Analytical expression of a long exposure coronagraphic point spread function. Application to the estimation of quasi-static aberrations in the presence of residual turbulence with COFFEE

Olivier Herscovici-Schiller[a], Laurent M. Mugnier[a], Jean-François Sauvage[a,b], Jean-Michel Le Duigou[c], and Faustine Cantalloube[a,d]

[a]Office national d'études et de recherches aérospatiales (ONERA), 29, avenue de la division Leclerc, 92322 Châtillon CEDEX, France
[b]Laboratoire d'astrophysique de Marseille (LAM), UMR 7326 CNRS – Aix-Marseille Université, 13388, Marseille, France
[c]Centre national d'études spatiales (CNES), 18 Avenue Edouard Belin, 31400 Toulouse, France
[d]Univ. Grenoble Alpes, Institut de Planétologie et d'Astrophysique de Grenoble (IPAG, UMR 5274), F-38000 Grenoble & CNRS, France

ABSTRACT

The resolution of coronagraphic high contrast exoplanet imaging devices such as SPHERE is limited by quasi-static aberrations. These aberrations produce speckles that can be mistaken for planets in the image. In order to design instruments, correct quasi-static aberrations or analyze data, the expression of the point spread function of a coronagraphic telescope in the presence of residual turbulence is useful. We have derived an analytic formula for this point spread function. We explain physically its structure, we validate it by numerical simulations and we show that it is computationally efficient. Finally, we use it in a simulation of focal-plane wave front estimation in the COFFEE method (based on coronagraphic phase diversity). The preliminary results, which give a sub-nanometric precision in the case of a SPHERE-like system, strongly suggest that quasi-static aberrations could be calibrated during observations.

Keywords: telescopes, wave front sensing, coronagraph, turbulence, high angular resolution, adaptive optics

1. INTRODUCTION

Direct detection of exoplanets is limited by the presence of speckles on scientific images. The origin of these speckles is quasi-static aberrations in the optical system.

In a high-contrast imaging system without residual turbulence, or for a ground-based adaptive optics-corrected system recording short exposures, the on-axis coronagraphic point spread function of the system can easily be computed as a function of the instantaneous aberrations (quasi-static and turbulence-induced). To the best of our knowledge, no expression has been previously published for the *long exposure* coronagraphic point spread function. Such an expression would be of great use for at least three different applications:

- The design, simulation and optimization of the coronagraphic system;
- The analysis and correction of the quasi-static aberrations,[1] which are the ultimate limitation of the performance of the coronagraph;
- The image reconstruction.[2]

Further author information: (Send correspondence to O. H.-S.)
O. H.-S.: E-mail: olivier.herscovici@onera.fr, Telephone: +33 (0)1 46 73 47 62

All those aspects demand that we know a direct model of image formation through a coronagraphic telescope. In the case of a ground-based telescope, with typical exposure times of several tens of seconds to minutes, turbulence, even in the form of residual turbulence of an extreme adaptive optics system, must be taken into account. We derive the long exposure coronagraphic point spread function as in Ref. 3, we interpret it, we validate it numerically, and we use it in the coronagraphic phase diversity COFFEE method, obtaining a nanometric reconstruction error.

2. INSTANTANEOUS POINT SPREAD FUNCTION OF A TELESCOPE EQUIPPED WITH A CORONAGRAPH

Let us consider a telescope equipped with a coronagraph as described in figure 1.

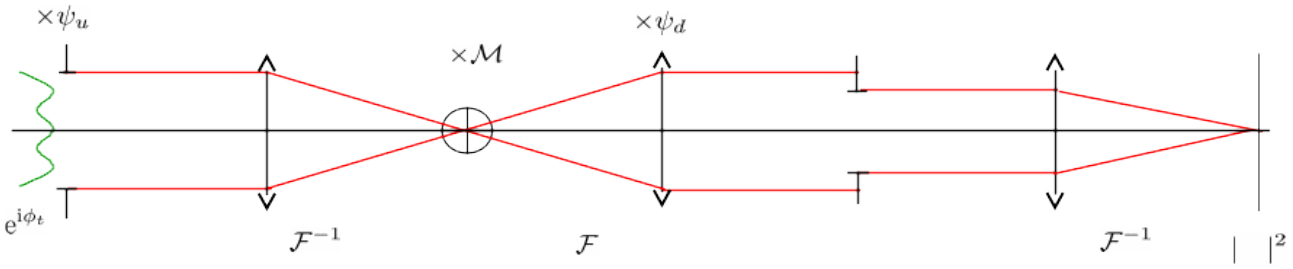

Figure 1. Sketch of a telescope equipped with a coronagraph

Let us denote by h_c the coronagraphic point spread function of the instrument without turbulence. It writes

$$h_c\left(\alpha;\psi_u,\psi_d\right)=\left|\mathcal{F}^{-1}\left\{\psi_d\times\mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u\right)\right]\right\}\left(\alpha\right)\right|^2. \tag{1}$$

This equation simply expresses the propagation from successive planes to the next ones, and the quadratic detection, in Fig. (1). The coordinate α is in the detector focal plane, $\psi_u = P_u e^{i\phi_u}$ where P_u is the upstream pupil and ϕ_u the quasi-static upstream phase aberration, $\mathcal{M}$ the coronagraphic focal plane mask, $\psi_d = P_d e^{i\phi_d}$ where P_d is the downstream pupil and ϕ_d the quasi-static downstream phase aberration, and $\mathcal{F}$ the bi-dimensional Fourier transform in space:

$$\mathcal{F}[f](r)=\iint_{\mathbb{R}^2} f(\alpha)\times e^{-i2\pi\alpha\cdot r}\,d\alpha. \tag{2}$$

For notational simplicity we take $\lambda = 1$ in the equations, that is to say we use the same coordinate system for the pupil phases and for the transfer functions.

In the presence of turbulence, the instantaneous point spread function of this instrument writes[1]

$$h_{sec}(\alpha,t;\psi_u,\psi_d)=h_c(\alpha;\psi_u\times e^{i\phi_t(t)},\psi_d). \tag{3}$$

Here $h_{sec}(\alpha,t,\psi_u,\psi_d)$ (*sec* stands for "short exposure coronagraphic") is the instantaneous point spread function taken at coordinate α in the detector focal plane at time t and $\phi_t(t)$ is the instantaneous turbulent phase at time t if the telescope is ground-based and equipped with adaptive optics.

For notational simplicity we take $\lambda = 1$ in the equations, that is to say we use the same coordinate system for the pupil phases and for the transfer functions.

3. LONG EXPOSURE CORONAGRAPHIC POINT SPREAD FUNCTION

In current high contrast systems such as SPHERE or GPI, the exposure time is large with respect to the typical evolution time of corrected turbulence, so that the signal that is actually recorded by the detector is the long exposure point spread function, which is an average over time of the short exposure point spread function.

$$h_{lec}(\alpha;\psi_u,\psi_d) = \langle h_{sec}(\alpha,t;\psi_u,\psi_d)\rangle_t \,. \tag{4}$$

Here h_{lec} is the long exposure coronagraphic point spread function, and $\langle\rangle_t$ is the averaging over time. The Wiener-Khintchine theorem states that

$$h_{sec}(\alpha,t;\psi_u,\psi_d) = \mathcal{F}^{-1}\left(\left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\} \otimes \left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\}\right)(\alpha), \tag{5}$$

where the bi-dimensional correlation product, noted $\otimes$, is defined by :

$$f\otimes g(\alpha) = \iint_{\mathbb{R}^2} f^*(\alpha_1)\times g(\alpha_1+\alpha)\,\mathrm{d}\alpha_1. \tag{6}$$

Here f^* is the complex conjugate of the function f. In developed form, the long exposure coronagraphic point spread function now writes:

$$h_{lec}(\alpha;\psi_u,\psi_d) = \left\langle \mathcal{F}^{-1}\left(\left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\} \otimes \left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\}\right)(\alpha)\right\rangle_t. \tag{7}$$

To somewhat simplify calculations, we will consider the optical transfer function instead of the point spread function. Let $\tilde{h} = \mathcal{F}(h)$ be the Fourier transform of the point spread function. The integral over space commutes with the averaging over time, so

$$\tilde{h}_{lec}(r;\psi_u,\psi_d) = \left\langle \left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\} \otimes \left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{\mathrm{i}\phi_t(t)}\right)\right]\right\}(r)\right\rangle_t. \tag{8}$$

In order to express the average over time, we have no choice but to fully develop this expression, which after some manipulations reads

$$\begin{aligned}\tilde{h}_{lec}(r;\psi_u,\psi_d) = &\iint \psi_d^*(r_1)\psi_d(r+r_1)\iint\iint \mathrm{e}^{\mathrm{i}2\pi r_1\cdot\alpha_1}\mathrm{e}^{-\mathrm{i}2\pi(r_1+r)\cdot\alpha_2}\mathcal{M}^*(\alpha_1)\mathcal{M}(\alpha_2)\\ &\times\iint\iint \mathrm{e}^{-\mathrm{i}2\pi r_2\cdot\alpha_1}\mathrm{e}^{\mathrm{i}2\pi r_3\cdot\alpha_2}\psi_u^*(r_2)\psi_u(r_3)\left\langle \mathrm{e}^{\mathrm{i}[\phi_t(r_3,t)-\phi_t(r_2,t)]}\right\rangle_t \mathrm{d}r_3\,\mathrm{d}r_2\,\mathrm{d}\alpha_2\,\mathrm{d}\alpha_1\,\mathrm{d}r_1.\end{aligned} \tag{9}$$

Following Roddier,[4] we assume that turbulence is an ergodic stationary process (which is a very reasonable assumption in the case of residual turbulence after an extreme adaptive optics system), so we write

$$\left\langle \mathrm{e}^{\mathrm{i}[\phi_t(r_3,t)-\phi_t(r_2,t)]}\right\rangle_t = \mathrm{e}^{-\frac{1}{2}D_\phi(r_3-r_2)}, \tag{10}$$

where D_ϕ is the turbulent phase structure function, defined by

$$D_\phi(r) = \left\langle [\phi_t(r') - \phi_t(r+r')]^2\right\rangle_{r'}. \tag{11}$$

In order to be able to separate variables r_2 and r_3 in Eq. (11), we take the inverse* Fourier transform of $\mathrm{e}^{-\frac{1}{2}D_\phi}$. We denote α' the conjugate variable of r_3-r_2, and h_a the inverse Fourier transform of $\mathrm{e}^{-\frac{1}{2}D_\phi}$, that is $h_a = \mathcal{F}^{-1}\left(\mathrm{e}^{-\frac{1}{2}D_\phi}\right)$ (we will come back later to the meaning of h_a).

*Note that we take the inverse Fourier transform and not the direct Fourier transform to be consistent with the convention to take the inverse Fourier transform when going from a pupil plane to a focal plane. However, since D_ϕ is real and even, taking its direct or inverse Fourier transform yields the same result.

We obtain that

$$\tilde{h}_{lec}(r;\psi_u,\psi_d,D_\phi) = \iint h_a(\alpha';D_\phi) \iint \psi_d^*(r_1) \iint \mathrm{e}^{\mathrm{i}2\pi r_1\cdot\alpha_1}\mathcal{M}^*(\alpha_1) \iint \mathrm{e}^{-\mathrm{i}2\pi r_2\cdot\alpha_1}\psi_u^*(r_2)\mathrm{e}^{-\mathrm{i}2\pi r_2\cdot\alpha'}\,\mathrm{d}r_2\,\mathrm{d}\alpha_1 \times \psi_d(r+r_1) \iint \mathrm{e}^{-\mathrm{i}2\pi(r_1+r)\cdot\alpha_2}\mathcal{M}(\alpha_2) \iint \mathrm{e}^{\mathrm{i}2\pi r_3\cdot\alpha_2}\psi_u(r_3)\mathrm{e}^{\mathrm{i}2\pi r_3\cdot\alpha'}\,\mathrm{d}r_3\,\mathrm{d}\alpha_2\,\mathrm{d}r_1\,\mathrm{d}\alpha'. \tag{12}$$

In a more compact form, this writes:

$$\tilde{h}_{lec}(r;\psi_u,\psi_d,D_\phi) = \iint h_a(\alpha';D_\phi)\left\{\psi_d\times\mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id})}\right)\right]\right\}\otimes\left\{\psi_d\times\mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id})}\right)\right]\right\}(r)\,\mathrm{d}\alpha'. \tag{13}$$

To obtain the point spread function back from this optical transfer function, we just have to take the inverse Fourier transform, then apply Wiener-Khintchine's theorem again. Finally, the long exposure coronagraphic point spread function reads

$$h_{lec}(\alpha;\psi_u,\psi_d,D_\phi) = \iint h_a(\alpha';D_\phi)\left|\mathcal{F}^{-1}\left\{\psi_d\times\mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id}}\right)\right]\right\}(\alpha)\right|^2\,\mathrm{d}\alpha'. \tag{14}$$

This expression gives the long exposure coronagraphic point spread function as a function of three deterministic parameters, namely: upstream aberrations, downstream aberrations, and residual turbulence-induced phase structure function.

4. PHYSICAL INTERPRETATION

In order to interpret Eq. (14) physically, we can re-write it in the more compact following form :

$$h_{lec}(\alpha;\psi_u,\psi_d,D_\phi) = \iint h_a(\alpha';D_\phi)h_c\left(\alpha;\psi_u\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id}},\psi_d\right)\,\mathrm{d}\alpha', \tag{15}$$

where $h_c\left(\alpha;\psi_u\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id}},\psi_d\right)$ is the coronagraphic point spread function of the instrument in the absence of turbulence, but with a tilt α' *added* to the upstream aberrations, that is to say with the light coming from the star being tilted by an angle α'.

The so-called atmospheric point spread function has been defined as $h_a = \mathcal{F}^{-1}\left(\mathrm{e}^{-\frac{1}{2}D_\phi}\right)$. it can be interpreted as a probability density, giving the probability for a photon coming from space to be scattered by turbulence in the direction $2\pi\alpha'$. Indeed, we have defined h_a as

$$h_a = \mathcal{F}^{-1}\left[\exp\left(-\frac{1}{2}D_\phi\right)\right] = \mathcal{F}^{-1}\left[\left\langle\mathrm{e}^{\mathrm{i}[\phi_t(r_3,t)-\phi_t(r_2,t)]}\right\rangle_t\right]. \tag{16}$$

If we note $\psi_t(r,t) = \exp\left(\mathrm{i}\phi_t(r,t)\right)$ the contribution to the electric field at position r and time t by the atmospheric turbulence, we can re-write h_a as :

$$h_a = \mathcal{F}^{-1}\left[\langle\psi_t^*(r_2,t)\psi_t(r_3,t)\rangle_t\right] \tag{17}$$

Assuming stationarity and ergodicity, we recognize the auto-correlation of ψ_t in $\langle\psi_t^*(r_2,t)\psi_t(r_3,t)\rangle_t$. So, thanks to Wiener-Khintchine's theorem, we can identify h_a as the energy spectral density of the turbulence-induced complex field, which is the point spread function associated with that field :

$$h_a = \left|\mathcal{F}^{-1}\left[\psi_t\right]\right|^2. \tag{18}$$

Finally, Eq. (15) can be interpreted as follows: the long exposure coronagraphic point spread function h_{lec} is a weighted sum of coronagraphic point spread functions h_c, without any turbulence, but with an upstream tilt. The weight on any of those tilted point spread functions is the probability that the atmosphere scatters light in the direction of the corresponding tilt.

Moreover, the formula Eq. (15) separates the turbulent part and the coronagraphic part of the equation: h_a codes for the characteristics of the turbulent atmosphere, while h_c codes for the characteristics of the telescope and instrument.

5. SPECIAL CASES

5.1 Non-turbulent point spread function

In the case where there is no turbulence, h_a is reduced to a Dirac distribution. Indeed, in the case where there is no turbulence, ϕ_t is a constant, so $D_\phi = \left\langle [\phi_t(r') - \phi_t(r+r')]^2 \right\rangle_{r'} = 0$. Then

$$h_a = \mathcal{F}^{-1}\left[\exp\left(-\frac{1}{2}D_\phi\right)\right] = \mathcal{F}\left[1\right] = \delta. \tag{19}$$

Hence, by use of Eq. (14),

$$\begin{aligned} h_{lec}(\alpha;\psi_u,\psi_d) &= \iint \delta(\alpha') \left|\mathcal{F}^{-1}\left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u \mathrm{e}^{-\mathrm{i}2\pi\alpha'\cdot\mathrm{Id}}\right)\right]\right\}(\alpha)\right|^2 \mathrm{d}\alpha' \\ h_{lec}(\alpha;\psi_u,\psi_d) &= h_c(\alpha;\psi_u,\psi_d) \\ h_{lec}(\alpha;\psi_u,\psi_d) &= \left|\mathcal{F}^{-1}\left\{\psi_d \times \mathcal{F}\left[\mathcal{M}\times\mathcal{F}^{-1}\left(\psi_u\right)\right]\right\}(\alpha)\right|^2. \end{aligned} \tag{20}$$

This is precisely the classical expression for a coronagraphic point spread function in the absence of turbulence.

5.2 Non-coronagraphic optical transfer function

If we consider a non-coronagraphic instrument, $\mathcal{M} = 1$, so Eq. (14) now reads

$$h_{lec}(\alpha;\psi_u,\psi_d,D_\phi) = \iint h_a(\alpha';D_\phi)\left|\mathcal{F}^{-1}\left\{\psi_d\psi_u \mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot\mathrm{Id}}\right\}(\alpha)\right|^2 \mathrm{d}\alpha'. \tag{21}$$

Once again we take the optical transfer function and use Wiener-Khintchine's theorem.

$$\begin{aligned} \tilde{h}_{lec}(r;\psi_u,\psi_d D_\phi) &= \iint h_a(\alpha';D_\phi) \iint \psi_d^*(r')\psi_u^*(r')\mathrm{e}^{-\mathrm{i}2\pi\alpha'\cdot r'}\psi_d(r+r')\psi_u(r+r')\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot(r+r')}\,\mathrm{d}r'\,\mathrm{d}\alpha' \\ &= \iint \psi_d^*(r')\psi_u^*(r')\psi_d(r+r')\psi_u(r+r')\,\mathrm{d}r' \iint h_a(\alpha';D_\phi)\mathrm{e}^{\mathrm{i}2\pi\alpha'\cdot r}\,\mathrm{d}\alpha' \\ \tilde{h}_{lec}(r;\psi_u,\psi_d) &= \psi_d\psi_u \otimes \psi_d\psi_u(r) \times \mathrm{e}^{-\frac{1}{2}D_\phi(r)}. \end{aligned} \tag{22}$$

This is the classic Roddier formula[4] of the long exposure transfer function for imaging through turbulence.

5.3 Approximation in the case of small turbulence

We consider the case of a telescope equipped with an extreme adaptive optics system. The residual turbulent phase after the adaptive optics can then be seen as a small perturbation. It is straightforward to see that

$$\frac{1}{2}D_\phi(r) = \sigma_\phi^2 - A_\phi(r), \tag{23}$$

where $A_\phi(r) = \langle \phi_t(r')\phi_t(r+r')\rangle_{r'}$ is the autocorrelation of the phase, hence

$$h_a = \mathrm{e}^{-\sigma_\phi^2}\mathcal{F}^{-1}\left[\mathrm{e}^{A_\phi}\right]. \tag{24}$$

Then, if the turbulence is small, A_ϕ is small compared to 1, and we can perform a first order MacLaurin expansion:

$$\begin{aligned} h_a(\alpha') &\simeq \mathrm{e}^{-\sigma_\phi^2} \mathcal{F}^{-1}\left[1 + A_\phi\right](\alpha') \\ h_a(\alpha') &\simeq \mathrm{e}^{-\sigma_\phi^2} \delta(\alpha') + \mathrm{e}^{-\sigma_\phi^2} S_\phi(\alpha') \end{aligned} \tag{25}$$

where S_ϕ is the power spectrum density of the turbulent phase ϕ_t, defined as the (inverse) Fourier transform of A_ϕ. Thus we can express h_{lec} :

$$h_{lec}(\alpha; \psi_u, \psi_d) \simeq \mathrm{e}^{-\sigma_\phi^2} \times \left[h_c(\alpha; \psi_u, \psi_d) + \iint S_\phi(\alpha') h_c\left(\alpha; \psi_u e^{\mathrm{i}2\pi\alpha'}, \psi_d\right) \, \mathrm{d}\alpha' \right] \tag{26}$$

This means that h_{lec} is approximately the non-turbulent coronagraphic point spread function, with a corrective additive term that takes into account the power spectrum density of the turbulent phase, all this dampened by the coherent energy $\mathrm{e}^{-\sigma_\phi^2}$

6. NUMERICAL VALIDATION AND EFFICIENCE CONSIDERATIONS

6.1 Numerical validation

To validate our point spread function model, we test it against an average of short exposure point spread functions, each one of them with a different outcome of the residual turbulence-induced phase. Each short exposure point spread function is computed using a matrix Fourier transform[5] in order to be accurate on the mask focal plane. A few averages of point spread functions are displayed on Fig. 2. They are simulated with a Lyot coronagraph of extension $3\lambda/D$, the residual turbulence of a SPHERE-like adaptive optics and upstream and downstream white noise phase aberrations of variance 0.1 rad^2.

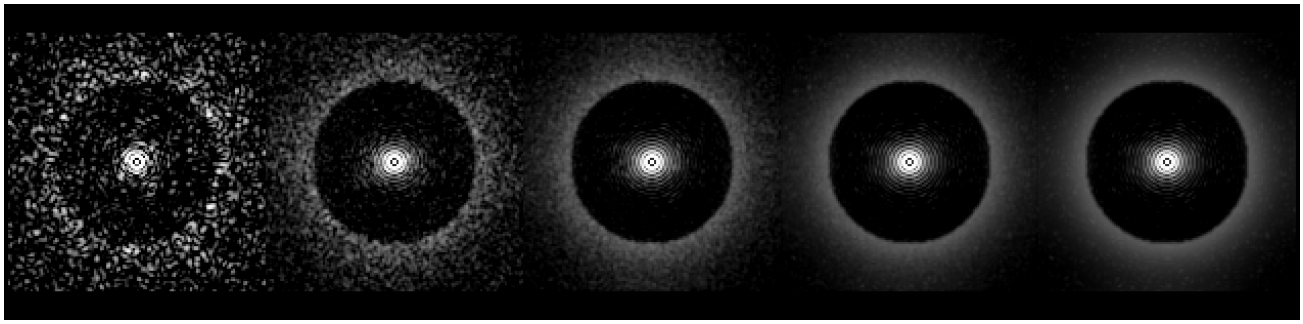

Figure 2. Point spread functions. From left to right: average of 1, 10, 100 and 1000 exposures, and analytic long exposure.

Let us define the convergence error as follows. We denote by h_{lec} the long exposure point spread function using the analytic formula of Eq. (14). We denote h_{sseN} the average of N short exposures. We define the error as:

$$\mathrm{err}_N = \max\left(\frac{|h_{sseN} - h_{lec}|}{h_{lec}}\right) \tag{27}$$

The division is taken pixel by pixel. We plot the evolution of the error as a function of the number of short exposures on Fig. 3. The evolution of the error is independent of the size of the images. From this evolution, we conclude without surprise that the average of short exposures tends to the analytical formula, with the error evolving as $N^{-1/2}$.

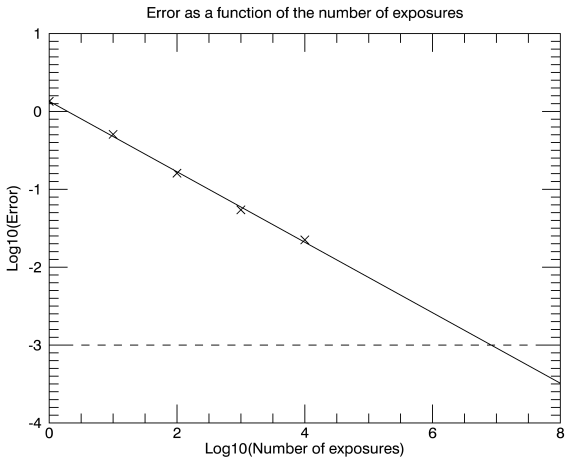

Figure 3. Convergence error as a function of the number of exposures (bi-logarithmic graph).

6.2 Computing cost

The evolution of the convergence error shown on Fig. 3 gives us an easy criterion to quantify the comparative computing costs of the long exposure coronagraphic point spread function and the empirical average of short exposures.

Let us take the computing cost of a short exposure point spread function as the unit computing cost. Then, if we want an error of less than 10^{-3} on the point spread function, Fig. 3 indicates that we must perform an average of approximately 10^7 short exposure point spread functions, for a cost of 10^7. The analytical formula for the long exposure point spread function has a total cost of the number of points on which the phase structure function is known. This implies that, for square images of 512×512 pixels, our exact formula is about 38 times less costly to evaluate than an average of short exposures.

In addition, it should be noted that, since the long exposure is an integral, and thus, in practice, a sum, it is very easy to compute it in parallel on several processors. We made all calculations in parallel on 16 cores.

Finally, it should also be noted that a useful approximation can accelerate the computing of the long exposure point spread function by a great deal. Indeed, when the tilt α' is greater than the radius of the mask, the point spread function can be well approximated by a shifted non-coronagraphic point spread function that can be computed once and for all, so the sum in Eq. (15) for the computation of the long exposure point spread function must actually only be computed on a square of side length typically $10\lambda/D$. That is only 20×20 short exposure point spread functions.

7. APPLICATION: ESTIMATION OF QUASI-STATIC ABERRATIONS IN THE PRESENCE OF RESIDUAL TURBULENCE WITH COFFEE

7.1 Motivation

In the context of direct exoplanet detection, during an observation from a terrestrial telescope, the quasi-static aberrations cause light from the star to leak on the focal plane and create speckles on the scientific image. Those speckles can be mistaken for planets, or hide planets. Hence, it is necessary to estimate and correct these aberrations. In current systems such as SPHERE these quasi-static aberrations are only corrected during daytime. For an upgrade, or for future high contrast instruments on extremely large telescopes, it would be useful to correct them during the scientific acquisition. To achieve this, we combine COFFEE,[1] which has already been demonstrated on SPHERE (with an internal calibration source),[6] with the long exposure phase diversity.[7]

7.2 Principle of COFFEE

COFFEE (for COronagraphic Focal-plane wave-Front Estimation for Exoplanet detection), consists in an extension of conventional phase diversity to a coronagraphic system: aberrations are estimated using two coronagraphic focal-plane images, recorded from the scientific camera itself, and thus without any differential aberration. The main idea is to find the phase that minimizes a regularized Maximum Likelihood (or Maximum A Posteriori) criterion. This criterion is the sum of several terms. The first term is the distance of the observed focused image to the focused image that is predicted knowing the model of image formation (this is where the phases are taken into account). The second term is also a distance, but calculated on the diversity images (the phase parameters are also taken into account). The other terms are regularisation terms. Mathematically, the criterion is of the form :

$$\begin{aligned} J(\phi_u, \phi_d) = & \frac{1}{2} \left|\left| \frac{i_{\text{foc}} - \alpha_{\text{foc}} h_{lec}(\phi_u, \phi_d; D_\phi) \star h_{\text{detector}} - \beta_{\text{foc}}}{\sigma_{\text{foc}}} \right|\right|^2 \\ & + \frac{1}{2} \left|\left| \frac{i_{\text{div}} - \alpha_{\text{div}} h_{lec}(\phi_u + \phi_{\text{div}}, \phi_d; D_\phi) \star h_{\text{detector}} - \beta_{\text{div}}}{\sigma_{\text{div}}} \right|\right|^2 + \mathcal{R}(\phi_u, \phi_d) \end{aligned} \tag{28}$$

In this formula, foc indicates a focused image, div a diversity image, i is the image on the detector, α the light flux, h_{lec} the point spread function of the telescope taking turbulence into account, h_{detector} the impulse response of the detector, β the background, σ the standard deviation of the noise, $\mathcal{R}$ a regularisation term, and ϕ_{div} the known diversity phase.

The two first terms represent the fidelity to data. The last term represents a priori information on the phases.

7.3 Simulation results

For a preliminary validation, we have made a reconstruction on simulated data. The parameters of the simulation were the following :
The phase images size was 64×64 pixels, the incoming flux was 10^9 photoelectrons, the standard deviation of the readout noise was 1 electron, the wavelength 1589 nm, the phase structure function is typical of the residual turbulence of the extreme adaptive optics system of the SPHERE instrument of the VLT, the coronagraph is a Roddier & Roddier phase mask, and there are no downstream aberrations. The upstream aberration is 50 nm RMS. Figure 4 shows the true upstream aberrations, the estimated upstream aberrations, and the difference between the upstream aberration and the estimated aberration. The error is in this case about 1.4 nm RMS.

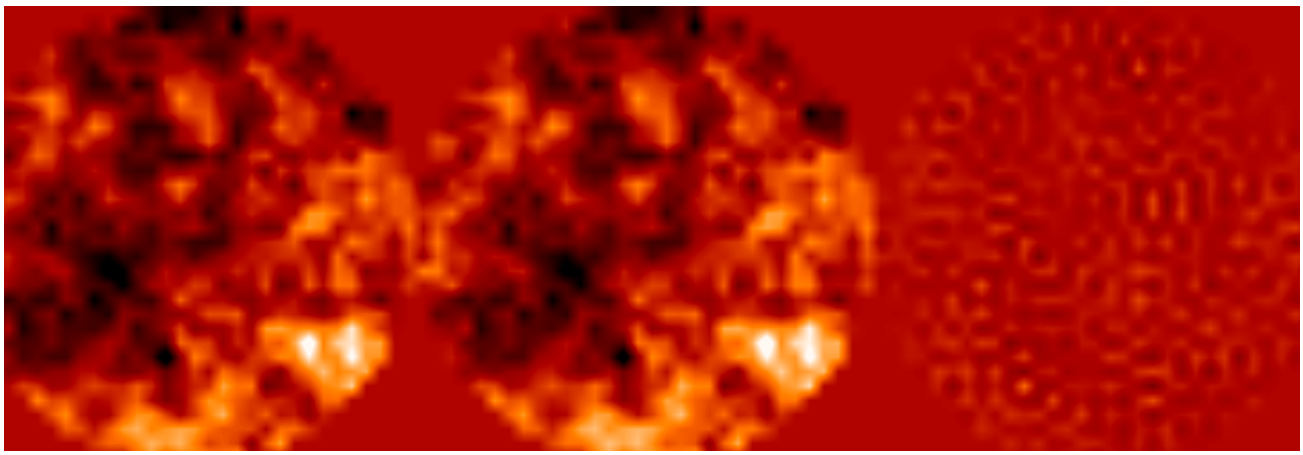

Figure 4. Left: upstream aberrations that we aim to reconstruct. Middle : reconstructed aberration. Right : difference, magnified ten times.

However, the usefulness of this reconstruction is even better than what the reconstruction error suggests. Indeed, let us consider the Fourier transform of the difference, as displayed on figure 5. We notice that the error is mainly located on the spatial frequencies that are on the border of the image, that is to say, those that are not corrected by the deformable mirror. If we take into account only the error on the corrected zone, the error drops to less than 0.35 nm RMS.

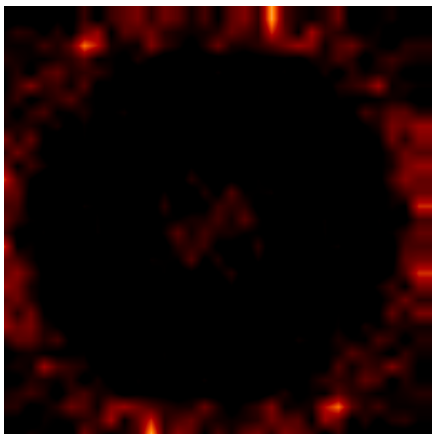

Figure 5. Spectral representation of the reconstruction error

8. CONCLUSION

We have derived an analytical expression to model image formation for a coronagraphic telescope through the residual turbulence of an extreme adaptive optics system. This model is general for any coronagraphic mask. It has a physical meaning: it separates between the atmospheric contribution and the contribution of the instrument, and can be seen as a plane wave decomposition. It is computationally effective, allowing a time gain of at least an order of magnitude on standard sized images. Lastly, we have shown that it can be used for the online measurement and correction of aberrations in the context of ground-based direct exoplanet detection.

ACKNOWLEDGMENTS

The PhD of O. Herscovici-Schiller is founded jointly by ONERA and CNES. The work presented here received support from the seventh Opticon Framework Program.

REFERENCES

[1] Paul, B., Mugnier, L. M., Sauvage, J.-F., Dohlen, K., and Ferrari, M., "High-order myopic coronagraphic phase diversity (COFFEE) for wave-front control in high-contrast imaging systems," *Optics Express* **21**(26), 31751–31768 (2013).

[2] Ygouf, M., Mugnier, L. M., Mouillet, D., Fusco, T., and Beuzit, J.-L., "Simultaneous exoplanet detection and instrument aberration retrieval in multispectral coronagraphic imaging," *Astron. Astrophys.* **551:A138** (2013).

[3] Herscovici-Schiller, O., Mugnier, L. M., and Sauvage, J.-F., "Analytical expression of a long exposure coronagraphic point spread function," *Optics Letters* (submitted).

[4] Roddier, F., "The effects of atmospheric turbulence in optical astronomy," *Progress in optics* **19**, 281–376 (1981).

[5] Soummer, R., Pueyo, L., Sivaramakrishnan, A., and Vanderbei, R. J., "Fast computation of lyot-style coronagraph propagation," *Optics Express* **15**, 15935–15951 (2007).

[6] Paul, B., Sauvage, J.-F., Mugnier, L. M., Dohlen, K., Petit, C., Fusco, T., Mouillet, D., Beuzit, J.-L., and Ferrari, M., "Compensation of high-order quasi-static aberrations on SPHERE with the coronagraphic phase diversity COFFEE," *Astron. Astrophys.* **572**, A32 (2014).

[7] Mugnier, L. M., Sauvage, J.-F., Fusco, T., Cornia, A., and Dandy, S., "On-line long-exposure phase diversity: a powerful tool for sensing quasi-static aberrations of extreme adaptive optics imaging systems," *Optics Express* **16**(22), 18406–18416 (2008).