Experimental validation of LIFT for estimation of low-order modes in low-flux wavefront sensing

C. Plantet, S. Meimon, J. -M. Conan, T. Fusco

Experimental validation of LIFT for estimation of low-order modes in low-flux wavefront sensing

C. Plantet,[1*] S.Meimon,[1] J.-M. Conan[1] and T. Fusco[1,2]

[1]*ONERA, The French Aerospace Lab*
92322 Chatillon, France
[2]*Aix Marseille Universite, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326*
13388 Marseille, France
[*]*cedric.plantet@onera.fr*

Abstract: Laser Tomographic and Multi-Conjugate Adaptive Optics systems rely on natural guide stars to sense low order aberrations (tip/tilt and focus). LIFT is a novel focal plane wavefront sensor (WFS), performing a maximum likelihood phase retrieval on a single image, with better sensitivity than a 2×2 Hartmann-Shack WFS. Its performance for the estimation of tip/tilt and focus is similar to a pyramid WFS without modulation, but with a simpler set-up. We present here the LIFT concept and associated data processing, as well as experimental results. We validate the estimation of tip/tilt and focus, with monochromatic and large bandwidth light, and verify the low noise sensitivity predicted by theory.

OCIS codes: (010.1080) Active or adaptive optics; (010.7350) Wave-front sensing.

References and links

1. T. Fusco, S. Meimon, Y. Clenet, M. Cohen, H. Schnetler, J. Paufique, V. Michau, J.-P. Amans, D. Gratadour, C. Petit, C. Robert, P. Jagourel, E. Gendron, G. Rousset, J.-M. Conan and N. Hubin, "ATLAS: the E-ELT laser tomographic adaptive optics system," Proc. SPIE **7736** 77360D (2010).
2. E. Diolaiti, J.-M. Conan, I. Foppiani, M. Lombini, E. Marchetti, C. Petit, C. Robert, L. Schreiber, A. Baruffolo, M. Bellazzini, G. Bregoli, P. Ciliegi, G. Cosentino, V. Biliotti, S. D'Odorico, T. Fusco, N. Hubin, S. Meimon and J.-F. Sauvage, "MAORY E-ELT multi-conjugate adaptive optics module," Proc. AO4ELT 02007 (2011).
3. G. Herriot, P. Hickson, B. L. Ellerbroek, D. A. Andersen, T. Davidge, D. A. Erickson, I. P. Powell, R. Clare, L. Gilles, C. Boyer, M. Smith, L. Saddlemyer and J.-P. Veran, "NFIRAOS: TMT narrow field, near-infrared facility adaptive optics," Proc. SPIE **6272** 62720Q (2006).
4. B. Neichel, F. Rigaut, M. Bec, M. Boccas, F. Daruich, C. D'Orgeville, V. Fesquet, R. Galvez, A. Garcia-Rissmann, G. Gausachs, M. Lombini, G. Perez, G. Trancho, V. Upadhya and T. Vucina, "The Gemini MCAO System GeMS: nearing the end of a lab-story," Proc. SPIE **7736** 773606-1 (2010).
5. G. Rousset, "Wave-front sensors," in *Adaptive Optics in Astronomy*, F. Roddier, (Cambridge University, 1999) 91–130.
6. F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," Appl. Opt. **27**(7), 1223–1225 (1988).
7. O. Guyon, "Limits of adaptive optics for high contrast imaging," Astrophys. J. **629**, 592–614 (2005).
8. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," J. Mod. Opt. **43**(2), 289–293 (1996).
9. R. Ragazzoni and J. Farinato, "Sensitivity of a pyramidic wave front sensor in closed loop adaptive optics," Astron. Astrophys. **350**, L23–L26 (1999).
10. C. Verinaud, "On the nature of the measurements provided by a pyramid wave-front sensor," Opt. Commun. **233**, 27–38 (2004).
11. R. Dekany, Caltech Optical Observatories, Physics, Math and Astronomy Division, California Institute of Technology, Pasadena, Calif. 91125, USA (personal communication).

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

12. C. Paterson and J. C. Dainty, "Hybrid curvature and gradient wave-front sensor," Opt. Lett. **25**(23), 1687–1689 (2000).
13. S. Barwick, "Performance comparison between Shack-Hartmann and astigmatic hybrid wavefront sensors," Appl. Opt. **48**(36), 6967–6972 (2009).
14. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829–832 (1982).
15. I. Mocoeur, 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).
16. 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).
17. C. Plantet, S. Meimon, T. Fusco, and J.-M. Conan, "Experimental validation of the linearized focal-plane technique (LIFT)," Proc. AO4ELT2 (2011).
18. R. A. Gonsalves, "Small-phase solution to the phase-retrieval problem," Opt. Lett. **26**(10), 684–685 (2001).
19. L.A. Poyneer and B. Macintosh, "Spatially filtered wave-front sensor for high-order adaptive optics," J. Opt. Soc. Am. **21**, 810–819 (2004).
20. S. Meimon, T. Fusco and F. Cassaing, "A focal plane sensor for low-order sensing on laser tomographic systems: LIFT," Proc. SPIE **7736** 773611 (2010).
21. D. K. Cohen, W. H. Gee, M. Ludeke and J. Lewkowicz, "Automatic focus control: the astigmatic lens approach," J. Opt. Soc. Am. **23**, 565–570 (1984).
22. M. Nicolle, T. Fusco, G. Rousset, and V. Michau, "Improvement of Shack-Hartmann wave-front sensor measurement for extreme adaptive optics," Opt. Lett. **29**(18), 2743–2745 (2004).
23. S. Thomas, T. Fusco, A. Tokovinin, M. Nicolle, V. Michau and G. Rousset, "Comparison of centroid computation algorithms in a Shack-Hartmann sensor," Mon. Not. R. Astron. Soc. **371**, 323–336 (2006).
24. P. Feautrier, J.-L. Gach, M. Downing, P. Jorden, J. Kolb, J. Rothman, T. Fusco, P. Balard, E. Stadler, C. Guillaume, D. Boutolleau, G. Destefanis, N. Lhermet, O. Pacaud, M. Vuillermet, A. Kerlain, N. Hubin, J. Reyes, M. Kasper, O. Ivert, W. Suske, A. Walker, M. Skegg, S. Derelle, J. Deschamps, C. Robert, N. Vedrenne, F. Chazalet, J. Tanchon, T. Trollier, A. Ravex, G. Zins, P. Kern, T. Moulin and O. Preis, "Advances in detector technologies for visible and infrared wavefront sensing," Proc. SPIE **8447**, 84470Q (2012).
25. C. Plantet, B. Neichel, S. Meimon, T. Fusco, and J.-M. Conan, "LIFT, a noise-effective low order focal-plane sensor: from theory to full experimental validation," Proc. SPIE (2012).

1. Introduction

Tomographic adaptive optics uses laser guide stars to build a 3D model of the turbulence. Nevertheless, natural guide stars are still needed to sense low order modes [1–4]. To get a maximum sky coverage, we need to sense stars as faint as possible. Hence, wavefront sensors must have a very low sensitivity to noise. Several wavefront sensors are already used or could be considered for the analysis of natural stars. We can cite the Shack-Hartmann (SH), curvature, pyramid, modified quad-cell sensors and phase diversity.

The Shack-Hartmann has been used and studied a lot, making it a reliable and widespread sensor. It is a pupil-plane sensor using microlenslets dividing the pupil into subapertures [5]. The local wavefront slope is measured in each subaperture to retrieve the wavefront surface. The signal to noise ratio is inversely proportional to the number of subapertures, since the total flux is separated among them. As it needs at least 4 subapertures to estimate the focus, it may not be effective enough for the faintest stars.

The curvature sensor, introduced by Roddier in 1988 [6], measures second derivatives of the wavefront instead of slopes. It has been shown that it is more sensitive to noise in low spatial frequencies than a Shack-Hartmann [7]. It is thus not adapted to the estimation of low order modes on natural guide stars.

Ragazzoni proposed in 1996 a new pupil-plane sensor using a pyramidal prism [8]. This sensor has been studied in terms of photon noise, and has better performance than a Shack-Hartmann [7, 9, 10]. This sensor could be a competitive one in the search for the best possible sky coverage. Nevertheless, it requires a complex setup and the pyramidal prism comes as an additional component.

On the other side of the WFS spectrum, we find focal-plane sensors. A simple and attractive solution, the modified quad-cell, has been proposed by Dekany for the Keck telescope [11]. It

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

consists in adding an astigmatism offset before focalization on a 4 pixel detector. This method allows the estimation of a focus in addition of the classical tip/tilt. Nevertheless, this sensor has poor linearity and is strongly sensitive to aliasing. Another concept, based on the modified quad-cell and the curvature sensor has been developed by Paterson and Dainty [12]. Unfortunately, this sensor seems to lose efficiency at low flux and does not outperform a Shack-Hartmann [13].

Phase diversity is another focal-plane sensor relying on two images [14]. One is taken at the focal plane and one with a focus offset. An algorithm deduces the aberrations from these images. Mocoeur *et al.* [15] proposed a method to make this sensor fitted for real-time analysis, and showed good results for small phase estimations of low-order aberrations. Though, the flux is divided between the two images, which makes the set-up a bit complex and potentially reduces the detection efficiency.

We recently proposed a wavefront sensor called LIFT (Linearized Focal-plane Technique), dedicated to faint natural guide stars [16], and requiring a very simple hardware. The estimation by LIFT is based upon a single image taken at the focal plane of the telescope. The relation between the aberration and the intensity pattern is linearized to make computations easier and faster. Besides, we add a known phase offset to avoid phase indeterminations. First experiments on ONERA's test bench have already helped validating this sensor [17]. The present article gives a detailed description of LIFT concept, as well as a theoretical analysis of its linearity and noise properties. The outcome of this analysis is compared with experimental results. The noise propagation of the Shack-Hartmann and pyramid sensors is also studied for comparison.

In section 2, we focus on the concept of LIFT and detail its algorithm. We then discuss in section 3 the optimization of the phase offset by means of analytical calculations on the convergence of estimation and noise propagation. This last analysis is used to compare LIFT with the Shack-Hartmann and pyramid sensors in section 4. We finally present in section 5 the full in-lab experimental validation of LIFT for the estimation of low order modes: tip, tilt and focus.

2. LIFT : a phase-retrieval wavefront sensor

LIFT is a focal plane WFS, comprising a focusing optical system introducing a known amount of astigmatism, a focal plane detection, and an original phase retrieval algorithm. LIFT, as a natural guide star low order WFS, is meant to be used in infrared (IR) light with a high order correction on laser guide stars. After a description of data formation, we go into the details of this algorithm.

2.1. Data formation

Let r be the vector of coordinates in the focal plane, ρ the vector of coordinates in the pupil plane, and λ the wavelength. The intensity pattern I on the imaging sensor is described by:

$$\begin{aligned} I(r,\phi+\phi_{off}) &= \int_\lambda F(\lambda)|\mathrm{FT}\{\underbrace{P(\rho)\exp[i\phi_{off}(\rho,\lambda)]}_{P_{off}}\times\exp[i\phi(\rho,\lambda)]\}_r|^2\mathrm{d}\lambda \qquad (1)\\ &= \int_\lambda F(\lambda)I_\lambda(r,\phi+\phi_{off})\mathrm{d}\lambda \end{aligned}$$

with $P(\rho)$ the pupil transmission, $F(\lambda)$ the incoming flux, ϕ the aberrated phase to be estimated and ϕ_{off} the known phase offset. $\mathrm{FT}\{f(\rho)\}_r$ is the Fourier Transform from the pupil plane to the focal plane as described by Fraunhofer diffraction. $I_\lambda(r,\phi+\phi_{off})$ is the intensity pattern at a given λ. The integral can be considered as a discrete sum over the wavelengths:

$$I(r,\phi+\phi_{off})=\sum_\lambda F_\lambda I_\lambda(r,\phi+\phi_{off}) \qquad (2)$$

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

ϕ can be decomposed on Zernike modes Z_i so that $\phi = \sum_i a_i Z_i$. Let $\mathbf{A}$ be the vector of coefficients a_i. We assume that the aberrations, in optical path difference ($\mathbf{A_{OPD}} = \mathbf{A} \times \lambda/2\pi$), do not depend on the wavelength.

For the needs of numerical computations, we consider now the intensity pattern as a vector comprising all the pixels of the imaging detector. It can then be expressed as a function of $\mathbf{A_{OPD}}$ only. Equation (2) becomes:

$$\mathbf{I}(\mathbf{A_{OPD}} + \mathbf{A_{off}})[p] = \sum_\lambda F_\lambda \mathbf{I}_\lambda(\mathbf{A_{OPD}} + \mathbf{A_{off}})[p] + \mathbf{n}[p] \tag{3}$$

with $\mathbf{I}(\mathbf{A_{off}})[p]$ corresponding to the integration on the p-th pixel of $I(r, \phi_{off})$. $\mathbf{n}$ is the noise on pixels. We use a first order Taylor expansion to linearize Eq. (3), as described by Gonsalves for small phase estimation [18]:

$$\sum_\lambda F_\lambda \mathbf{I}_\lambda(\mathbf{A_{OPD}} + \mathbf{A_{off}}) \simeq \sum_\lambda F_\lambda [\mathbf{I}_\lambda(\mathbf{A_{off}}) + \sum_k a_k \mathbf{I}'_{\mathbf{k}_\lambda}(\mathbf{A_{off}})] + \mathbf{n} \tag{4}$$

which is equivalent to:

$$\sum_\lambda F_\lambda [\mathbf{I}_\lambda(\mathbf{A_{OPD}} + \mathbf{A_{off}}) - \mathbf{I}_\lambda(\mathbf{A_{off}})] \simeq \sum_\lambda \sum_k a_k F_\lambda \mathbf{I}'_{\mathbf{k}_\lambda}(\mathbf{A_{off}}) + \mathbf{n} \tag{5}$$

with $\mathbf{I}'_{\mathbf{k}_\lambda}(\mathbf{A_{off}}) = \frac{\partial \mathbf{I}_\lambda(\mathbf{A_{OPD}} + \mathbf{A_{off}})}{\partial a_k}\Big|_{\mathbf{A_{OPD}} = \mathbf{0}}$.

Let $\Delta\mathbf{I}_\lambda$ be the vector so that $\Delta\mathbf{I}_\lambda[p] = \mathbf{I}_\lambda(\mathbf{A_{OPD}} + \mathbf{A_{off}})[p] - \mathbf{I}_\lambda(\mathbf{A_{off}})[p]$, and $\mathbf{H}_\lambda(\mathbf{A_{off}})$ be the matrix of which the element at the p-th row and k-th column is defined by $\mathbf{H}_\lambda(\mathbf{A_{off}})[p,k] = \mathbf{I}'_{\mathbf{k}_\lambda}(\mathbf{A_{off}})[p]$. Then Eq. (5) becomes:

$$\sum_\lambda F_\lambda \Delta\mathbf{I}_\lambda = \Big(\sum_\lambda F_\lambda \mathbf{H}_\lambda(\mathbf{A_{off}})\Big)\mathbf{A_{OPD}} + \mathbf{n} \tag{6}$$

By analogy with Eq. (3), we define $\Delta\mathbf{I} = \sum_\lambda F_\lambda \Delta\mathbf{I}_\lambda$ and $\mathbf{H}(\mathbf{A_{off}}) = \sum_\lambda F_\lambda \mathbf{H}_\lambda(\mathbf{A_{off}})$. The previous relation writes:

$$\Delta\mathbf{I} = \mathbf{H}(\mathbf{A_{off}})\mathbf{A_{OPD}} + \mathbf{n} \tag{7}$$

The estimation of $\mathbf{A_{OPD}}$ is made by maximum likelihood (ML). If the optical bandwidth is narrow enough, we can consider the central wavelength only:

$$\Delta\mathbf{I}_{\bar{\lambda}} = \mathbf{H}_{\bar{\lambda}}(\mathbf{A_{off}})\mathbf{A_{OPD}} + \mathbf{n} \tag{8}$$

with $\bar{\lambda} = (\sum_\lambda F_\lambda \lambda)/(\sum_\lambda F_\lambda)$. In practice and in this article, we always use a monochromatic model at the central wavelength, but the polychromatic model could be easily implemented. The accuracy of estimation with this approximation in polychromatic light is evaluated in section 5.2.2.

LIFT is meant to sense low orders (tip/tilt/focus) in a context where higher orders are corrected based on laser guide star measurements. The images to be analyzed should therefore present a low amount of high orders. Besides, in order to limit aliasing and by analogy with the pinhole filtering method in a Shack-Hartmann WFS [19], only a limited square zone (~ 5 Airy diameters) centered on the image spot is taken into account. This lowers computation time as well. Also, LIFT offers the possibility to estimate a few modes beyond tip/tilt/focus, which can help reduce aliasing [20]. As a first approximation, aliasing effects are assumed negligible, and consequently higher modes are set to zero in the present paper. A detailed study of LIFT's behavior in the presence of high orders will come in a future publication.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

2.2. Algorithm

We approximate the k-th pixel photon noise by an additive zero-mean Gaussian noise, with a variance equal to the mean flux $\bar{I_k}$. Read-out noise follows a zero-mean Gaussian distribution with a standard deviation σ_e considered uniform on the detector, hence the global noise variance for the pixel k:

$$\sigma_{n_k}^2 = \bar{I_k} + \sigma_e^2 \tag{9}$$

This expression defines the elements of the covariance matrix $\mathbf{R_n} = \langle \mathbf{n}\mathbf{n}^t \rangle$, which is assumed to be diagonal. The Gaussian approximation for photon noise is accurate with at least 10 photons per pixel. Under 10 photons, the read-out noise is dominant in IR imaging.

We search $\mathbf{A_{OPD}}$ so that the likelihood $p(\mathbf{I}|\mathbf{A_{OPD}}, \mathbf{H}(\mathbf{A_{off}}), \mathbf{I}(\mathbf{A_{off}}))$ is maximum. Based on the assumed noise statistics, the log likelihood takes the form:

$$-\ln\left[p(\mathbf{I}|\mathbf{A_{OPD}}, \mathbf{H}(\mathbf{A_{off}}), \mathbf{I}(\mathbf{A_{off}}))\right] = \frac{1}{2}(\Delta\mathbf{I} - \mathbf{H}(\mathbf{A_{off}})\mathbf{A_{OPD}})^t \mathbf{R_n}^{-1}(\Delta\mathbf{I} - \mathbf{H}(\mathbf{A_{off}})\mathbf{A_{OPD}}) + \text{cst} \tag{10}$$

The maximum likelihood solution is thus:

$$\hat{\mathbf{A}}_{\mathbf{OPD}} = \mathbf{P_{ML}}(\mathbf{A_{off}})\Delta\mathbf{I} \quad \text{with} \quad \mathbf{P_{ML}}(\mathbf{A_{off}}) = (\mathbf{H}(\mathbf{A_{off}})^t \mathbf{R_n}^{-1} \mathbf{H}(\mathbf{A_{off}}))^{-1} \mathbf{H}(\mathbf{A_{off}})^t \mathbf{R_n}^{-1} \tag{11}$$

and the variance of the estimation error $\mathbf{E} = \mathbf{A_{OPD}} - \hat{\mathbf{A}}_{\mathbf{OPD}}$ is given by:

$$\sigma_{error}^2 = Tr\{\langle \mathbf{E}\mathbf{E}^t \rangle\} = Tr\{(\mathbf{H}(\mathbf{A_{off}})^t \mathbf{R_n}^{-1} \mathbf{H}(\mathbf{A_{off}}))^{-1}\} \tag{12}$$

Each diagonal element of $\langle \mathbf{E}\mathbf{E}^t \rangle$ gives the variance of the estimation error for one mode.

Of course, first order Taylor expansion as seen in Eq. (4) has a limited validity range, which defines the sensor linearity domain. However, this domain can be extended by iterating the estimation. We use the following algorithm:

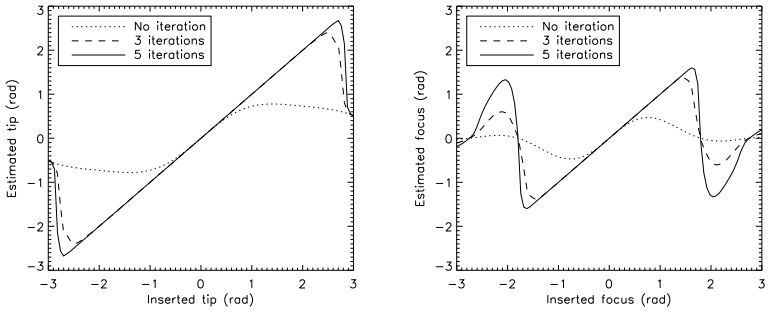
- Iteration 0: $\Delta\mathbf{I} = \mathbf{I_{data}} - \mathbf{I}(\mathbf{A_{off}})$; $\hat{\mathbf{A}}^0_{\mathbf{OPD}} = \mathbf{P_{ML}}(\mathbf{A_{off}})\Delta\mathbf{I}$
- Iteration 1: $\Delta\mathbf{I} = \mathbf{I_{data}} - \mathbf{I}(\hat{\mathbf{A}}^{\mathbf{0}}_{\mathbf{OPD}} + \mathbf{A_{off}})$; $\hat{\mathbf{A}}^1_{\mathbf{OPD}} = \mathbf{P_{ML}}(\hat{\mathbf{A}}^{\mathbf{0}}_{\mathbf{OPD}} + \mathbf{A_{off}})\Delta\mathbf{I} + \hat{\mathbf{A}}^0_{\mathbf{OPD}}$
- Iteration 2: $\Delta\mathbf{I} = \mathbf{I_{data}} - \mathbf{I}(\hat{\mathbf{A}}^{\mathbf{1}}_{\mathbf{OPD}} + \mathbf{A_{off}})$; $\hat{\mathbf{A}}^2_{\mathbf{OPD}} = \mathbf{P_{ML}}(\hat{\mathbf{A}}^{\mathbf{1}}_{\mathbf{OPD}} + \mathbf{A_{off}})\Delta\mathbf{I} + \hat{\mathbf{A}}^1_{\mathbf{OPD}}$

$$\vdots$$

with $\mathbf{I_{data}}$ being the image recorded on the detector and $\mathbf{I}(\mathbf{X})$ a simulated image from our model.

The extension of the linearity domain with the number of iterations is shown on Fig. 1 for tip/tilt and focus. A few iterations are sufficient to improve linearity.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013



(a) Tip (b) Focus

Fig. 1. Simulation of the linearity domain with different numbers of iterations for tip and focus. Phase offset: 0.4 rad of astigmatism. Monochromatic light.

We may discuss the stopping criterion used to determine the number of iterations. Without the approximation made by linearization, the maximum likelihood estimator $\hat{\mathbf{A}}_{\mathbf{OPD}}$ verifies:

$$\hat{\mathbf{A}}_{\mathbf{OPD}}^{\mathbf{ML}} = \underset{\mathbf{A_{OPD}}}{\arg\min} \left([\mathbf{I_{data}} - \mathbf{I}(\mathbf{A_{OPD}} + \mathbf{A_{off}})]^t {\mathbf{R_n}}^{-1} [\mathbf{I_{data}} - \mathbf{I}(\mathbf{A_{OPD}} + \mathbf{A_{off}})] \right) \tag{13}$$

We thus chose to evaluate, at each iteration k, the criterion:

$$C_k = [\mathbf{I_{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\mathbf{OPD}}^{\mathbf{k-1}} + \mathbf{A_{off}})]^t {\mathbf{R_n}}^{-1} [\mathbf{I_{data}} - \mathbf{I}(\hat{\mathbf{A}}_{\mathbf{OPD}}^{\mathbf{k-1}} + \mathbf{A_{off}})] \tag{14}$$

The algorithm iterates until the relative difference $|C_k - C_{k-1}|/C_k$ is lesser than 10^{-6} or until the mean of the difference between two successive estimations $\langle ||\hat{\mathbf{A}}_{\mathbf{OPD}}^{k} - \hat{\mathbf{A}}_{\mathbf{OPD}}^{k-1}|| \rangle$ is lesser than 10^{-3} rad.

3. Algorithm tuning

In this section, we discuss the choice of the algorithm parameters. In particular, we choose the phase offset in order to have an efficient focus estimation. We studied two possibilities: focus and astigmatism. A theoretical analysis of the minimization criterion and of the noise propagation helps us choose the appropriate aberration as well as its amplitude. In a major part of this section, we discuss the estimation of focus. The values of aberrations are given in radians rms in all the following.

3.1. Choice of the sampling

Too much oversampling would lead to important noise propagation, as flux is divided into more pixels. We thus chose to use a Shannon sampling, i.e. the pixel size is $\lambda/2D$ with D the diameter of the pupil. However, sampling is a parameter that can be changed to fit any system. A Shannon/2 sampling could also be considered, but is not discussed in this paper.

3.2. Choice of the offset mode

Without any phase offset, the focus estimation is not unique: two values of opposite sign lead to the same focal image. Adding a known phase to the aberrated wavefront will help us remove this indetermination.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

A natural choice would be a focus offset, as used in phase diversity. But this only shifts the position of the focal plane. The symmetry with respect to the focused image is still present. For example, if we choose a 0.4 rad focus offset, a -0.3 rad or a -0.5 rad focus will lead to the exact same focal image with an absolute amount of focus equal to 0.1 rad. To understand better the algorithm behavior with respect to the inserted focus, we can analyze the structure of the criterion to be minimized (see Eq. (13)), which can be rewritten:

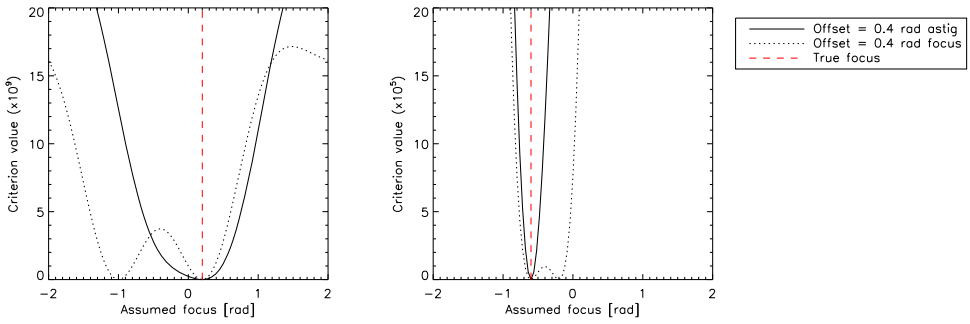
$$\mathbf{C}(\mathbf{A_{OPD}}) = \left([\mathbf{I_{data}} - \mathbf{I}(\mathbf{A_{OPD}} + \mathbf{A_{off}})]^t \mathbf{R_n}^{-1} [\mathbf{I_{data}} - \mathbf{I}(\mathbf{A_{OPD}} + \mathbf{A_{off}})]\right) \tag{15}$$

We recall that $\mathbf{A_{off}}$ denotes the phase offset, for now a focus offset. $\mathbf{A_{OPD}}$ is called the assumed focus, and the focus to be estimated is called true focus. Hence, $\mathbf{I_{data}}$ is formed by addition of the focus offset and the true focus.

At first iteration, the criterion takes the value at 0 rad of assumed focus. We call it the starting point, which corresponds to an image with the focus offset only $\mathbf{I}(\mathbf{A_{off}})$. During minimization, the criterion follows the curve toward lower values. The estimated focus is the assumed focus for which $\mathbf{C}(\mathbf{A_{OPD}})$ is minimum. If several minima co-exist, the estimated focus is at the minimum closest to the starting point.

Figure 2(a) plots in dotted line the criterion value depending on the assumed focus. The focus offset is 0.4 rad, and the true focus is 0.2 rad. The criterion should thus be minimum at 0.2 rad.

Two minima are visible on the curve. One is at 0.2 rad and the other at -1 rad. As expected, they are symmetrical with respect to -0.4 rad, where the true focus and the focus offset cancel each other out. In this case, as 0.2 rad is closer to the starting point, the focus estimation is correct. On the other hand, if the true focus is less than -0.4 rad, the closest minimum may be the wrong one. Figure 2(b) demonstrates that the estimated focus is -0.2 rad if we insert a -0.6 rad focus. Hence, adding a focus offset does not totally remove the indetermination.



(a) True focus = 0.2 rad (b) True focus = -0.6 rad

Fig. 2. Simulation of the minimization criterion value depending on the assumed focus. The curves are plotted for an offset phase at 0.4 rad, in focus or astigmatism. In the case of a focus offset, the estimated focus is 0.2 rad in (a) and -0.2 in (b). In the case of an astigmatism offset, the estimated focus is the true focus in (a) and (b).

Another intuitive choice is an astigmatism offset, since it gives particular shapes to the focal images when introducing a focus. The spot will stretch out in one direction with a positive focus, and in the perpendicular direction with a negative focus. Besides, astigmatism has already been used in focus control systems, and has shown to be effective [21]. If we compare it to a focus

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

offset in terms of minimization criterion, we observe only one minimum, at the true focus (Fig. 2). Hence, the astigmatism offset gives a better behavior which will lead to better linearity.

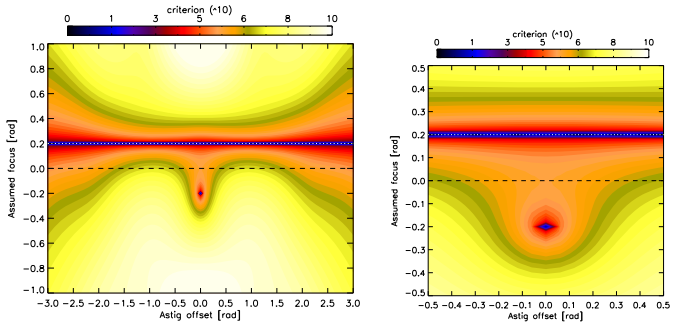
3.3. Choice of the offset amplitude

3.3.1. Impact on the minimization criterion

We have just seen that the criterion to be minimized has a better shape with an astigmatism offset than with a focus offset. But this shape actually depends on the amount of astigmatism introduced as a phase offset.

Figure 3(a) represents the criterion depending on the assumed focus and on the astigmatism offset. The true focus is still 0.2 rad. Each column corresponds to the criterion curve for a given astigmatism offset. For example, the column at the abscissa 0.4 rad gives the solid line in Fig. 2(a). The white dotted line at 0.2 rad of assumed focus is the criterion minimum that gives the true estimation for all astigmatism offsets, and the dashed line at 0 rad corresponds to the starting point. We confirm here that without any astigmatism, the estimation is undetermined, since we have two minima at 0.2 rad and -0.2 rad that are symmetrical with respect to the starting point.

This figure allows us to know the shape of the criterion for different astigmatism offsets. First of all, we checked that the offset has to be larger than 0.15 rad to avoid a minimum duality. The range 0.5-1.5 rad may seem attractive since the criterion is steeper, which should lead to better convergence of the algorithm. Though, a high astigmatism leads to a large spot, with less flux per pixel, and thus more sensitivity to noise. The following section shows how the noise propagation constrains the choice of the amplitude.



(a) Criterion (b) Zoom

Fig. 3. Simulation of the minimization criterion value depending on the assumed focus and on the astigmatism offset. Dashed line: starting point. White dotted line: true focus to estimate.

3.3.2. Impact on noise propagation

The behavior of a wavefront sensor with respect to noise is described by the variance of its estimation error:

$$\sigma^2_{error} = \sum_i \sigma^2(\hat{a}_i - a_i) = \overbrace{\left(\sum_i \alpha_i\right)}^{\alpha} \frac{1}{n^{tot}_{ph}} + \overbrace{\left(\sum_i \beta_i\right)}^{\beta} \left(\frac{\sigma_e}{n^{tot}_{ph}}\right)^2 \tag{16}$$

with n^{tot}_{ph} the total number of detected photons and σ_e the read out noise standard deviation (in ph.e^-). The variance is computed as a sum over the estimated modes. The coefficients α_i and

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

β_i are respectively the photon noise and read out noise coefficients for the Zernike polynomial Z_i.

For $\mathbf{A_{OPD}} = \mathbf{0}$, we can get the theoretical variance from Eq. (12), and then compute easily the coefficients. If we set the read-out noise to 0, we find

$$\alpha = \sum_i (\mathbf{H}(\mathbf{A_d})^t \mathbf{R_n}^{-1} \mathbf{H}(\mathbf{A_d}))^{-1}[i,i] \times n_{ph}^{tot} \tag{17}$$

and if we remove photon noise, we find

$$\beta = \sum_i (\mathbf{H}(\mathbf{A_d})^t \mathbf{R_n}^{-1} \mathbf{H}(\mathbf{A_d}))^{-1}[i,i] \times \left(\frac{n_{ph}^{tot}}{\sigma_e}\right)^2 \tag{18}$$

One can similarly compute the modal noise coefficients α_i and β_i . Low values mean a low sensitivity to noise.

Analytical calculations have shown that the amount of astigmatism we introduce can be optimized to lower α and β (Fig. 4) for the estimation of tip/tilt and focus. The coefficients decrease dramatically at low astigmatisms, because of the indetermination removal. At higher astigmatisms, we get less flux per pixels, hence the coefficients rise again. We find a different optimum for the two coefficients: 0.42 rad leading to $\alpha = 1.64$ and 0.6 rad leading to $\beta = 84$. The best amount of astigmatism to be introduced is thus between these two values, depending on the detector and signal characteristics. It is closer to 0.42 rad at high flux, where photon noise is predominant, and closer to 0.6 rad at low flux, where read-out noise is more important. For our simulations and experiments, we chose an astigmatism offset approximately at the middle of the bracket, at 0.5 rad, but this value can be tuned depending on needs. This value should also lead to good criterion properties, as discussed in section 3.3.1.

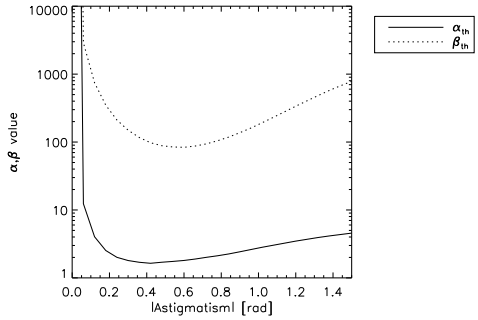


Fig. 4. α and β computed over 3 modes (tip/tilt/focus) depending on the amount of astigmatism introduced as phase offset (in absolute value).

We have justified the choice of astigmatism as a phase offset and demonstrated that it can be optimized depending on the conditions of experimentation. We now focus on the noise propagation analysis to compare LIFT with other sensors.

4. Comparison of LIFT's noise propagation with the Shack-Hartmann 2x2 and pyramid sensors

In the previous section, we related the variance of a WFS estimation error to noise through the coefficients α and β, which respectively quantify the sensitivity to photon noise and the sensitivity to read-out noise.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

We compare the values of α and β computed over tip/tilt and focus for LIFT with a 0.5 rad astigmatism, for a Shack-Hartmann 2×2 and for a fixed (e.g. without modulation) pyramid sensor with a pupil sampled on 4 pixels (Table 1). The parameters and formulas for computing the Shack-Hartmann coefficients are those used in references [16, 22]. In particular, for the Weighted Center of Gravity (WCoG), we consider a fixed Gaussian weighting, and the estimation is corrected by an unbiasing factor, as described by Thomas *et al.* [23]. The Full Width at Half Maximum (FWHM) of the weighting is equal to the Airy spot FWHM, i.e. 2 pixels at Shannon sampling. In our pyramid simulation, we performed a maximum likelihood estimation, and were able to compute α and β with the same analysis we used for LIFT (see Eq. (12)).

Table 1 shows that LIFT is approximately 5 times less sensitive to photon noise and 4 times less sensitive to read-out noise than a SH 2×2 using a WCoG. In LIFT, the photon noise sensitivity is reduced by the full aperture gain. The SH 2×2 divides the flux by 4, and each spot is imaged at the resolution of a subaperture. LIFT thus gets more flux and better precision. Moreover, sensitivity to detector noise is limited thanks to the maximum likelihood estimation properties [16]. As a result, for the estimation of tip/tilt, LIFT takes the better from the Center of Gravity (CoG), which is efficient in photon noise regime and the weighted center of gravity, which is efficient in read-out noise regime [22]. This effect is clearly visible on Fig. 5, which plots the variance of estimation error of LIFT, a CoG and a WCoG for the estimation of a tip only.

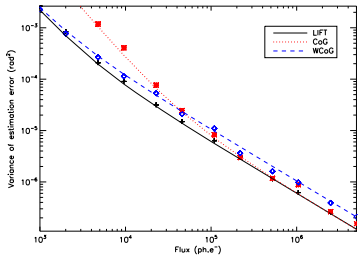


Fig. 5. Variance of the error on tip estimation for LIFT, the center of gravity and the weighted center of gravity at Shannon sampling on 8×8 pixels. Read-out noise = 10 e^-. Lines : theory. Symbols : simulations. The weighting used for the WCoG is a Gaussian, and the estimation is corrected by an unbiasing factor. The parameters are $N_t = N_w = 2$ (N_t : FWHM of the aberrated spot, N_w : FWHM of the weighting function, see reference [23]).

The pyramid sensor benefits from the full aperture gain as well, and sensitivity to detector noise is reduced by the use of only 4 pixels. The similarity of noise coefficients could thus be expected.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

Table 1. Comparison of noise propagation coefficients between LIFT, a Shack-Hartmann 2×2 (using either a center of gravity or a weighted center of gravity), and a fixed pyramid with a pupil sampled on 4 pixels for the estimation of tip/tilt and focus. Coefficients for the WCoG are computed in the case of an unbiased Gaussian weighting, with parameters $N_t = N_d = N_w = 2$ (see reference [23]).

	SH 2×2 (CoG)	SH 2×2 (WCoG)	LIFT	Pyramid
α	4.61	8.19	1.71	2.3
β	$> 10^4$	334	87	62

As an example of performance, we plot the variance of estimation error on tip/tilt and focus depending on flux for a read-out noise of 2 e^- (Fig. 6), which is the typical noise on high performance infrared cameras [24]. In this case, a Shack-Hartmann 2×2 using a weighted center of gravity would need a flux of 958 photons to reach a total variance of 0.01 rad^2 over the three modes, whereas LIFT and a pyramid sensor would respectively need 291 and 310 photons. The ratio between LIFT and the SH is 3.3, which amounts to more than 1 star magnitude. Table 2 sums up the gains in star magnitude obtained with LIFT and the pyramid compared to the SH, for the case above and an ideal case with photon noise only.

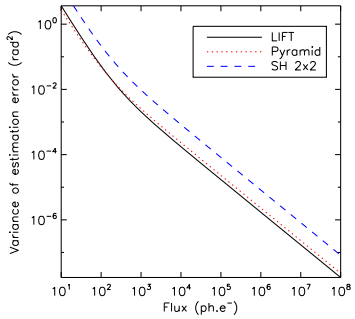


Fig. 6. Simulation of the total variance of tip/tilt/focus estimation error for LIFT, the SH 2×2 using a WCoG (with an unbiased Gaussian weighting) and the fixed pyramid. Read-out noise = 2 e$^-$.

Table 2. Gain in star magnitude obtained with LIFT and the pyramid compared to the Shack-Hartmann for the estimation of tip/tilt/focus, with a read-out noise of 2 e$^-$ and without read-out noise.

	LIFT	Pyramid
RON = 2 e$^-$	1.3	1.2
RON = 0 e$^-$	1.7	1.4

We were able to prove that LIFT is much more suited for low flux estimation than a Shack-Hartmann WFS. It has a general performance similar to a pyramid WFS, and is less sensitive to photon noise. In the following section, we validate our concept through lab experiments.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

5. Experimental validation

The previously mentioned performance of LIFT assumes a perfect knowledge of the astigmatism offset, a point source and a perfectly linear camera with uniform read-out noise, gain and dark current, as well as a known sampling. Robustness to uncertainties on these parameters had to be demonstrated. We want to test the linearity of LIFT in narrow and wide spectral band, as well as the noise propagation. To do that, we must be able to control the flux, the aberrations, the read-out noise, the sampling and the spectrum of the source. We have therefore performed experiments in realistic conditions. In the following, we discuss the obtained results.

5.1. LIFT's test bed optical setup

Figure 7 describes the optical setup. The object a single mode optical fiber connected to a laser diode light at 635 nm ($\Delta\lambda < 0.5$ nm). It can also be connected to a white light for wide band experiments. Two lenses collimate and then focus again the beam. Finally, we form the image with a CCD at Shannon sampling. The pupil is defined by a stop in the collimated beam, and the 0.5 rad astigmatism offset is inserted thanks to a tilted glass plate in the diverging beam coming from the source. The tip/tilt and focus are inserted by translating the camera along X, Y or Z. A Shack-Hartmann sensor was used to calibrate the astigmatism offset and the static aberrations. We neglect here differential aberrations.

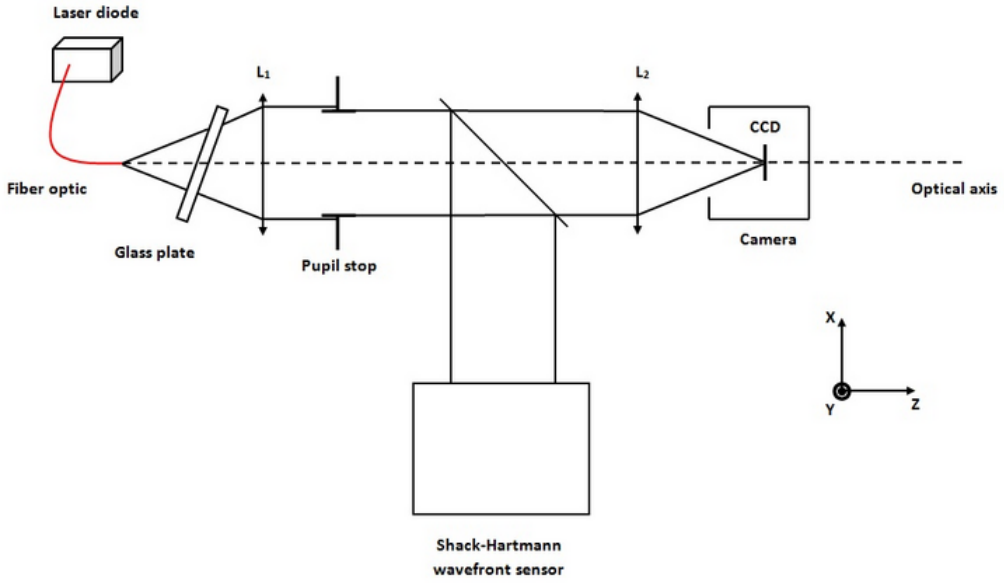


Fig. 7. Bench dedicated to the experimental validation of LIFT.

5.2. Linearity

5.2.1. Tip/tilt estimation

Without other aberrations, tip/tilt can easily be calibrated thanks to a center of gravity. This method is known to be effective at high flux, where the read-out noise becomes negligible [5]. The astigmatism offset does not bias the estimation, since the resulting spot is symmetrical with respect to both x and y axis. We thus formed images at high flux with different amounts of tip, in monochromatic light at 635 nm.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

Figure 8 shows the tip estimated by LIFT as a function of the inserted tip, calibrated with a center of gravity. The solid line follows the equation y=x. The estimations given by LIFT and the center of gravity match each other over more than ± 2 rad. The maximum error is 0.04 rad (less than a $\lambda/100$ rms wavefront error) at -2.6 rad of tip.

This demonstrates that LIFT's estimation of tip is linear and accurate on a wide range.

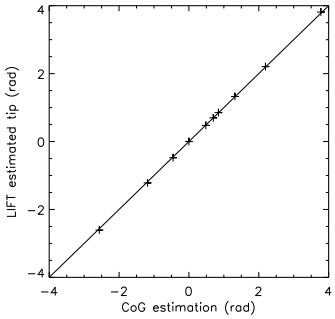


Fig. 8. Experimental estimation of a tip by LIFT versus a center of gravity estimation. Solid line: y=x.

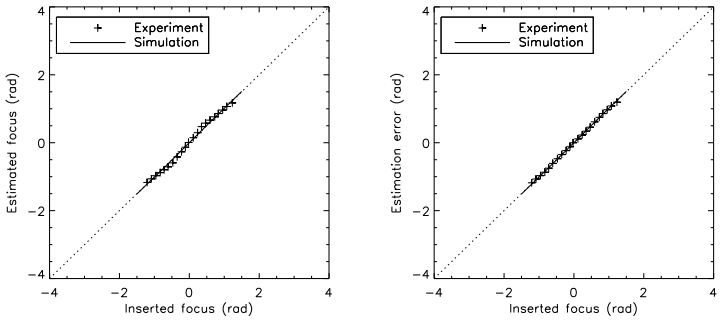
5.2.2. Focus estimation

Monochromatic case The Shack-Hartmann WFS could also be used for the calibration of focus. At high flux, its estimation is reliable. Similarly to the previous experiment, we took images with different amounts of focus. The astigmatism offset measured by the Shack-Hartmann and used in our model is 0.16 rad on Z_5 and 0.43 rad on Z_6. The phase offset is different from the optimum because of the experimental setup, which only allowed a rough adjustment. Nevertheless, as seen in Fig. 4, noise propagation coefficients do not vary significantly around the optimum. The combination of Z_5 and Z_6 is not far from 0.5 rad here, and is thus acceptable.

In Fig. 9(a), we have plotted the estimation of LIFT versus the Shack-Hartmann estimation. The dotted line is y=x and the solid line is a simulation. We globally observe a good linearity, with a maximum error of 0.15 rad ($\sim \lambda/40$ rms).

The error is symmetrical with respect to 0 and can be explained by a difference between the astigmatism seen on the detector and the one used in our model. The data acquired for this experiment also allowed to apply a phase diversity algorithm [14], so as to obtain the aberrations on the imaging camera. By performing this method, we found that differential aberrations added 0.1 rad on both Z_5 and Z_6. On this last mode, this means a model error of $\sim 25\%$. Figure 9(b) shows the estimation with correction of the astigmatism in the model. The maximum error is then 0.04 rad ($\sim \lambda/160$ rms).

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013



(a) Without astig correction (b) With astig correction

Fig. 9. Experimental estimation of a focus by LIFT with monochromatic light. In (a), the astigmatism offset used in the model is the one measured by the Shack-Hartmann. In (b), the model is corrected by adding the differential aberrations found by phase diversity.

The focus estimation is thus accurate in the range [-1, 1] rad, even with a significant model error on the astigmatism offset.

Polychromatic case The estimation of focus was also made with a polychromatic source. Its FWHM is 200 nm for a central wavelength at 676 nm (R=3.4). As seen in section 2, LIFT can make an estimation in polychromatic light by summing the contributions from the different wavelengths. But simulations have shown that assuming solely a contribution of the central wavelength, it still gives a good performance [17], as long as the spectral bandwidth is narrow enough. Moreover, we checked that the coefficients of noise propagation do not vary significantly in this case, and we thus have the same optimum as in monochromatic light. We thus estimated the focus in this configuration.

The astigmatism offset measured by the Shack-Hartmann and used in our model is 0.11 rad on Z_5 and 0.5 rad on Z_6 at the central wavelength. The inserted focus in radians was also computed from the Shack-Hartmann measurements at the central wavelength.

Figure 10 shows the source spectrum and the result of the estimation. As for monochromatic light, a linear estimation is observed. The maximum error goes up to 0.12 rad ($\sim \lambda/50$ rms) in negative foci, and 0.03 rad ($\sim \lambda/200$ rms) in positive foci. The higher error for negative values is due to lateral chromatic aberration coming along with the tilted plate, which was not taken into account in our model. Again, the robustness of LIFT is demonstrated here.

As a conclusion, the estimation of LIFT in wide band can be as accurate as in monochromatic light, with no additive calculations.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

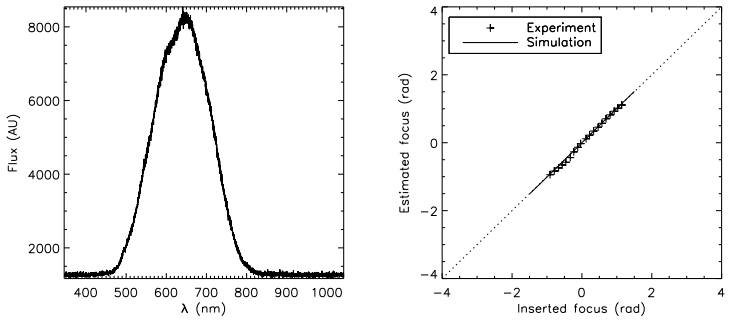


Fig. 10. Left: Spectrum of the source used in the estimation with large bandwidth light. Right: Experimental estimation of a focus by LIFT in large bandwidth.

5.3. Noise propagation

We intend here to validate the simulations shown in section 5.1 by observing the evolution of the estimation error variance with the flux. To get interpretable data, where we could see both the read-out noise and the photon noise regimes, we changed the conditions of experimentation. The astigmatism was raised to 1 rad to increase α and β. The equivalent read-out noise in this experiment is 14 ph.e^-. The estimation is made on a zone of 20×20 pixels.

We took several series of images at different flux and computed an empirical variance for tip/tilt and focus (Fig. 11). Given the conditions of experiment (sampling, phase offset...), we are able to find the noise propagation coefficients α and β thanks to analytical calculations (see section 3.3.2). Once we have these coefficients, we can compute a theoretical variance of estimation error at any flux.

The theoretical variance and the empirical one match over several decades of flux, confirming our model of noise propagation.

A saturation can be seen on both figures. It is estimated from the value at the highest flux, and is equal to $5.9\pm3.4\times10^{-5}rad^2$ for tip and $7\pm3.3\times10^{-7}rad^2$ for focus. We checked that the ratio between these saturations is compatible with vibrations of the bench.

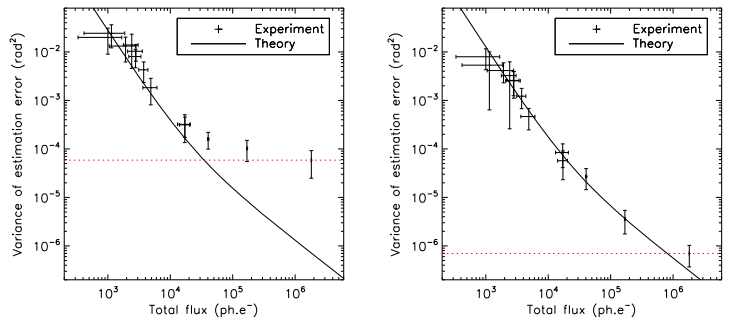


Fig. 11. Experimental variance of estimation error for tip/tilt (left) and focus (right). RON = 14 $ph.e^-$. Red dotted line: saturation due to vibrations.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013

6. Conclusion

We detailed here the concept of LIFT. An astigmatism offset is added to the incoming phase to remove the indetermination on focus, then a maximum likelihood algorithm is performed on the focal image. Iterations compensate for the approximation of small phase. A theoretical study of the minimization criterion and noise propagation has let us understand better the algorithm and justify the choice of an astigmatism offset rather than focus, as well as the choice of the offset amplitude.

We have demonstrated that LIFT is robust to setup uncertainties and provides a linear focus estimation, even with large spectral bandwidth light (R=3.4). The validation of noise propagation confirms the superiority of LIFT over a Shack-Hartmann 2×2: LIFT has more than a 1 magnitude gain over a Shack-Hartmann 2×2 for a variance of 0.01 rad^2 and a read-out noise of 2 e^-. This performance is similar to what one could have with a fixed pyramid sensor. Moreover, LIFT is light in hardware, easy to set up, and thanks to its algorithm, automatically adapts the weighting on pixels depending on noise properties. Its offset phase and sampling can be tuned to match most of systems. Finally, it relies on well sampled focal images that can be interpreted easily, which eases the debugging stage.

LIFT has already been tested on an operational system: Gemini Southern Observatory [25]. An on-sky validation is planned for spring 2013. Also, a deeper study of the algorithm will be made to get a further comparison with the pyramid and other wavefront sensors. However, LIFT can already be considered as an effective and robust low-order WFS in the context of wide field adaptive optics for very large and extremely large telescopes.

Acknowledgments

The authors would like to thank Bruno Fleury for his help on understanding some experimental issues, and Baptiste Paul for helping in the setup of the bench. The work was partly funded by the European Commission under FP7 Grant Agreement No. 312430 Optical Infrared Coordination Network for Astronomy.

#185587 - $15.00 USD Received 19 Feb 2013; revised 29 Apr 2013; accepted 29 Apr 2013; published 2 Jul 2013