Maximum likelihood approach for the adaptive optics point spread function reconstruction

J. Exposito[a], Damien Gratadour[a], Gérard Rousset, Yann Clénet, Laurent Mugnier[b] and Éric Gendron[a]

[a]LESIA, Observatoire de Paris, CNRS, UPMC, Université Paris-Diderot, 5 place Jules Janssen, 92195 Meudon, France;
[b]ONERA - The French Aerospace Lab, F-92322 Châtillon, France

ABSTRACT

This paper is dedicated to a new PSF reconstruction method based on a maximum likelihood approach (ML) which uses as well the telemetry data of the AO system (see Exposito et al. (2013)[1]). This approach allows a joint-estimation of the covariance matrix of the mirror modes of the residual phase, the noise variance and the Fried parameter r_0.

In this method, an estimate of the covariance between the parallel residual phase and the orthogonal phase is required. We developed a recursive approach taking into account the temporal effect of the AO-loop, so that this covariance only depends on the r0, the wind speed and some of the parameters of the system (the gain of the loop, the interaction matrix and the command matrix). With this estimation, the high bandwidth hypothesis is no longer required to reconstruct the PSF with a good accuracy.

We present the validation of the method and the results on numerical simulations (on a SCAO system) and show that our ML method allows an accurate estimation of the PSF in the case of a Shack-Hartmann (SH) wavefront sensor (WFS).

Keywords: Adaptive optics, PSF reconstruction, maximum likelihood

1. INTRODUCTION

To get the best image quality from an optical system in terms of contrast, it is required to deconvolve the image from the point spread function (PSF). For this method to be effective, it is necessary that the optical transfer function (OTF) of the system telescope + atmospheric residuals is well characterized. The optimal situation is to have a nearby star in the isoplanatic field around the science target which provides a PSF with the same correction and the same conditions of turbulence. The OTF can also be obtained by observing separately, but close in time and space, a star with similar characteristics than the science target to get an estimate of the PSF consistent with the observation. Unfortunately, this situation requires a large amount of integration time for the PSF calibration.

PSF reconstruction is another solution. The AO loop measures continuously the state of the turbulence. It is then possible to use a method to reconstruct the OTF (or the PSF) of the residual atmosphere during the observation using data from real-time telemetry. The reconstruction is done in post-processing and does not require extra time.

Véran et al. (1997)[2] developed the first method for PSF reconstruction and applied successfully the method to the PUEO system at CFHT. This method is based on a least-squares (LS) approach and requires an estimate of the noise variance and an estimate of the covariance of the aliasing on the mirror modes so as to unbias the covariance matrix of the modes. This method relies on the high bandwidth hypothesis and an accurate estimation of the noise variance. However, for a Shack-Hartmann (SH) wavefront sensor (WFS), the temporal sampling can be as low as 15 Hz and the noise variance estimation can be inaccurate and this can significantly impact

Further author information: (Send correspondence to J. Exposito)
J. Exposito: E-mail: jonathan.expositocano@obspm.fr

the quality of the reconstruction. Various studies successfully applied the LS method using a bright natural guide star (NGS) and a Shack-Hartmann (SH) wavefront sensor (WFS)(Harder & Chelli 2000[3] ; Jolissaint et al. 2012[4]). But the accuracy of this method is limited if the system is operating at low bandwidth. A second method was recently developed using a maximum likelihood (ML) approach (Exposito et al. 2013[1]). It was successfully tested using a simple linear model for the SH measurements without noise and this paper presents the improvement of the method using a model of the SH with diffraction. Our numerical simulations demonstrate that the ML method allows an accurate PSF estimation when the adaptive optics system is operating either at high or low temporal bandwidth and either at low or high photons level.

2. MODELS AND THEORETICAL APPROACH OF THE PSF RECONSTRUCTION

Let us define the pupil function of the telescope as:

$$P(\boldsymbol{x}) = \begin{cases} 1 & \text{si } \frac{d}{2} < |\boldsymbol{x}| < \frac{D}{2} \\ 0 & \text{sinon} \end{cases}$$

with D its diameter and d the diameter of its central obscuration.

The electric field in the image plane Ψ_{im} is generally expressed as a function of the Fourier transform of the electric field of the incident wave of the source $\Psi(\boldsymbol{x})$:

$$\Psi_{\text{im}}(\boldsymbol{\alpha}) \propto \int P(\boldsymbol{x})\Psi(\boldsymbol{x})e^{-\frac{2i\pi}{\lambda}\boldsymbol{x}\boldsymbol{\alpha}}d\boldsymbol{x} \tag{1}$$

where $\boldsymbol{\alpha}$ is the coordinates vector in the image plane and $\boldsymbol{x}$ the coordinates vector in the pupil plane, and λ the wavelength. The structure function of the residual phase is given by Roddier (1981):[5] The structure function of the phase is given by:

$$D_{\Phi\varepsilon}(\boldsymbol{x},\boldsymbol{\rho}) =< \left|\Phi_\varepsilon(\boldsymbol{x}) - \Phi_\varepsilon(\boldsymbol{x}+\boldsymbol{\rho})\right|^2 > \tag{2}$$

with $\boldsymbol{\rho}$ the vector of the separation in the pupil and Φ_ε the residual phase (i.e. corrected by the AO system).

The DM may compensate only for the mirror component of the turbulent phase. Thus the residual phase has the same high-order component as the turbulent phase. $\Phi_\varepsilon(\boldsymbol{x})$ is then decomposed such as

$$\Phi_\varepsilon(\boldsymbol{x}) = \Phi_{\varepsilon\parallel}(\boldsymbol{x}) + \Phi_{\varepsilon\perp}(\boldsymbol{x}). \tag{3}$$

The stationary phase approximation Conan (1994)[6] allows us to express the long exposure OTF as:

$$< B(\boldsymbol{\rho}) >\propto \underbrace{e^{-\frac{1}{2}\bar{D}_{\Phi\varepsilon}(\boldsymbol{\rho})}}_{B_{\text{atm}}} \underbrace{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})d\boldsymbol{x}}_{B_{\text{tel}}}, \tag{4}$$

where $\bar{D}_{\Phi\varepsilon}(\boldsymbol{\rho})$ is the structure function of the stationary residual phase and is given by:

$$\bar{D}_{\Phi\varepsilon}(\boldsymbol{\rho}) = \frac{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})D_{\Phi\varepsilon}(\boldsymbol{x},\boldsymbol{\rho})d\boldsymbol{x}}{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})d\boldsymbol{x}}, \tag{5}$$

B_{atm} is the atmospheric OTF and B_{tel} is the OTF of the telescope (i.e the autocorrelation of the pupil function). Several study show that the stationary phase approximation has a minor impact on the PSF reconstruction process (Véran at al. 1997,[2] Exposito et al. 2012[7]). We assume hereafter the validity of the approximation.

Using Eq. 3, $\bar{D}_{\Phi\varepsilon}(\boldsymbol{\rho})$ can be written as:

$$\bar{D}_{\Phi\varepsilon}(\boldsymbol{\rho}) = \bar{D}_{\Phi\varepsilon\parallel}(\boldsymbol{\rho}) + \bar{D}_{\Phi\perp}(\boldsymbol{\rho}) + 2\Gamma_\varepsilon(\boldsymbol{\rho}), \tag{6}$$

$$\Gamma_\varepsilon(\boldsymbol{\rho}) = \frac{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho}) < [\Phi_{\varepsilon\parallel}(\boldsymbol{x}) - \Phi_{\varepsilon\parallel}(\boldsymbol{x}+\boldsymbol{\rho})][\Phi_\perp(\boldsymbol{x}) - \Phi_\perp(\boldsymbol{x}+\boldsymbol{\rho})] > d\boldsymbol{x}}{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})d\boldsymbol{x}} \tag{7}$$

Simulations shows that the cross term $\Gamma_\varepsilon(\boldsymbol{\rho})$ as no effect in the PSF reconstruction process (Véarn et al. 1997[2] Exposito et al. 2012[7]). In the followings, this term will be considered as negligible.

The stationary structure function of the parallel phase $\bar{D}_{\Phi\varepsilon\|}(\boldsymbol{\rho})$ can be expressed with the covariance matrix of the modes as described in Gendron et al. (2006):[8]

$$\bar{D}_{\Phi\varepsilon\|}(\boldsymbol{\rho}) = \sum_{i=1}^{N} \boldsymbol{\Lambda}_{ii} V_{ii}(\boldsymbol{\rho}), \tag{8}$$

with $\boldsymbol{\Lambda}$ the eigenvalues matrix of $\mathbf{C}_{\varepsilon\|}$, the covariance matrix of the modes, such as $\mathbf{C}_\varepsilon = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{t}}$, N is the number of the mirror modes, and,

$$V_{ii}(\boldsymbol{\rho}) = \frac{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})[\mathbf{M}'_i(\boldsymbol{x}) - \mathbf{M}'_i(\boldsymbol{x}+\boldsymbol{\rho})][\mathbf{M}'_i(\boldsymbol{x}) - \mathbf{M}'_i(\boldsymbol{x}+\boldsymbol{\rho})]d\boldsymbol{x}}{\int P(\boldsymbol{x})P(\boldsymbol{x}+\boldsymbol{\rho})d\boldsymbol{x}}, \tag{9}$$

with $\mathbf{M}' = \mathbf{M}.\mathbf{V}$, the new mirror modes on which $\mathbf{C}_{\varepsilon\|}$ is diagonal and $\mathbf{M}$ the original mirror modes.

The long exposure PSF is then given by:

$$< K(\boldsymbol{\alpha}) > = \mathcal{F}^{-1}[< B(\boldsymbol{\rho}) >], \tag{10}$$

$$< B(\boldsymbol{\rho}) > = \underbrace{e^{-\frac{1}{2}\bar{D}_{\Phi\varepsilon\|}(\boldsymbol{\rho})}}_{<B_{\varepsilon\|}>} \times \underbrace{e^{-\frac{1}{2}\bar{D}_{\Phi\perp}(\boldsymbol{\rho})}}_{<B_{\perp}>} \times B_{\mathrm{tel}}. \tag{11}$$

$\bar{D}_{\Phi\perp}(\boldsymbol{\rho})$ is the structure function of the high-order turbulent phase not corrected by the DM. This component can be computed once for all by simulations for a given system and for a $D/r_0 = 1$ then scaled for a given r_0. In brief, to estimate the AO PSF, the knowledge of $\mathbf{C}_{\varepsilon\|}$ is crucial. The first method was developed by Véarn et al. (1997).[2] This method is based on a least-squares estimation of $\mathbf{C}_{\varepsilon\|}$ using the measurements provided by the WFS and successfully tested on PUEO/CFHT. However, a large temporal bandwidth of the system is necessary for the estimation to be accurate such as PUEO, a curvature WFS system running at 1 kHz. Moreover, the measurements are perturbed by the noise and the aliasing caused by the spatial sampling of the WFS. Thus, the noise must be measured to unbias the covariance matrix, and the covariance matrix of the aliasing must be simulated to be unbiased from $\mathbf{C}_{\varepsilon\|}$. The latter is dependent of the high frequency phase and thus dependent of an estimation of r_0. Previous works applied this method to a Shack-Hartmann WFS (Harder & Chelli 2000[3] ; Jolissaint et al. 2012[4]). However, SH WFS are able to run at frequencies as low as 15 Hz. In this case, the bandwidth of the system is not large enough to ensure an accurate estimation of $\mathbf{C}_{\varepsilon\|}$ using the LS method. Subsequently we present the new method developed, using a maximum likelihood approach, allowing us a accurate estimation of $\mathbf{C}_{\varepsilon\|}$.

3. THE MAXIMUM LIKELIHOOD APPROACH

The maximum likelihood estimation is based on the notion of plausibility of a given set of observations with respect to a model of the measurements:

$$\boldsymbol{w} = \mathbf{D}\boldsymbol{\varepsilon} + \boldsymbol{n}, \tag{12}$$

with $\boldsymbol{w}$ the vector of instant measurements given by the WFS, $\mathbf{D}$ the interaction matrix describing the linear behavior of the WFS, $\boldsymbol{\varepsilon}$ the vector of decomposition of the phase on the modes of the phase and $\boldsymbol{n}$ the vector of instant noise on the measurements. The noise is explicitly taken into account in this equation. The effects of the aliasing in the measurements is taken into account by considering a large amount of modes in the interaction matrix $\mathbf{D}$. The interaction matrix is built by recording the response of the WFS to a mirror mode. In this approach, we extend the number of modes to include in the interaction matrix the theoretical response from the WFS to higher order than the mirror modes so that the aliasing is explicitly taken into account in the measurements.

$\boldsymbol{w}$ is a random variable following the normal distribution with zero mean. The probability density function is given by:

$$f(\boldsymbol{w}) = \frac{1}{(2\pi)^{\frac{N}{2}}[\det(\mathbf{C}_w)]^{\frac{1}{2}}} \exp[-\frac{1}{2}\boldsymbol{w}^{\mathrm{t}}\mathbf{C}_w^{-1}\boldsymbol{w}] \tag{13}$$

Following the model given Eq. 12, the covariance of the measurements is given by:

$$\mathbf{C}_w =< \boldsymbol{w}\boldsymbol{w}^{\mathrm{t}} >= \mathbf{D}\mathbf{C}_\varepsilon\mathbf{D}^{\mathrm{t}} + \mathbf{C}_n \tag{14}$$

The likelihood function associated to the probability density function is defined as:

$$\mathcal{L}(\boldsymbol{w}|\mathbf{C}_\varepsilon, \sigma_n^2) = \frac{1}{(2\pi)^{\frac{N}{2}}[\det(\mathbf{C}_w)]^{\frac{1}{2}}} \exp[-\frac{1}{2}\boldsymbol{w}^{\mathrm{t}}\mathbf{C}_w^{-1}\boldsymbol{w}] \tag{15}$$

Moreover, assuming the independency of each vector in the temporal sample, the likelihood is given by:

$$\mathcal{L}(\boldsymbol{w}_i|\mathbf{C}_\varepsilon, \sigma_n^2) = \frac{1}{(2\pi)^{\frac{nN}{2}}[\det(\mathbf{C}_w)]^{\frac{n}{2}}} \prod_{i=1}^{n} \exp[-\frac{1}{2}\boldsymbol{w}_i^{\mathrm{t}}\mathbf{C}_w^{-1}\boldsymbol{w}_i] \tag{16}$$

where n is the number of vectors of measurements of the series and i is the iteration at which the measurement vector $\boldsymbol{w}$ was acquired. Of course, the measurements are not strictly statistically independent. Correction of the AO system being partial, measurement at iteration i is dependent on the measurement at iteration i?1. However, the correction strongly attenuates low spatial and temporal frequencies of the phase which, in the case of the frozen phase, reduces the correlation time. Therefore, between two successive iterations, there is no statistical independence between two vectors of measurements, although the dependence is low, but between iteration i and $i+2$, the correlation is very low. The likelihood function given above is a valid approximation. This function can be modified if necessary to take into account the statistical dependence. We aim to maximize the probability of obtaining the covariance of the measurements relative to the elements of $\mathbf{C}_\varepsilon$, σ_{n}^2 et r_0. That is to minimize (using an algorithm minimizing multivariate functions):

$$\begin{aligned}
\mathcal{J}(\mathbf{C}_\varepsilon, \sigma_n^2) = &- \ln(\mathcal{L}) = \frac{nN}{2}\ln(2\pi) + \frac{n}{2}\ln|\det[\mathbf{C}_w]| \\
&+ \frac{1}{2}\sum_{i=1}^{n} \boldsymbol{w}_i^{\mathrm{t}}\mathbf{C}_w^{-1}\boldsymbol{w}_i \\
&= \frac{nN}{2}\ln(2\pi) + \frac{n}{2}\ln\left|\det[\mathbf{D}\mathbf{C}_\varepsilon\mathbf{D}^{\mathrm{t}} + \mathbf{C}_n]\right| \\
&+ \frac{1}{2}\sum_{i=1}^{n} \boldsymbol{w}_i^{\mathrm{t}}(\mathbf{D}\mathbf{C}_\varepsilon\mathbf{D}^{\mathrm{t}} + \mathbf{C}_n)^{-1}\boldsymbol{w}_i \\
\Rightarrow \mathcal{J}(\mathbf{C}_\varepsilon, \sigma_n^2) \propto &\ N\ln(2\pi) + \ln|\det[\mathbf{C}_w]| \\
&+ \mathrm{Tr}[\left(\mathbf{D}\mathbf{C}_\varepsilon\mathbf{D}^{\mathrm{t}} + \mathbf{C}_n\right)^{-1}\mathbf{C}_w].
\end{aligned} \tag{17}$$

Finally, we chose to model the variance of the noise with 1 scaling parameter given the relative noise level per subaperture ($\mathbf{C}_n^{\mathrm{rel}}$) such as:

$$\mathbf{C}_n = \alpha\mathbf{C}_n^{\mathrm{rel}}. \tag{18}$$

Moreover, C_n^{rel} is adjusted so that wether $\alpha = 1$, $\mathbf{C}_n = C_n^{\mathrm{rel}}$. So, α is a direct measure of the relative error on the noise estimation in our simulations.

3.1 The cross term and the large bandwidth approximation

One of the limitation of the LS method is the large bandwidth approximation. The approximation consists in simplifying the coupling terms between the residual parallel phase and the aliasing term in the mirror. This approximation is verified for a system operating at high frequency such as PUEO but a SH WFS is able to operate at a frequency as low as 15 Hz, such as NaCo/VLT, and the approximation is no more verified. In the ML method, this term is explicitly taken into account through the term $\mathbf{C}_{\varepsilon\parallel\perp} =< \boldsymbol{\varepsilon}_{\parallel}\boldsymbol{\varepsilon}_{\perp}^{\mathrm{t}} >$ and a recursive method was developed in Exposito et al. (2013)[1] to compute this term with a good precision. The ML method for PSF reconstruction is thus robust even at low frequency of the loop.

3.2 Modal basis

The choice of the modal basis is essential in this approach. Our aim is to estimate the covariance matrix of the modes of the phase which minimizes the likelihood criterion. A limited number of parameters to estimate is required in any method of minimization. We must find a modal basis which reduces the number of elements in $\mathbf{C}_{\varepsilon}$.

We chose to place this work in the context of the Canary demonstrator in its single conjugated adaptive optics (SCAO) configuration where about 30 modes are corrected by the deformable mirror. For the simulation purpose, we describe the turbulence on the 200 first Karhunen-Loève (KL) modes which are orthonormal and statistically independent. We consider that the deformable mirror is able to compensate for the first 30 modes. Thus for the modes 31 to 200, the matrix $\mathbf{C}_{\varepsilon}$ is a diagonal matrix.

3.3 The covariance matrix of the modes $\mathbf{C}_{\varepsilon}$

The simulated phase, at each iteration, is decomposed on the basis of atmospheric KL thanks to the projector (in the least squares senses) from the phase to the modes, defined as,

$$\Phi_{\varepsilon} = \mathbf{K}\boldsymbol{\varepsilon}, \tag{19}$$

$$\mathbf{K}^{+} = (\mathbf{K}^{\mathrm{t}}\mathbf{K})^{-1}\mathbf{K}^{\mathrm{t}} \tag{20}$$

where $\mathbf{K}$ is the matrix defining the surfaces of the modes, such that,

$$\mathbf{K}^{+}\Phi_{\varepsilon} = (\mathbf{K}^{\mathrm{t}}\mathbf{K})^{-1}\mathbf{K}^{\mathrm{t}}(\mathbf{K}\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}. \tag{21}$$

The vector $\boldsymbol{\varepsilon}$ is the vector of the coefficients of the decomposition of the phase (Φ_{ε}) on the KL. The covariance matrix is then defined such that

$$\mathbf{C}_{\varepsilon} =< \boldsymbol{\varepsilon}^{\mathrm{t}}\boldsymbol{\varepsilon} > . \tag{22}$$

This matrix is represented in part on the figure 1. It is divided into four quadrants, limited by the number of corrected modes (30 modes):

- $\mathbf{C}_{\perp}$: this term represents the correlations between modes of high order non-corrected by the mirror. The KL modes are statistically independent, the covariance matrix is diagonal. Assuming a Kolmogorov phase, the covariance of these modes depends only on the parameter r_0. This term can then be calculated once and for all by simulation assuming a $D/r_0 = 1$ and then infer the covariance for a particular case:

$$\mathbf{C}_{\perp}\Big|_{\frac{D}{r_0}} = \mathbf{C}_{\perp}\Big|_{\frac{D}{r_0}=1} \times \left(\frac{D}{r_0}\right)^{5/3}, \tag{23}$$

- $\mathbf{C}_{\varepsilon\parallel}$: In the presence of the correction, correlations between modes appear. In the mirror space, the covariance matrix is not diagonal. These terms are contained in the subspace generated by mirror modes an contain thereby the residual parallel component, the aliasing in the measurements wrongly reconstructed with the mirror and the propagation of noise on the mirror. Furthermore, they depend of the correction provided by the system. These are the terms that the method is primarily intended to estimate

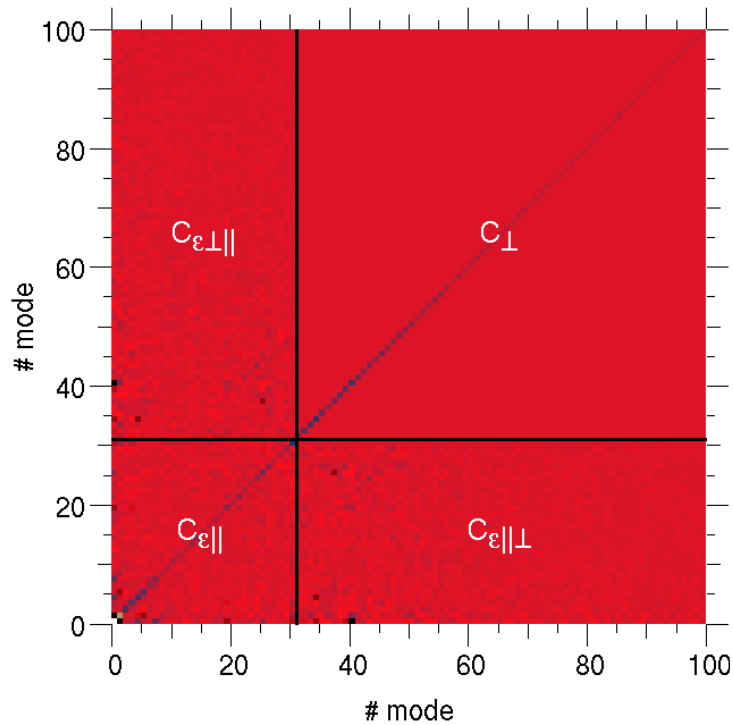

Figure 1. Covariance matrix of the first 100 KL modes showing the four quadrants. The complete matrix contains 200 modes.

- $\mathbf{C}_{\varepsilon_{\parallel\perp}}$: the cross term is the coupling between the modes of high spatial frequencies and modes of the mirror space (corrected) in the covariance matrix of the modes. Out of correction, these terms should be zero because the orthogonal phase is statistically independent of the parallel phase. However, because of the high spatial frequency aliased on the sensor, erroneous commands are sent to the DM that introduces low spatial frequencies in phase. For the estimation of $\mathbf{C}_{\varepsilon_{\parallel}}$ it is mandatory to have an estimation of the terms $\mathbf{C}_{\varepsilon_{\parallel\perp}}$ because the aliasing has a strong impact on the high orders of $\mathbf{C}_{\varepsilon_{\parallel}}$ and therefore the covariance measurements. This term can be computed with a recursive approach such as describe in Exposito et al. (2013).[1]

- $\mathbf{C}_{\varepsilon_{\perp\parallel}}$ is the transposed matrix of $\mathbf{C}_{\varepsilon_{\parallel\perp}}$: and has the same properties.

  In Exposito et al. (2013)[1] we have demonstrated that the cross term can be fully computed using turbulence spatial and temporal covariance matrices on the KL modes. These covariance matrices only depend on r_0 and wind speed considering the Kolmogorov and Taylor models. But in the present paper, we will use the numerical simulation to determine the cross term.

3.4 Parameter space

The matrix $\mathbf{C}_{\varepsilon}$ consists in 200x200 elements and is symmetrical. Only one parameter is necessary to estimate $\mathbf{C}_{\varepsilon_{\perp}}$ and $\mathbf{C}_{\varepsilon_{\parallel\perp}}$: r_0.

However, each non redondant element of $\mathbf{C}_{\varepsilon_{\parallel}}$ have to be estimated, including the noise variance, to minimize the ML criterion. Thus, the parameter space consists in 1 parameter (r_0) to scale $\mathbf{C}_{\varepsilon_{\perp}}$ and $\mathbf{C}_{\varepsilon_{\parallel\perp}}$, 1 parameter to scale the noise variance (Eq. 18) and $N(N+1)/2 = 465$ with $N = 30$ mirror modes which corresponds to $\mathbf{C}_{\varepsilon_{\parallel}}$. Hence, the total number of parameters is 467.
We use a Minimum Variance Bounds (Nocedal 1980[9]) algorithm to minimize the criterion

In brief, the parameters used to minimize the criterion are:

- r_0, acting on $\mathbf{C}_{\varepsilon\perp}$ and $\mathbf{C}_{\varepsilon\parallel\perp}/\mathbf{C}_{\varepsilon\perp\parallel}$,
- $\mathbf{C}_{\varepsilon\parallel ij}$, the covariances of the parallel modes,
- α, the scaling factor of the noise variance,

The constants used are:

- $\mathbf{C}_{\varepsilon\perp}\Big|_{\frac{D}{r_0}=1}$ which is scaled with respect to r_0 (Eq. 23),
- $\mathbf{C}_w$, the covariance matrix opt the on-sky measurements,
- $\mathbf{C}_{\varepsilon\parallel\perp}$ which is scaled with respect to r_0 and $\mathbf{C}_{\varepsilon\parallel ii}$, the variances of the parallel modes,
- and the system matrices: $\mathbf{D}_\infty$, $\mathbf{M}$, $\mathbf{D}^+$.

4. MODEL TUNING

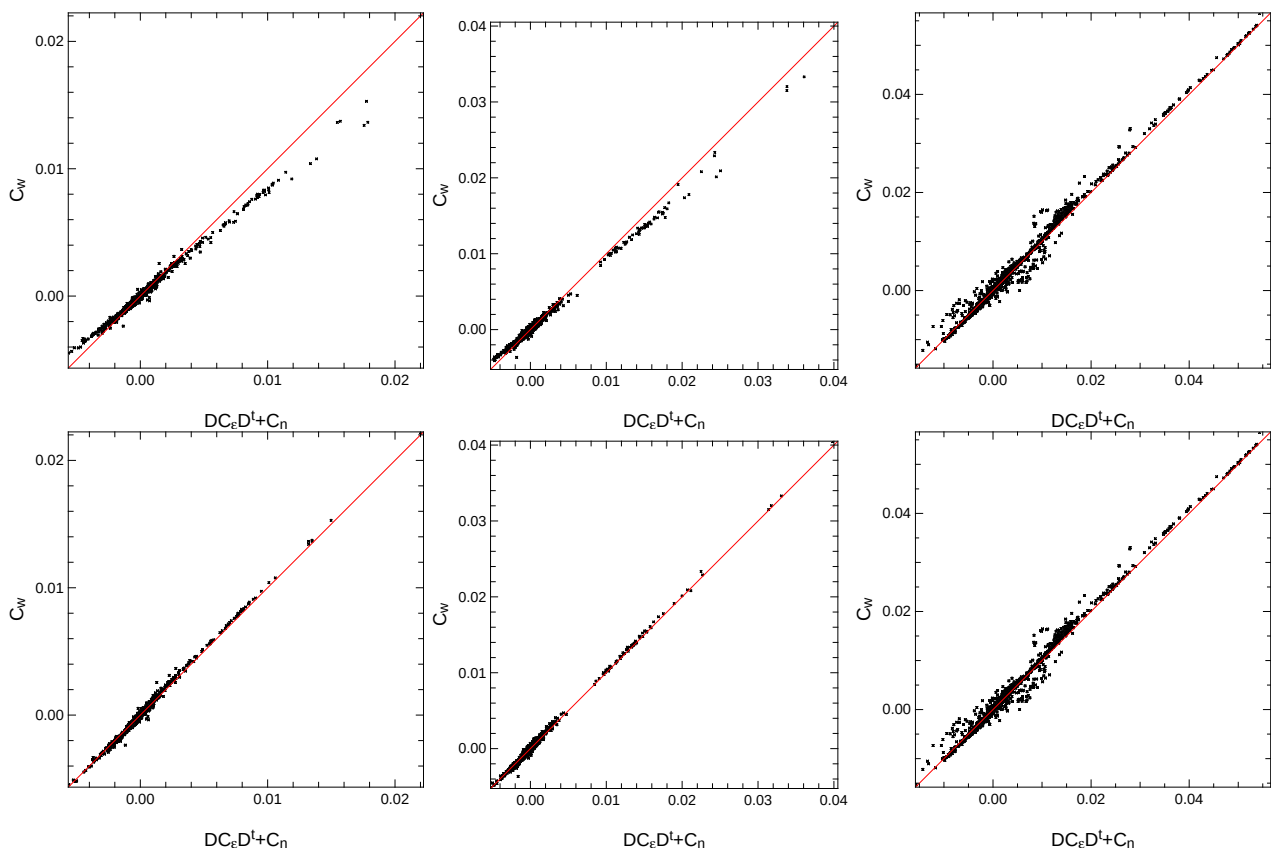

Figure 2. Covariance matrix of the measurements versus model of measurements with and without taking into account the centroid gain in the interaction matrix, respectively on top and bottom. *Left*: the NGS is at magnitude 0 and the loop is operating at 150 Hz, *middle*: the NGS is at magnitude 12 and the loop is operating at 150 Hz and *right*: NGS at magnitude 12 and the loop is operating at 15 Hz.

The maximum likelihood estimation is based on the notion of plausibility of a given set of observations with respect to a model. The model of measurements must be as representative as possible of the real measurements. In order to assess the validity of the model, all the values of $\mathbf{C}_w$ can be represented as a function of the corresponding values of the model (Eq. 12). Doing so, we obtain the scatter plot on top of Fig. 2. It shows that

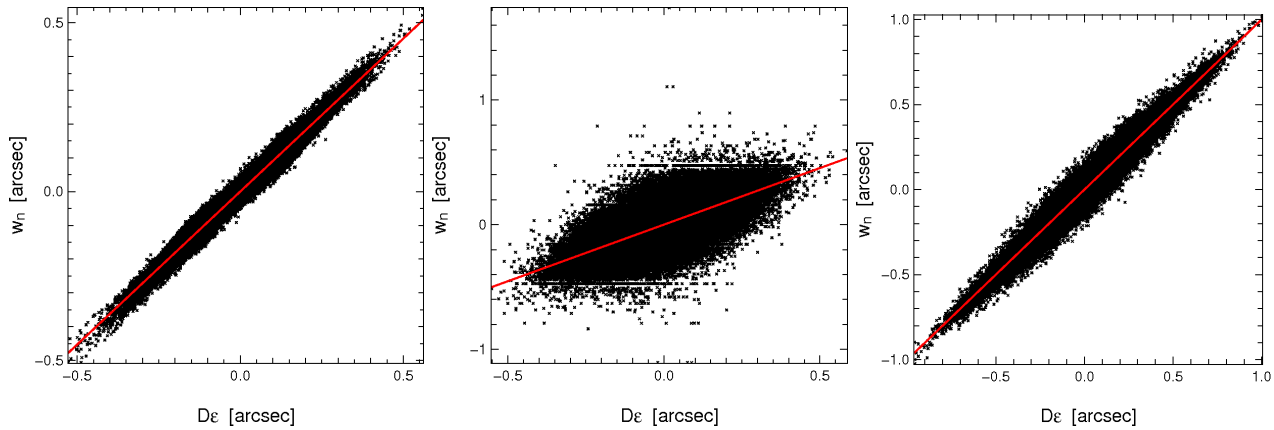

Figure 3. Estimation of the centroid gain with a linear regression between measurements and the SH model.

the calibration method of **D** used in the simulation is not fully representative of the on-sky measurements because speckles on-sky are opposed to the ideal calibration source). This error introduces a gain of which the model has to be compensated to provide a high fidelity representation of the measurements. This gain is equivalent to a centroid gain (CG) which can be estimated on sky as describe in Gratadour & Rigaut (2007).[10]

In this work, we determine the CG by a linear regression between the measurements and the model of SH, as shows in Fig. 3. We then compensate the centroid gain, subaperture by subaperture, from the interaction matrix. We present in the bottom of the Fig. 2 the results of the compensation on the model of measurements. We can clearly see that taking into account the CG provides a more accurate model. At 15 Hz, we can note that taking into account the centroid gain has a weak impact on the model.

5. SIMULATIONS

This work is placed in the context of the Canary project (Gendron et al. 2011[11]) installed at William Herschel Telescope. Canary is working in *open loop* but it can also work in SCAO. The simulation was configured to simulate the closed loop system. The detail of the simulation parameters are presented in the table. 1

The simulated system is composed of a WFS Shack-Hartmann 7x7 subapertures observing at a wavelength of 589 nm. The deformable mirror is capable of reproducing the first thirty Karhunen-Loève atmospheric modes. The telescope has a diameter of 4.2 m and a central obstruction of 1.20 m or 0.286 in fraction of the pupil diameter.

This system observes a natural guide star on the axis, subject to a turbulent layer with a $r_0 = 0.12\ m$ and a wind speed of 10 m s^{-1}. For these tests, the phase is filtered of very high order modes. Thus, the phase is a linear combination of 200 KL. Phase screens are infinite size as described in Assémat et al.(2006)[12] to avoid cyclic redundancy in measurements.

Concerning the loop, the simulated system operates at a frequency of 150 Hz (maximum) with a gain $g = 0.4$ for 5000 iterations (a time comparable to a real observation exposures). These parameters can be varied to estimate the performance of the ML method based on the bandwidth of the system. The frame delay parameter is set to 0 in the simulation but Yao considers that there is a basic frame delay between the analysis and the application of the commands to the deformable mirror. Hence, the overall frame delay is 1.

At the end of the simulation, it retrieves the data ($\boldsymbol{w}$), the commands send to the DM ($\boldsymbol{V}$) and the differential commands ($\Delta\boldsymbol{V}$) at each iteration, and the average PSF.

Yao has been modified so that it can also output the turbulent phase $\boldsymbol{a}$ at each iteration, the decomposition of the residual phase $\boldsymbol{\varepsilon}$ and the aliasing on measurements(by injecting the high-order phase in the WFS during the simulation). This allows us to get all the variables necessary to perform our tests.

Table 1. Table of the parameters of the simulations. The hacked parameters correspond to modifications added in Yao for this work.

	Parameters	Value
Atmosphere	r_0 @ 0.5 μm	0.12 m
	Wind speed	7.5 m s^{-1}
	Altitude of the turbulent layer	0 m
WFS	Type	diffraction SH
	Wavelength	0.589 μm
	Number of subapertures	7x7 (36 lightened)
	Pixel size on sky	0.3”
	Pixels per subapertures	16x16
	Noise	Photon noise only
DM	Type	KL
	Number of mirror modes	30 (including tip-tilt)
Telescope	Diameter	4.2 m
	Central obscuration	0.25
	Imaging wavelength	1.65 μm
Loop	Frequency of the loop	150–15 Hz
	Number of iterations	5000
	Yao frame delay	0 (equivalent 1 frame)
Yao outputs	$\boldsymbol{w}$	-
	$\boldsymbol{V}$	-
	$\Delta\boldsymbol{V}$	-
	$\mathbf{D}$	-
	Average PSF	-
	Command matrix	-
Hacked outputs	$\boldsymbol{a}$	-
	$\boldsymbol{\varepsilon}$	-
	$\bar{D}_{\Phi_\perp}$	-

6. PSF RECONSTRUCTION RESULTS

We simulate an observation with the parameters described in Tab. 1. This simulation gives us the measurements vectors $\boldsymbol{w}$ nd the on-sky PSF. We also recover the simulated covariance matrix of the modes $\mathbf{C}_\varepsilon$.

Following this, we perform a second simulation using different screens of phase, but with the same r_0. This method requires constants presented in subsec. 3.4. With this second simulation, we extract $\mathbf{C}_{\varepsilon\parallel\perp}$ and $\mathbf{C}_{\varepsilon\perp}$ of the residual phase normalized at $D/r_0 = 1$.
We also extract the first guess for the minimization process from the second simulation.

Then, we estimate with the ML method the covariance matrix of the modes $\mathbf{C}_\epsilon^{\mathrm{ML}}$ by minimizing the criterion with respect to the measurements provided by the first simulation.

Finally, we reconstruct the PSF using the latter covariance and we compare it with the on-sky PSF and the PSF reconstructed with the hacked covariance matrix of the modes $\mathbf{C}_\epsilon^{\mathrm{exact}}$ from the first simulation (i.e. synchronized with the observation). Thus, the difference between the on-sky pSF and the latter is mainly the convergence error and the effects of the different approximations. The convergence error with 5000 iteration is about 0.01 on the OTF.

We present results on three cases: one case with a NGS at magnitude 0 so that the noise on the measurements is negligible and a loop operating at 150 Hz, one case with a NGS at magnitude 12 leading to a noise level at $\sim$200 μm RMS and a loop at 150 Hz, and one case with a NGS at magnitude 12 but a low frequency loop, i.e. 15 Hz, leading to small temporal bandwidth.

For each result, we show the X-cut and the Y-cut of the PSFs as well as the OTF. In black, we show the on-sky PSF, in red the ML estimated PSF and in green the PSF reconstructed using $\mathbf{C}_\epsilon^{\text{exact}}$. In dashed lines, we show the corresponding residuals i.e. the absolute difference between the on-sky PSF and the reconstructed. We also show the corresponding images, the on-sky PSF in the left, the ML estimated in the middle and the PSF reconstructed using $\mathbf{C}_\epsilon^{\text{exact}}$ in the right and the corresponding residuals in the bottom i.e. the absolute difference between the on-sky PSF and the reconstructed.

6.1 NGS at magnitude 0 and loop at 150 Hz

With a NGS at magnitude 0, the signal-to-noise ratio in each subaperture is extremely high. The only limitations on the performances of the AO is the fitting error and the aliasing. In this case only, we set the parameter $\alpha = 0$ so that the noise is not taken into account in the minimization of the ML criterion. Result on the reconstruction is presented on Fig. 4. We find an estimated $r_0 = 11.80$ cm, compared to the 12 cm simulated. The on-sky Strehl ratio is $\text{SR}_{\text{on-sky}} = 45.6\%$, the ML estimated $\text{SR}_{\text{ML}} = 42.5\%$ and the Strehl ratio find by using the matrix $\mathbf{C}_\epsilon^{\text{exact}}$ in the reconstruction process is $\text{SR}_{\text{exact}} = 44.7\%$.

The PSF is estimated with a fairly good precision as shown in Fig. 4. The rings and the wings are well reproduced as well as the global shape of the PSF in the images.

However, it exists a residual error of about 3% on the SR estimation. This error is not caused by the error on the estimated r_0 (i.e. an error of 0.2 cm) or by the approximation $alpha = 0$ corresponding to a non-noisy SH. This error is caused by a local minimum of the criterion which stops the convergence of the minimization.

6.2 NGS at magnitude 12 and loop at 150 Hz

For this case, the noise is no more negligible and is taken into account. We estimate the Fried parameter at $r_0 = 12.20$ cm and the noise is 3.35% over-estimated. Withal, the PSF is fairly accurate. Such as the first case, the rings and the wings are well reproduced as well as the global shape of the PSF in the images. The error on the Strehl ratio is close to the convergence error. The on-sky Strehl ratio is $\text{SR}_{\text{on-sky}} = 40.7\%$, the ML estimated $\text{SR}_{\text{ML}} = 39.4\%$ and the Strehl ratio find by using the matrix $\mathbf{C}_\epsilon^{\text{exact}}$ in the reconstruction process is $\text{SR}_{\text{exact}} = 39.9\%$. The results are presented in Fig. 5.

6.3 NGS at magnitude 12 and loop at 15 Hz

With a frequency of 15 Hz, the correction performances are close to the atmospheric conditions. However, we obtain with the ML method a really good accuracy on the reconstruction. We estimate the Fried parameter at $r_0 = 12.80$ cm and the noise is 24.7% under-estimated.

Since the variance of the turbulence is high because of the low frequency of the loop, the variance of the noise on the measurements is negligible. Thus, the important error found on the noise estimation is weak compared to the turbulent measurements. The on-sky Strehl ratio is $\text{SR}_{\text{on-sky}} = 2.89\%$, the ML estimated $\text{SR}_{\text{ML}} = 2.85\%$ and the Strehl ratio find by using the matrix $\mathbf{C}_\epsilon^{\text{exact}}$ in the reconstruction process is $\text{SR}_{\text{exact}} = 2.81\%$.

Despite those extreme conditions, the PSF reconstruction provides a PSF with a good precision. We residuals in the images shown in Fig. 6 consists mainly in speckles.

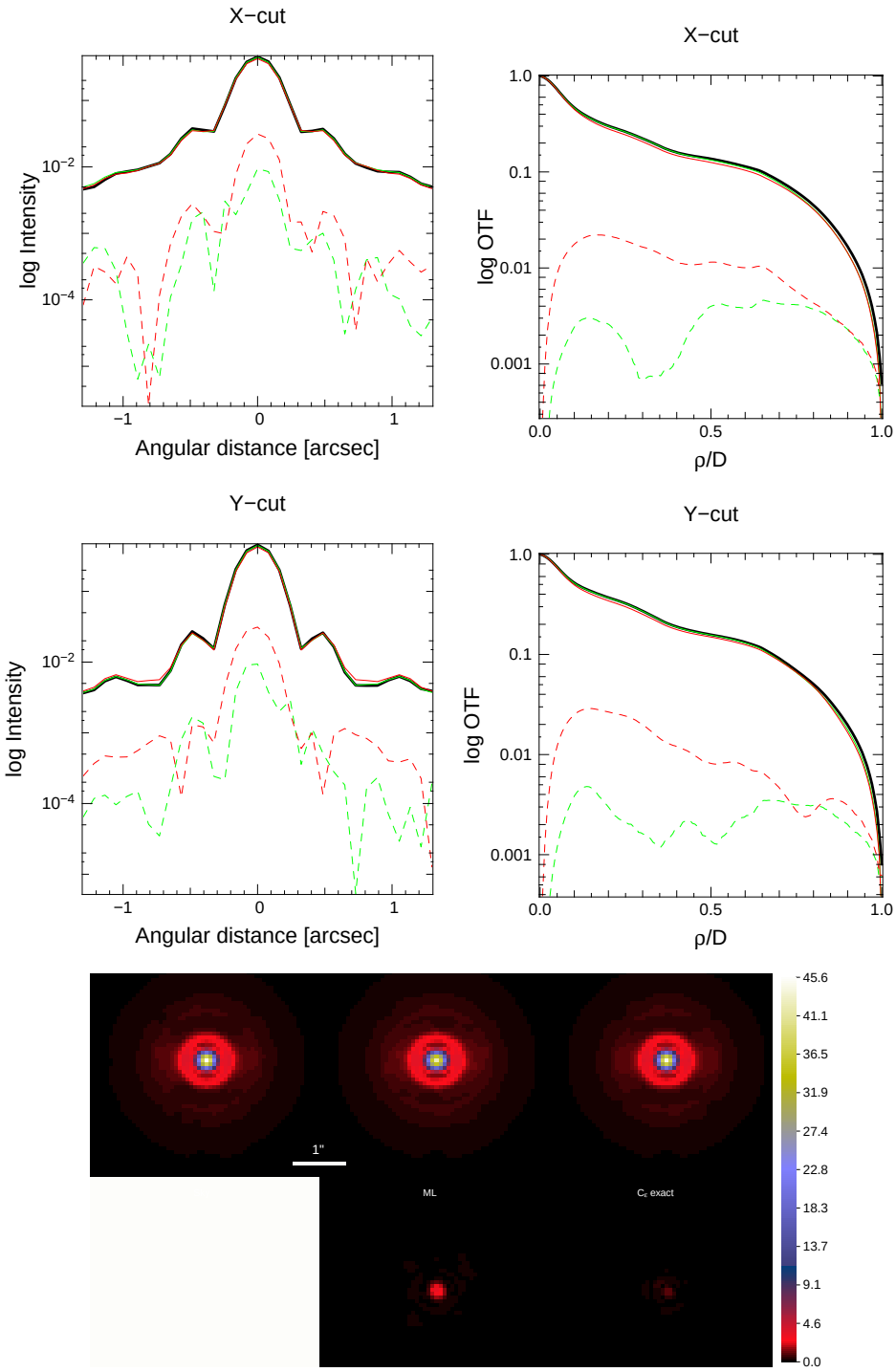

Figure 4. Comparison between PSF (lines) on-sky (*black*), ML reconstructed (*red*) and reconstructed with the exact $\mathbf{C}_\varepsilon$ matrix (*green*) and the residuals (dashed) for each reconstruction (ML: *red*, exact $\mathbf{C}_\varepsilon$: *green*) with respect to the on-sky PSF. In the top-left, the X-cut and Y-cut of the PSF (log scale). In the top-right, the X-cut and Y-cut of the OTF (log scale). The simulation was operated with a temporal frequency of 150 Hz for a NGS at magnitude 0. The Strehl ratio is given by the maximum of intensity of the PSF. In the bottom, we present the images of the PSF on-sky (left), ML reconstructed (middle), and the reconstructed with the exact $\mathbf{C}_\varepsilon$ (right) and the corresponding residuals (bottom).

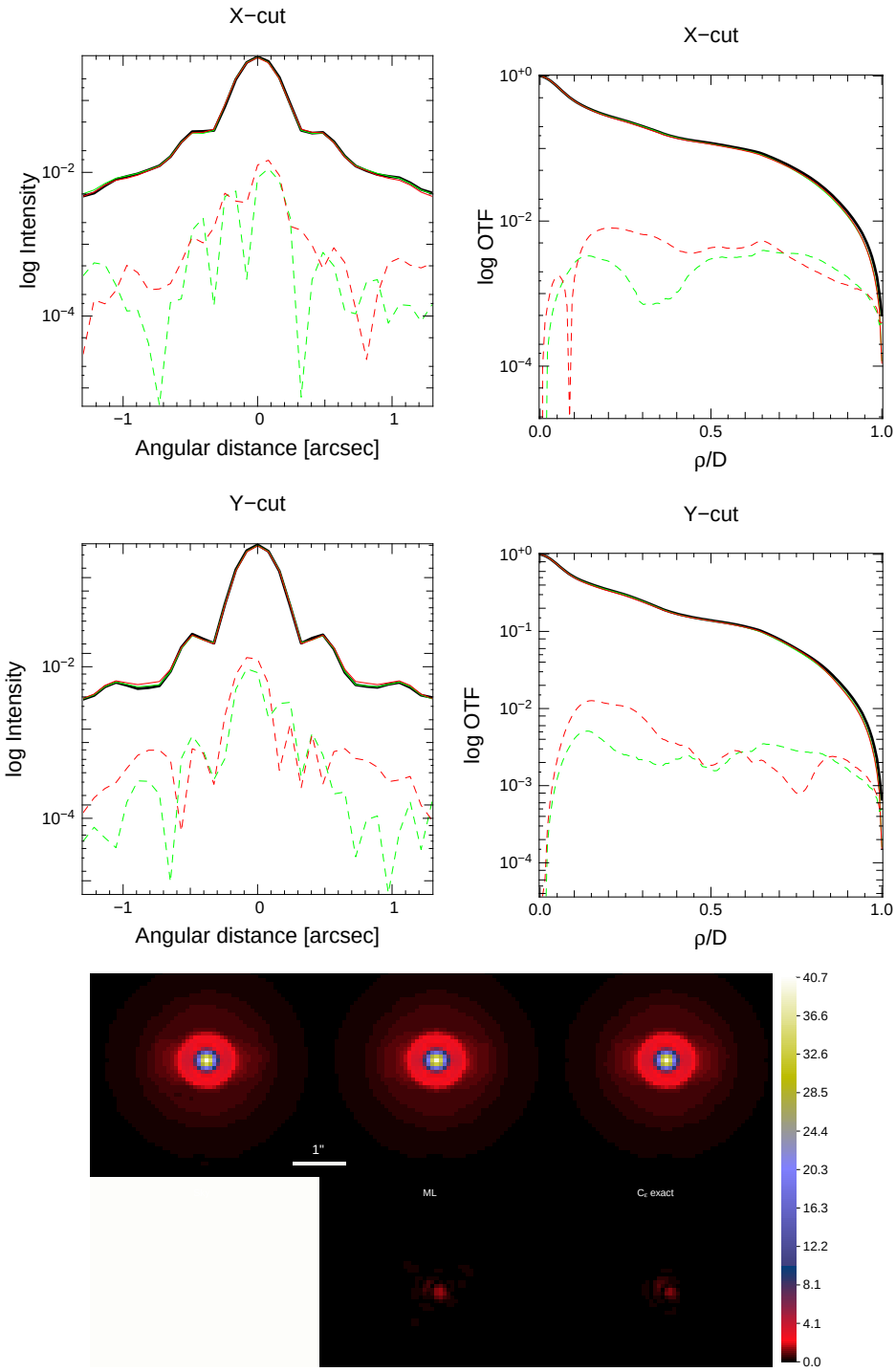

Figure 5. Comparison between PSF (lines) on-sky (*black*), ML reconstructed (*red*) and reconstructed with the exact $\mathbf{C}_\varepsilon$ matrix (*green*) and the residuals (dashed) for each reconstruction (ML: *red*, exact $\mathbf{C}_\varepsilon$: *green*) with respect to the on-sky PSF. In the top-left, the X-cut and Y-cut of the PSF (log scale). In the top-right, the X-cut and Y-cut of the OTF (log scale). The simulation was operated with a temporal frequency of 150 Hz for a NGS at magnitude 12. The Strehl ratio is given by the maximum of intensity of the PSF. In the bottom, we present the images of the PSF on-sky (left), ML reconstructed (middle), and the reconstructed with the exact $\mathbf{C}_\varepsilon$ (right) and the corresponding residuals (bottom).

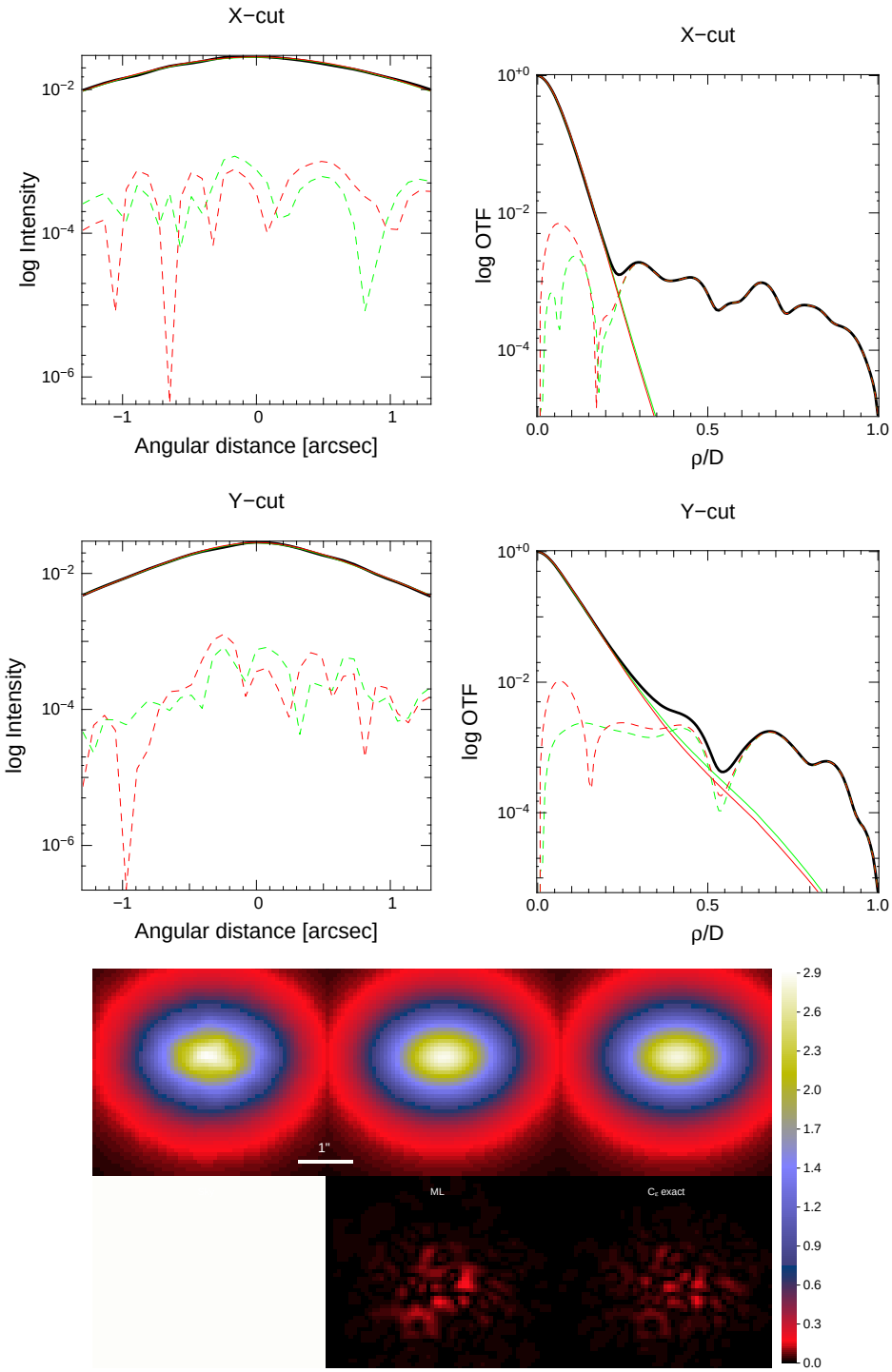

Figure 6. Comparison between PSF (lines) on-sky (*black*), ML reconstructed (*red*) and reconstructed with the exact $\mathbf{C}_\varepsilon$ matrix (*green*) and the residuals (dashed) for each reconstruction (ML: *red*, exact $\mathbf{C}_\varepsilon$: *green*) with respect to the on-sky PSF. In the top-left, the X-cut and Y-cut of the PSF (log scale). In the top-right, the X-cut and Y-cut of the OTF (log scale). The simulation was operated with a temporal frequency of 15 Hz for a NGS at magnitude 12. The Strehl ratio is given by the maximum of intensity of the PSF. In the bottom, we present the images of the PSF on-sky (left), ML reconstructed (middle), and the reconstructed with the exact $\mathbf{C}_\varepsilon$ (right) and the corresponding residuals (bottom).

7. CONCLUSION & PERSPECTIVE

In this paper, we presented the improvement to the maximum likelihood method for the adaptive optics point spread function reconstruction. We simulated a diffraction Shack-Hartmann wavefront sensor as close as possible to a real system. We were confronted to the interaction matrix calibration issue. Indeed, significant differences between the on-sky measurements and our model of measurements were identified as caused by a centroid gain. We estimated this centroid gain by a linear regression using the model and the measurements and we compensate the interaction matrix from this gain.

Although the method is only tested on simulations, the results obtained are very promising. In the presented cases, we estimate with a good accuracy:

- the Strehl ratio, with a precision lower than 3%, even for short bandwidth systems. The global shape of the PSF is as well reconstructed with a fairly good precision.
- The Fried parameter, with a precision lower than 1 cm,
- and the noise level when the latter is significant.

In the future, we aim to apply the ML method to real data obtained with the Multi Object Adaptive Optics demonstrator Canary operating in its close-loop mode.

REFERENCES

[1] Exposito, J., Gratadour, D., Rousset, G., Clenet, Y., Gendron, E., and Mugnier, L., "A new method for adaptive optics point spread function reconstruction," in [*Proceedings of the Third AO4ELT Conference*], Esposito, S. and Fini, L., eds. (Dec. 2013).

[2] Veran, J.-P., Rigaut, F., Maitre, H., and Rouan, D., "Estimation of the adaptive optics long-exposure point-spread function using control loop data.," *J. Opt. Soc. Am. A, Vol. 14, No. 11, p. 3057 - 3069* **14**, 3057–3069 (Nov. 1997).

[3] Harder, S. and Chelli, A., "Estimating the point spread function of the adaptive optics system ADONIS using the wavefront sensor measurements," *A&A Supplement* **142**, 119–135 (Feb. 2000).

[4] Jolissaint, L., Neyman, C., Christou, J., and Wizinowich, P., "Adaptive optics point spread function reconstruction project at W. M. Keck Observatory: first results with faint natural guide stars," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **8447** (July 2012).

[5] Roddier, F., "Atmospheric limitations to high angular resolution imaging," in [*Scientific Importance of High Angular Resolution at Infrared and Optical Wavelengths*], Ulrich, M. H. and Kjaer, K., eds., 5–22 (1981).

[6] Conan, J.-M., *Etude de la correction partielle en optique adaptative*, PhD thesis, Ph. D. thesis, Université Paris XI Orsay, (1994) (1994).

[7] Exposito, J., Gratadour, D., Clénet, Y., Rousset, G., and Mugnier, L., "Estimation of errors on the PSF reconstruction process for myopic deconvolution," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **8447** (July 2012).

[8] Gendron, E., Clénet, Y., Fusco, T., and Rousset, G., "New algorithms for adaptive optics point-spread function reconstruction," *A&A* **457**, 359–363 (Oct. 2006).

[9] Nocedal, J., "Updating quasi-newton matrices with limited storage," *Mathematics of Computaion* **35**, 773–782 (1980).

[10] Gratadour, D. and Rigaut, F., "Online centroid gain determination for lgs ao systems," in [*Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM*], *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM* **PMA4**, PMA4, Optical Society of America (2007).

[11] Gendron, E., Vidal, F., Brangier, M., Morris, T., Hubert, Z., Basden, A., Rousset, G., Myers, R., Chemla, F., Longmore, A., Butterley, T., Dipper, N., Dunlop, C., Geng, D., Gratadour, D., Henry, D., Laporte, P., Looker, N., Perret, D., Sevin, A., Talbot, G., and Younger, E., "MOAO first on-sky demonstration with CANARY," *A&A* **529**, L2 (May 2011).

[12] Assémat, F., Wilson, R., and Gendron, E., "Method for simulating infinitely long and non stationary phase screens with optimized memory storage," *Optics Express* **14**, 988–999 (Feb. 2006).