shown in [8], the optimal diversity depends on the present aberration and can generally not be considered small. In the present paper we overcome this shortcoming by the use of an alternative approximation of the PSF. The linearization of the PSF is done around the diversity for small values of the phase aberration. A similar approximation is used in [4] in an iterative manner based on only one measurement - linearized focal-plane technique (LIFT). At each iteration a new linearization of the PSF is done around the current phase estimate and a new least squares (LS) problem is solved to obtain the estimate. Using only the PSF approximation around zero aberration, we present a novel method called iterative linear phase diversity (ILPD), which consists in collecting two diversity images, solving a LS problem and correcting for the wavefront aberrations by the LS estimate. As opposed to LIFT, we use the approximation around zero and collect two new images at each iteration. In this way we eliminate the sign ambiguity in the recovered phase and also increase the speed of the algorithm due to the fact that the linear coefficients of the PSF do not change from iteration to iteration.

[a] Corresponding author: r.m.marinica@tudelft.nl

The paper is organized as follows. In Section 2 we introduce the optical system. In Section 3 we review linear and quadratic PSF approximations. In Sections 4 and 5 we present the ILPD solution for static and time varying aberrations, respectively. In Section 6 we discuss results of numerical simulations. We end with conclusions in Section 7.

The mathematical notations used are standard: $\bullet^T$, $\bullet^*$, and $\star$ denote transposition, transpose conjugation, and the convolution operator, respectively. $\| \bullet \|$ denotes the vector 2-norm, $\mathcal{O}(\| \bullet \|^a)$ is the a-th order Lagrange residue, the infinity norm of a vector v is defined as $\|v\|_\infty := \max(|v_1|, \ldots, |v_n|)$, and the $\mathcal{H}_\infty$ norm of a rational transfer function $M(z)$ is defined as $\|M\|_\infty := \sup_\omega \overline{\sigma}(M(j\omega))$, where $\overline{\sigma}$ is the maximum singular value of the matrix $M(j\omega)$.

2 The optical system

The incoherent image formation of a point source in the presence of phase aberrations $\phi_i(u_j, v_j) = Z(u_j, v_j)^T \alpha$ and a phase diversity β_i is given by [9]

$$y_{i,j} = \mu_i h(s_j, t_j, \alpha, \beta_i) + n_i(s_j, t_j), \tag{1}$$

where $y_{i,j}$ denotes the j-th pixel of the i-th diversity image, μ_i is the number of photons, h denotes the spatially invariant PSF expressed in the spatial coordinates (s_j, t_j), and $n_i(s_j, t_j)$ is Gaussian white noise with standard deviation $\sigma_{i,j}$, which we assume to be equal for all pixels by dropping the index j. The phase aberrations are approximated using a normalized Zernike basis [10], where Z contains the n Zernike polynomials evaluated in the pupil plane coordinates (u_j, v_j). The spatially invariant PSF of the i-th optical path in Eq. (1) can be written as

$$h(s_j, t_j, \alpha, \beta_i) = \mathcal{F}(\Pi(u, v) \exp(i\phi_i(u, v)))(s_j, t_j) \, \mathcal{F}(\Pi(u, v) \exp(i\phi_i(u, v)))^*(s_j, t_j) \tag{2}$$

where $\mathcal{F}$ is the Fourier transform and Π is the pupil function. The quantity $p(u_j, v_j, \alpha, \beta_i) = \Pi(u_j, v_j) \exp(i\phi_i(u_j, v_j))$ in Eq. (2) is the GPF. It is assumed that the measurements are dominated by Gaussian read-out noise. The signal to noise (SNR) level is given by

$$\mathrm{SNR} = {}^1\!/_{m^2} \sum_{j=1}^{m^2} \mu_i h_j(\alpha, \beta_i) \left({}^1\!/_{m^2} \sum_{j=1}^{m^2} \sigma_i^2 \right)^{-1/2} \left(= \frac{\mu_i}{m^2 \sigma_i} \right), \tag{3}$$

where we have omitted the coordinates in brackets and added the pixel index j for h_j.

3 Approximations of the PSF

The PSF in Eq. (2) is computationally expensive to evaluate and needs to be approximated. The generally used approach is the Born approximation, which is valid for small phase aberrations (up to 0.5 rad rms). It has been shown in [8] that the lower bound on the variance of any unbiased estimator of the wavefront aberration is much lower for large diversities. Therefore, it is of high importance to investigate approximations of Eq. (2) that allow us to use large diversities.

We start with a first order approximation based on the Taylor expansion of the GPF. The Born approximation assumes a small phase (*small total phase approximation*), $\phi_i = Z^T(\alpha + \beta_i)$, such that the GPF can be approximated using only a first order Taylor expansion around $\alpha + \beta_i = 0$. Substituting this first order approximation in Eq. (2) yields a quadratic PSF

$$h_j(\alpha, \beta_i) = A_{0,j} + A_{1,j}(\alpha + \beta_i) + (\alpha + \beta_i)^T A_{2,j}(\alpha + \beta_i) + \mathcal{O}\left(\|\alpha + \beta_i\|^2\right), \tag{4}$$

where $A_{0,j} := h_j(\alpha,\beta_i)|_{\alpha+\beta_i=0}$, $A_{1,j} := \frac{\partial h_j(\alpha,\beta_i)}{\partial(\alpha+\beta_i)}\Big|_{\alpha+\beta_i=0}$, $A_{2,j} := \mathcal{F}\left(\frac{\partial p_j(\alpha,\beta_i)}{\partial(\alpha+\beta_i)} \star \frac{\partial p^*_{-j}(\alpha,\beta_i)}{\partial(\alpha+\beta_i)^T} + \frac{\partial p_j(\alpha,\beta_i)}{\partial(\alpha+\beta_i)} \star \frac{\partial p^*_{-j}(\alpha,\beta_i)}{\partial(\alpha+\beta_i)^T}\right)(s_j,t_j)\Big|_{\alpha+\beta_i=0}$. The linear term of the approximated PSF in Eq. (4) is invariant in the even aberrations, which makes it impossible to neglect the quadratic term of the PSF when the Born approximation is used to formulate a linear estimation problem.

Another first order model is obtained by directly approximating the PSF in Eq. (2) for small aberrations and non-zero diversities (*small aberration approximation*)

$$h_j(\alpha,\beta_i) = B_{0,j}(\beta_i) + B_{1,j}(\beta_i)\alpha + \mathcal{O}\left(\|\alpha\|^2\right), \tag{5}$$

where $B_{0,j}(\beta_i) := h_j(\alpha,\beta_i)|_{\alpha=0}$, and $B_{1,j}(\beta_i) := \frac{\partial h_j(\alpha,\beta_i)}{\partial\alpha}\Big|_{\alpha=0}$. The approximation in Eq. (5) has the property that for $\phi \neq 0$ the even modes do not cancel in this linear term.

It has been shown in [5] that using a second order Taylor expansion of the GPF in $\phi_i = 0$ and neglecting the 3rd and the 4th orders of the resulting PSF (*small total phase approximation*), a more accurate model than Eq. (4) is obtained

$$h_j(\alpha,\beta_i) = C_{0,j} + C_{1,j}(\alpha+\beta_i) + (\alpha+\beta_i)^T C_{2,j}(\alpha+\beta_i) + \mathcal{O}\left(\|\alpha+\beta_i\|^3\right), \tag{6}$$

where $C_{0,j} := A_{0,j}$, $C_{1,j} := A_{1,j}$ and $C_{2,j} := \frac{\partial^2 h_j(\alpha,\beta_i)}{\partial(\alpha+\beta_i)\partial(\alpha+\beta_i)^T}\Big|_{\alpha+\beta_i=0}$.

4 ILPD for static aberrations

Here, we aim to obtain a linear relationship between the measured intensity and the unknown aberration. A linear model has low computational complexity and gives rise to fast algorithms. For the approximation in Eq. (5), we stack the measurements as follows

$$\begin{bmatrix} Y_1^T & Y_2^T \end{bmatrix}^T = b_S + A_S\alpha + \Delta b_S(\alpha) + n, \quad \Delta b_S(\alpha) := \mathcal{O}\left(\|\alpha\|^2\right). \tag{7}$$

where $Y_i := \begin{bmatrix} y_{i,1} \dots y_{i,j} \dots \end{bmatrix}^T$, $b_S := \begin{bmatrix} \tilde{b}_{S,1} \dots \tilde{b}_{S,i} \dots \end{bmatrix}^T$, $\tilde{b}_{S,i} := \begin{bmatrix} B_{0,i,1} \dots B_{0,i,j} \dots \end{bmatrix}$, $A_S := \begin{bmatrix} \tilde{A}_{S,1}^T \dots \tilde{A}_{S,i}^T \dots \end{bmatrix}^T$, $\tilde{A}_{S,i} := \begin{bmatrix} B_{i,1}(\beta_i)^T \dots B_{i,j}(\beta_i)^T \dots \end{bmatrix}^T$, and $n = \begin{bmatrix} n_1^T n_2^T \end{bmatrix}^T$. Using Eq. (6), a set of linear equations is obtained by subtracting 2 intensity measurements

$$Y_1 - Y_2 = b_D + A_D\alpha + \Delta b_D + \Delta n, \quad \Delta b_D := \mathcal{O}\left(\|\alpha+\beta_1\|^3\right) - \mathcal{O}\left(\|\alpha+\beta_2\|^3\right), \tag{8}$$

where $b_D := \begin{bmatrix} \tilde{b}_{D,1} \dots \tilde{b}_{D,j} \dots \end{bmatrix}^T$, $\tilde{b}_{D,j} := C_{1,j}(\beta_1 - \beta_2) + \beta_1^T C_{2,j}\beta_1 - \beta_2^T C_{2,j}\beta_2$, $A_D := \begin{bmatrix} \tilde{A}_{D,1}^T \dots \tilde{A}_{D,j}^T \dots \end{bmatrix}^T$, $\tilde{A}_{D,j} := 2(\beta_1 - \beta_2)^T C_{2,j}$, and $\Delta n = n_1 - n_2$.

In static ILPD, the residual aberration is repeatedly estimated and corrected with a DM using the residual aberration estimate. Denoting the residual aberration estimate at step $k-1$ by $\hat{\alpha}_{k-1}$, and the residual aberration at step $k-1$ by α_{k-1}, ILPD gives $\alpha_k = \alpha_{k-1} - \hat{\alpha}_{k-1}$. At the k-th correction step, two images $Y_{1,k}$, $Y_{2,k}$ are recorded for two different diversities β_1 and β_2 assuming the wavefront aberration does not change. A new estimate of α_k is obtained via the solution of a LS problem based on Eqs. (8) and (7). The algorithm continues until the strength of the aberration decreases under a certain threshold. Let Eqs. (7) and (8) be rewritten as

$$\overline{b}_{S,k} - \Delta b_S(\alpha_k) = A_S\alpha_k + n_k, \tag{9}$$

$$\overline{b}_{D,k} - \Delta b_D(\alpha_k) = A_D\alpha_k + \Delta n_k, \tag{10}$$

respectively, where $\overline{b}_{D,k} := Y_{1,k} - Y_{2,k} - b_D$, $\overline{b}_{S,k} := \begin{bmatrix} Y_{1,k}^T & Y_{2,k}^T \end{bmatrix}^T - b_S$, $\Delta n_k = n_{1,k} - n_{2,k}$, $n_k = \begin{bmatrix} n_{1,k}^T & n_{2,k}^T \end{bmatrix}^T$, with $n_{i,k}$ the stacked measurement noise at time instant k and the i-th diversity image. Then the two LS problems that we need to solve are

$$M_1 : \min_{\alpha_k} \|\overline{b}_{S,k} - A_S\alpha_k\|^2, \quad M_2 : \min_{\alpha_k} \|\overline{b}_{D,k} - A_D\alpha_k\|^2. \tag{11}$$

A convergence analysis can be made using the relative residue after each correction step $r_{\mathrm{LS}} := {\|\alpha_{k+1}\|}/{\|\alpha_k\|}$. The relative residue has to be smaller than 1 to ensure that the rms is reduced.

5 Correction with dynamic aberrations

The statistical properties of atmospheric turbulence can be modeled by the Kolmogorov spectrum and we assume a static model for the DM correction. The linear dynamics of the Kolmogorov model combined with a linear perturbed measurement equation yield

$$\begin{aligned} \alpha_{k+1} &= A_{\text{tur}}\alpha_k + B_{\text{act}}u_k + w_k, \\ \overline{b}_{S,k} &= A_S\alpha_k + \Delta b_S(\alpha_k) + n_k, \end{aligned} \tag{12}$$

where n_k has the covariance matrix $N = \sigma^2 I$. The system in Eq. (12) is in the standard state space form for $\Delta b_S(\alpha_k) \equiv 0$. Since the output $\overline{b}_{S,k}$ has the dimension equal to the number of pixels, we reduce the computational complexity of the calculations using a QR decomposition of $A_S = \begin{bmatrix} Q_1 \, Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ (R_1 upper triangular, and Q_1, Q_2 with orthogonal columns)

$$z_k = R_1\alpha_k + Q_1^T(n_k + \Delta b_S(\alpha_k)), \tag{13}$$

where $z_k := Q_1^T\overline{b}_{S,k}$. Next, we use the Kalman filter theory [11] for the system in Eq. (12), where the output is replaced by Eq. (13). The Kalman filter prediction and update equations are

$$\begin{aligned} \hat{\alpha}_{k|k-1} &= A_{\text{tur}}\hat{\alpha}_{k-1|k-1} + B_{\text{act}}u_k, & P_{k|k-1} &= A_{\text{tur}}P_{k-1|k-1}A_{\text{tur}}^{\text{T}} + Q_k, \\ K_k &= P_{k|k-1}R_1^{\text{T}}\left(R_1P_{k|k-1}R_1^T + Q_1^TNQ_1\right)^{-1}, & \hat{\alpha}_{k|k} &= \hat{\alpha}_{k|k-1} + K_k\left(z_k - R_1\hat{\alpha}_{k|k-1}\right), \\ P_{k|k} &= (I - K_kR_1)P_{k|k-1}. \end{aligned} \tag{14}$$

It is assumed that the atmospheric disturbance model $\mathcal{S}(z)$ is in innovation form [11] with respect to z_k

$$\begin{aligned} \alpha_{k+1} &= A_{\text{tur}}\alpha_k + K_kv_k \\ z_k &= \quad R_1\alpha_k + \quad v_k \end{aligned} \tag{15}$$

where A_{tur}–K_kR_1 and A_{tur} are stable, and v_k is a zero-mean white process with covariance $R_v = R_1P_{k|k-1}R_1^T + Q_1^TNQ_1$. The innovation process v_k incorporates both the effect of atmospheric turbulence and measurement noise. A schematic representation of the control problem in the so-

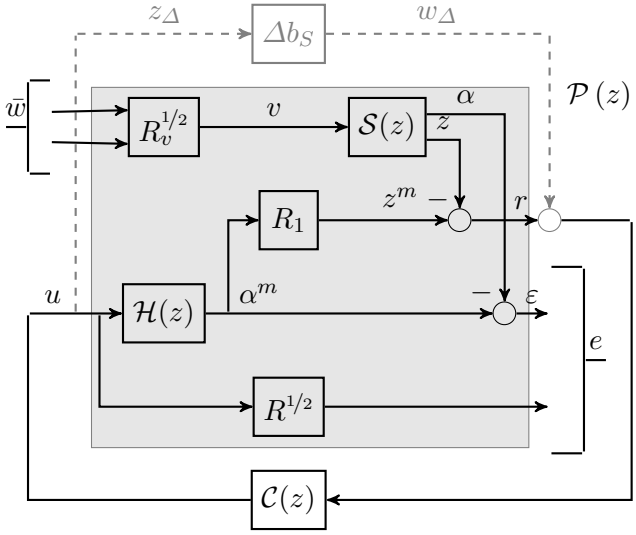

Fig. 1: Block diagram of the control system

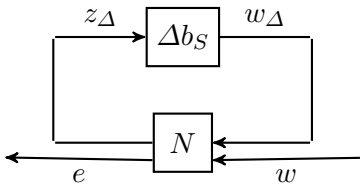

Fig. 2: Uncertain closed-loop interconnection

called generalized plant framework is given in Fig. 1. The generalized plant $\mathcal{P}(z)$ (the shaded block), makes a distinction between exogenous zero-mean white noise inputs $\bar{w}_k$ and control

inputs u_k on one hand and the measurement outputs r_k and performance outputs e_k on the other hand. The white noise v_k does not have unit covariance, so we need to augment the atmospheric disturbance model with a static matrix $R_v^{1/2}$ multiplication to normalize the input covariance. α^m is the applied phase correction, $\varepsilon_k = \alpha_k - \alpha_k^m$ is the residual phase error, and the reduced signal is $r_k = z_k - z_k^m$, where z_k and z_k^m denote the contributions due to turbulence and the DM. We consider the case when the DM dynamics consist of a pure delay, i.e., $\mathcal{H}(z) = z^{-1} B_{\text{act}}$.

In the following, we design an optimal controller that minimizes the cost function $J = \mathcal{E}\{\varepsilon_k^T \varepsilon_k\} + \mathcal{E}\{u_k^T R u_k\}$, where $R \geq 0$ is a matrix which makes a trade off between the mean-square residual phase error and the amount of control effort. This is equivalent to finding a controller $\mathcal{C}(z)$ that minimizes the $\mathcal{H}_2$-norm of the closed loop transfer function $S(\mathcal{P}, \mathcal{C}) = \mathcal{P}_{11} + \mathcal{P}_{12}\mathcal{C}(I - \mathcal{P}_{22}\mathcal{C})^{-1}\mathcal{P}_{21}$, where the four blocks of the generalized plant $\mathcal{P}$ are given by

$$\left[\begin{array}{c} \frac{\xi_{k+1}}{e_k} \\ \hdashline r_k \end{array}\right] = \left[\begin{array}{cc:c:c} 0 & 0 & 0 & I \\ 0 & A_{\text{tur}} & K_k R_v^{1/2} & 0 \\ \hdashline -I & I & 0 & 0 \\ 0 & 0 & 0 & R^{1/2} \\ \hdashline -R_1 & R_1 & R_v^{1/2} & 0 \end{array}\right] \left[\begin{array}{c} \frac{\xi_k}{\bar{w}_k} \\ \hdashline u_k \end{array}\right]. \tag{16}$$

A simplified version of Corollary 1 in [12], adapted to our framework, gives an analytic formulation of the optimal controller for the problem mentioned above.

Theorem 1 *Let α_k and z_k be characterized by the stochastic process in Eq. (15) with $R_v > 0$. Assume that the DM can be modeled as a pure delay and that B_{act} or R has full column rank. The optimal feedback $\mathcal{C}(z)$ that internally stabilizes $\mathcal{P}(z)$ has a state-space representation*

$$\left[\begin{array}{c} \frac{\hat{\xi}_{k+1}}{u_k} \end{array}\right] = \left[\begin{array}{c:c} \tilde{A} + K_k R_1 F & K_k \\ \hdashline F\left(\tilde{A} + K_k R_1 F\right) & F K_k \end{array}\right] \left[\begin{array}{c} \frac{\hat{\xi}_k}{r_k} \end{array}\right], \tag{17}$$

where $\tilde{A} = A_{\text{tur}} - K_k R_1$, $F = H_R^\dagger$, and $H_R^\dagger = (I+R)^{-1}$. The controller output is $u_k = H_R^\dagger \hat{\alpha}_{k+1|k}$.

The controller computed in Theorem 1 stabilizes the nominal generalized plant $\mathcal{P}(z)$, but as seen from Eq. (12), $\mathcal{P}(z)$ is affected by an additive uncertainty $\Delta b_S(\alpha_k)$. We further investigate under which conditions the controller in Eq. (17) also stabilizes the uncertain system. For this, we look at the influence which the uncertainty can exert on the interconnection. The transfer function seen by Δb_S is nothing but the transfer function $w_\Delta \to z_\Delta$, $M(z) = (I + C(z) R_1 \mathcal{H}(z))^{-1} C(z)$. Schematically, the system including the uncertainty is given in Fig. 2, where $N = S(\mathcal{P}, \mathcal{C}) = \begin{bmatrix} N_{11} N_{12} \\ N_{21} N_{22} \end{bmatrix}$. The transfer matrix seen by Δb_S is $N_{11} = M$. The set of uncertainties is $\mathbf{\Delta} := \{\Delta b_S \in \mathbb{R}^{m^2 \times 1} | \Delta b_S \in \mathbf{\Delta}_\star\}$, where $\mathbf{\Delta}_\star$ is a value set of real matrices that defines the structure and the size of the uncertainties. We now state two theorems [13] which determine under which conditions the closed loop system remains stable for $\Delta b_S \neq 0$.

Theorem 2 *If $\mathcal{C}$ stabilizes $\mathcal{P}$, and if $I - M\Delta b_S$ has a proper and stable inverse for all $\Delta b_S \in \mathbf{\Delta}$, then $\mathcal{C}$ robustly stabilizes $S(\Delta b_S, \mathcal{P})$ against $\mathbf{\Delta}$.*

Theorem 3 (Small gain theorem) *If any $\Delta b_S^\star \in \mathbf{\Delta}_\star$ satisfies $\|\Delta b_S^\star\|_\infty < r$, and if $\|M\|_\infty \leq \frac{1}{r}$, then $I - M\Delta b_S$ has a proper and stable inverse for all $\Delta b_S \in \mathbf{\Delta}$.*

From Theorems 3 and 2, $\|M\|_\infty \leq \frac{1}{r}$ is a sufficient condition for $\mathcal{C}$ to robustly stabilize $S(\Delta b_S, \mathcal{P})$.

6 Simulations

Static aberrations. We present numerical simulations for the aberration correction problem using M_1, M_2 in Eq. (11), and the method presented in [5] (M_3). For M_1, we take 2 images at

defocus -2 and 2 rad and precompute the coefficients in Eq. (5). The coefficients in Eq. (6) are independent of the chosen diversity and can also be precomputed. M_3 assumes that the DM itself introduces the phase diversity, namely the correction, that minimizes the wavefront error and also minimizes the image acquiring effort. We consider a pupil of radius $r = 32$. The static aberration (and the initial state in the dynamic case) consisting of $n = 14$ modes is generated using the Kolmogorov model [10]. We show the convergence of the algorithms (6 correction steps) for a wavefront rms of 1 rad and two noise realizations, with SNR ∞ and 10 dB in Figs. 3, and 4, respectively. Because M_1 and M_2 use 2 images per iteration and M_3 uses only 1 image starting with the second iteration, the images corresponding to the error in M_3 are shown by skipping one intermediate step. For M_2, we choose a diversity that in simulation minimizes the residual error. M_1 has a robust performance using the chosen fixed diversity.

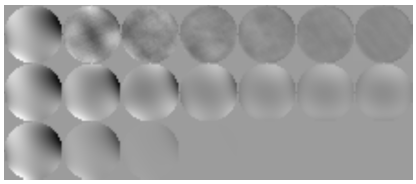

Fig. 3: Wavefront error: 1 rad rms, no noise - M_3 (top), M_2 (middle), M_1 (bottom)

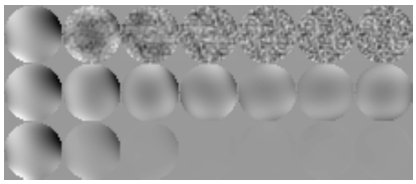

Fig. 4: Wavefront error: 1 rad rms, 10 dB SNR - M_3 (top), M_2 (middle), M_1 (bottom)

Repeating the previous experiment 1024 times for random aberrations of 1 rad rms and no noise, we plot in Fig. (5) the residual error in the aberration vector at each correction step versus the number of images. Plotted are the median, first and third quartile, and the minimal and maximal values. M_1 converges to ~ 0 residual error. M_2 and M_3, as expected, converge to a constant value not equal zero (~ 0.2). In Fig. 6 we plot, for the same data, the relative residue. The error bound depends on the model error for well conditioned matrices (the condition numbers of A_S and A_D are 4.18 and 12.70, respectively). M_1 converges to zero residual error for different realizations of $\Delta b_S(\alpha_k)$. M_2 converges to a residual error not equal to zero and depends on the realizations of $\Delta b_D(\alpha_k)$. M_3 also converges to a residual error not equal to zero.

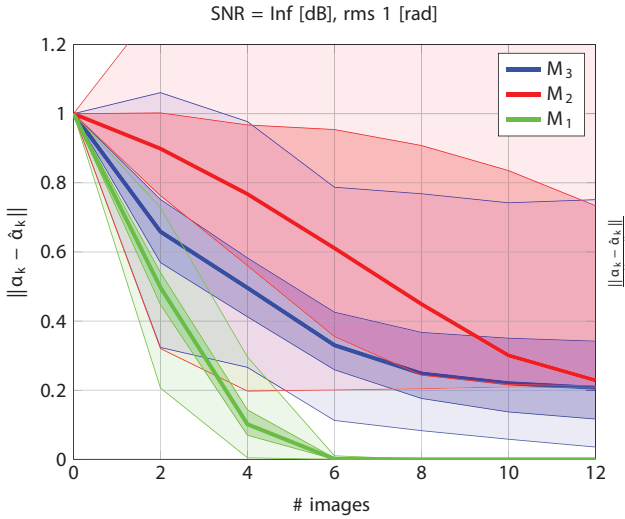

Fig. 5: Residual error

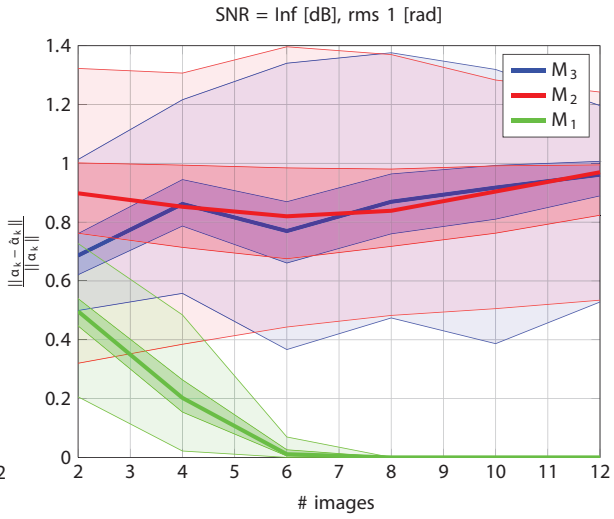

Fig. 6: Relative residue

Next, we perform a Monte Carlo simulation for increasing SNR and rms in Figs. 7 and 8, respectively. For each SNR, we repeat in Fig. 7 the experiment 1024 times. The initial aberration

has 1 rad rms. For SNR between -5 and 10 dB, the relative residue corresponding to M_2 and M_3 is sometimes larger then 1, which means that the rms value of the wavefront increases after one iteration. For increasing SNR, all three methods have a relative residue smaller than 1, although of different values from which the lowest can be observed for M_1. Besides the fact that M_2 gives very different results with respect to the chosen diversity, the estimates also have a larger variance than the other two methods and this is due to the decrease in SNR by subtracting images. The same type of analysis is made in Fig. 8 for increasing rms and an SNR of 10 dB. All three methods are based on either a small-phase or small-aberration assumption, so the bias of the estimation increases with increasing rms. Overall, M_1 has the smallest relative residue.

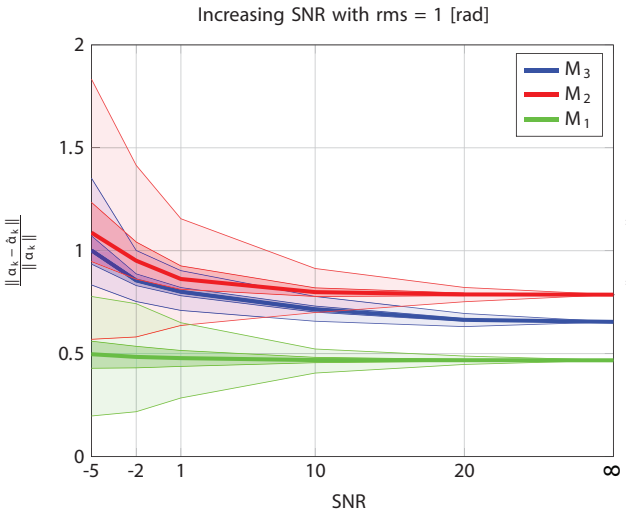

Fig. 7: Relative residue vs SNR

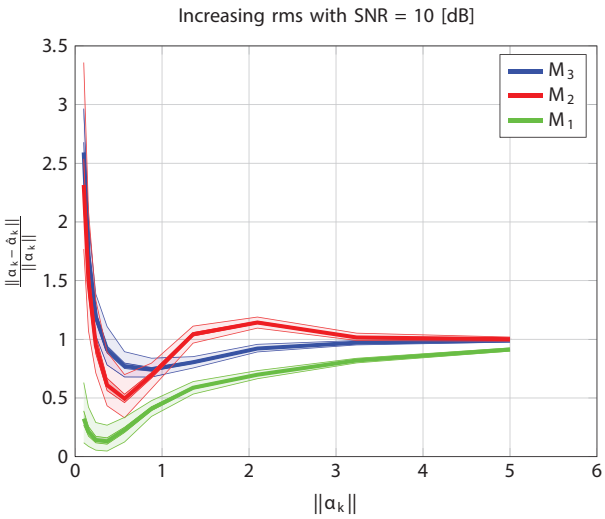

Fig. 8: Relative residue vs rms

Dynamic aberrations. We choose a turbulence model using telescope diameter $D = 1[\mathrm{m}]$, outer scale $L_0 = 42[\mathrm{m}]$, Fried parameter $r_0 = 0.4[\mathrm{m}]$, wind velocity $10[\mathrm{m/s}]$, sampling time $0.01[\mathrm{s}]$ and we perform a Monte Carlo simulation over 128 turbulence realizations. The turbulence model matrices in Eq. (12) are derived as in [14]. We investigate the performance of the optimal $\mathcal{H}_2$ controller in Theorem 1 with $R{=}I$. In Fig. 9, the aberration is significantly reduced to $\sim 1/7$ from an initial state of $\sim 3/2$ rad. In Fig. 10, we investigate the robust stability of the system in Eq. (15) as a function of turbulence strength. For each value of r_0, we run 128 simulations and compute the maximum value of $\|\Delta b_S\|_\infty$ and of $\|M\|_\infty$. If we choose r in Theorem 3 as $\max \|\Delta b_S\|_\infty$ over a simulated time interval we ensure that $\|\Delta b_S\|_\infty \leq r$. In Fig. 10 we plot $\max\left(\|\Delta b_S\|_\infty\right) \max\left(\|M\|_\infty\right)$ for increasing r_0. Stability is ensured if the quantity plotted in Fig. 10 is smaller than 1.

7 Conclusions

We have presented a method for wavefront estimation and correction. For a point-like object, we have used a least squares approach for aberration estimation. The linearization of the PSF used is valid for small phase aberrations and a general diversity, which is the typical situation in a control loop. We have also adapted the algorithm to the dynamic turbulence model, and have corrected for the aberration with an H_2 optimal controller.

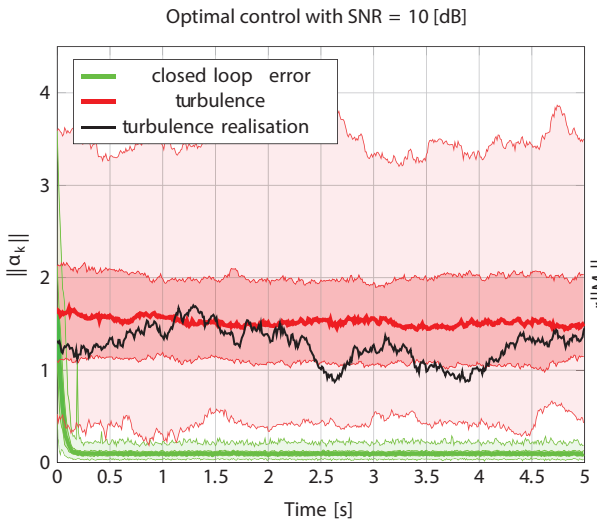

Fig. 9: Residual error in closed loop

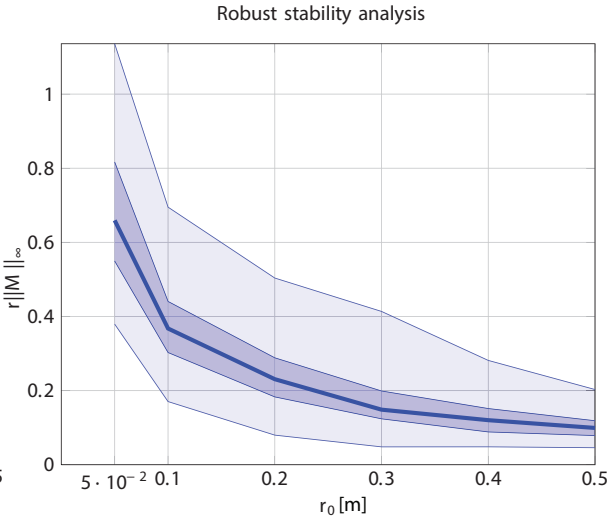

Fig. 10: Robust stability analysis

References

1. L. M. Mugnier, A. Blanc, and J. Idier, Phase diversity: a technique for wave-front sensing and for diffraction-limited imaging, *Adv. Imag. Electr. Phys.* 141, 1–76, 2006.
2. D. R. Gerwe, M. M. Johnson, and B. Calef, Local minima analysis of phase diverse phase retrieval using maximum likelihood, In *AMOS Technical Conference, Kihei HI*, 2008.
3. R. A. Gonsalves, Small-phase solution to the phase-retrieval problem, *Opt. Lett.* 26(10), 684–685, 2001.
4. S. Meimon, T. Fusco, and L. M. Mugnier, Lift: a focal-plane wavefront sensor for real-time low-order sensing on faint sources, *Opt. Lett.* 35(18), 3036–3038, 2010.
5. C. U. Keller, V. Korkiakoski, N. Doelman, R. Fraanje, R. Andrei, and M. Verhaegen, Extremely fast focal-plane wavefront sensing for extreme adaptive optics, In *Proc. SPIE 8447, Adaptive Optics Systems III, 844721*, 2012.
6. W. J. Wild, Linear phase retrieval for wave-front sensing, *Opt. Lett.* 23(8), 573–575, 1998.
7. I. Mocœur, L. M. Mugnier, and F. Cassaing, Analytical solution to the phase-diversity problem for real-time wavefront sensing, *Opt. Lett.* 34(22):3487–3489, 2009.
8. D. J. Lee, M. C. Roggemann, and B. M. Welsh, Cramer-rao analysis of phase-diverse wavefront sensing, *J. Opt. Soc. Am. A* 16(5), 1005–1015, 1999.
9. Joseph W. Goodman, *Introduction To Fourier Optics*, Roberts & Company Publishers, USA, 2005.
10. R.J. Noll, Zernike polynomials and atmospheric turbulence, *J. Opt. Soc. Am. A* 66(3), 207–211, 1976.
11. B. D. O. Anderson and J. B. Moore, *Optimal filtering*, volume 11, Prentice-hall Englewood Cliffs, NJ, 1979.
12. K. Hinnen, M. Verhaegen, and N. Doelman, A data-driven H_2-optimal control approach for adaptive optics, *IEEE T. Contr. Syst. T.* 16(3), 381–395, 2008.
13. K. Zhou, J. C. Doyle, K. Glover, *Robust and optimal control*, volume 40, Prentice Hall Upper Saddle River, NJ, 1996.
14. C. Kulcsár, H. F. Raynaud, C. Petit, and J. M. Conan, Minimum variance prediction and control for adaptive optics, *Automatica*, 2012.