



HAL
open science

Optimal a posteriori fringe tracking in optical interferometry

Ferréol, Soulez, Éric Thiébaud, Michel, Tallon, Isabelle, Tallon-Bosc, Paulo Garcia

► To cite this version:

Ferréol, Soulez, Éric Thiébaud, Michel, Tallon, Isabelle, Tallon-Bosc, Paulo Garcia. Optimal a posteriori fringe tracking in optical interferometry. SPIE Optical and Infrared Interferometry IV., SPIE, 2014, Montreal, Canada. pp.91462Y, <10.1117/12.2056590>. <hal-01184508>

HAL Id: hal-01184508

<https://hal.science/hal-01184508v1>

Submitted on 13 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/269320226>

Optimal a posteriori fringe tracking in optical interferometry

Conference Paper in Proceedings of SPIE - The International Society for Optical Engineering · July 2014

DOI: 10.1117/12.2056590

CITATIONS

0

READS

15

5 authors, including:



Ferréol Soulez

Claude Bernard University Lyon 1

64 PUBLICATIONS 439 CITATIONS

SEE PROFILE



Éric Thiébaud

Claude Bernard University Lyon 1

251 PUBLICATIONS 2,047 CITATIONS

SEE PROFILE



Michel Tallon

Claude Bernard University Lyon 1

131 PUBLICATIONS 1,648 CITATIONS

SEE PROFILE



Paulo Garcia

University of Porto

106 PUBLICATIONS 693 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



FRIEND project [View project](#)



European Solar Telescope MCAO [View project](#)

All content following this page was uploaded by [Paulo Garcia](#) on 07 April 2016.

The user has requested enhancement of the downloaded file.

Optimal a posteriori fringe tracking in optical interferometry

Ferréol Soulez^a, Éric Thiébaud^a Michel Tallon^a, Isabelle Tallon-Bosc^a, and Paulo Garcia^b

^aCentre de Recherche Astrophysique de Lyon - Observatoire de Lyon, Lyon, France, and with the Laboratoire Hubert Curien, Université Jean Monnet, Saint-Étienne, France;

^bUniversity of Porto, Departamento de Engenharia Física

ABSTRACT

The so-called “*phase delay tracking*” attempts to estimate the effects of the turbulence on the phase of the interferograms in order to numerically cophase the measured complex visibilities and to coherently integrate them. This is implemented by the “*coherent fringe analysis*” of MIDI instrument¹ but has only been used for high SNR data. In this paper, we investigate whether the sensitivity of this technique can be pushed to its theoretical limits and thus applied to fainter sources. In the general framework of the maximum likelihood and exploiting the chromatic behavior of the turbulence effects, we propose a global optimization strategy to compute various estimators of the differential pistons between two data frames. The most efficient estimators appear to be the ones based on the phasors, even though they do not yet reach the theoretical limits.

Keywords: optical interferometry, phase delay tracking cophasing

1. INTRODUCTION

Monomode optical interferometers get rid of atmospheric effects by integrating observables which are insensitive to the differential pistons between the interfering telescopes. For instance, the Fourier phase of the object is preserved in the so-called phase closures. This is however to the detriment of the amount of measured information. Exploiting the chromatic behavior of the differential pistons, it is in principle possible to perform “*a posteriori fringe tracking*” to coherently add the complex visibilities of the short exposure interferograms and thus not only enhance the signal to noise ratio but also minimize the loss of information. To benefit from these improvements, there are several important issues. First, the ultimate performances of the method depend on the quality of the criterion introduced to measure the co-phasing the errors. Second, no closed form solution exists to the problem of minimizing these errors and global optimization is mandatory to estimate the instantaneous differential piston parameters from the available data. These may explain why the “*coherent fringe analysis*” of MIDI instrument¹ has only been used for high SNR data. For instance, the study of Tatulli & le Bouquin² clearly shows that model-based estimators are superior to more simple ones computed by fast Fourier transforming the data. In this paper, we investigate whether the sensitivity of a posteriori fringe tracking can be pushed to its theoretical limits and thus applied to fainter sources.

2. FRINGE DATA

2.1 Available raw data

We consider an interferometric instruments like AMBER or PIONIER³ which exploit spatial filtering to convert phase fluctuations due to the turbulence into photometric and piston fluctuations. Using matrix notation, a general model of the raw interferometric data $\tilde{\mathbf{y}}$ provided by the detector of such interferometers takes the form:⁴

$$\tilde{\mathbf{y}} = \mathbf{H} \cdot \mathbf{x} + \tilde{\mathbf{n}}, \quad (1)$$

where \mathbf{H} is a generalization of the “*visibility to pixel matrix*” (V2PM^{2,3}), $\tilde{\mathbf{n}}$ accounts for the noise and \mathbf{x} combines photometric parameters, the complex visibility of the object and atmospheric piston errors. For an ideal two telescope monomode recombiner, the measured interferogram writes:

$$\tilde{y}_{k,\ell,m} = \frac{N_{\ell,m}}{\gamma K} [1 + \rho_{\ell,m} \cos(\varphi_{k,\ell} + \psi_{\ell,m})] + \tilde{n}_{k,\ell,m} = (\mathbf{H}_{\ell} \cdot \mathbf{x}_{\ell,m})_k + \tilde{n}_{k,\ell,m}, \quad (2)$$

Further author information: E-mail: ferreol.soulez@univ-lyon1.fr

with $N_{\ell,m}$ the expected total number of photons of the fringe pattern in ℓ -th spectral channel of m -th frame, K the number of samples along such a pattern, γ the detector gain (in e^- per digital level), $\rho_{\ell,m}$ and $\psi_{\ell,m}$ the contrast and phase of the fringes and $\varphi_{k,\ell}$ the instrumental phase offset to sample the fringes. In this paper, to maintain the consistency of the notations, k , ℓ and m are respectively the pixel, spectral channel and frame indexes. Here, to simplify the equations, we assumed that K and γ are the same for all spectral channels. Using a matrix notation, the components of the deterministic term of Eq. (2) are:

$$\mathbf{x}_{\ell,m} = N_{\ell,m} \begin{pmatrix} 1 \\ \rho_{\ell,m} \cos \psi_{\ell,m} \\ \rho_{\ell,m} \sin \psi_{\ell,m} \end{pmatrix} \quad \text{and} \quad \mathbf{H}_\ell = \frac{1}{K \gamma} \begin{pmatrix} 1 & \cos(\varphi_{1,\ell}) & -\sin(\varphi_{1,\ell}) \\ \vdots & \vdots & \vdots \\ 1 & \cos(\varphi_{k,\ell}) & -\sin(\varphi_{k,\ell}) \\ \vdots & \vdots & \vdots \\ 1 & \cos(\varphi_{K,\ell}) & -\sin(\varphi_{K,\ell}) \end{pmatrix}. \quad (3)$$

Assuming L spectral channels and M frames and going back to Eq. (1), \mathbf{H} is very sparse with $3 \times K$ blocks and $\mathbf{x} = \{\mathbf{x}_{\ell,m} \in \mathbb{R}^3 \mid \ell = 1, \dots, L; m = 1 \dots, M\}$. The second and third components of $\mathbf{x}_{\ell,m}$ form what is commonly called the “*raw complex visibilities*” which we denote as:

$$\mathbf{z}_{\ell,m} = \left[(\mathbf{x}_{\ell,m})_2, (\mathbf{x}_{\ell,m})_3 \right]^\top \in \mathbb{R}^2 \quad \iff \quad z_{\ell,m} = (\mathbf{x}_{\ell,m})_2 + i (\mathbf{x}_{\ell,m})_3 = N_{\ell,m} \rho_{\ell,m} e^{i \psi_{\ell,m}} \in \mathbb{C}. \quad (4)$$

As formally expressed above, the 2-D variables $\mathbf{z}_{\ell,m}$ can also be seen as complexes denoted as $z_{\ell,m}$.

The contrast $\rho_{\ell,m}$ and the phase $\psi_{\ell,m}$ are related to the complex visibility of the object and are what one wants to recover from the data. These quantities are however affected by time varying effects mostly due to the atmospheric turbulence. The phase of the fringes is randomly shifted due to optical delay not compensated by the delay lines:

$$\psi_{\ell,m} = \phi_\ell + \kappa_\ell p_m \quad (5)$$

where ϕ_ℓ is the phase of the object complex visibility at λ_ℓ , the effective wavelength in ℓ -th channel, p_m is the optical path difference between the two interfering telescopes during m -th frame and κ_ℓ is a chromatic factor. If the delay lines are in the vacuum, $\kappa_\ell = 2\pi/\lambda_\ell$. The speckle noise and the motion of the fringes during a frame yield variations of the flux (and thus $N_{\ell,m}$) and of the fringe contrast $\rho_{\ell,m}$. This paper focuses on getting rid of the phase jitter.

2.2 Statistics of the raw data

With current detectors (CCD or CMOS), we can safely assume that measurements provided by distinct detector elements or from different detector frames are statistically independent. The covariance matrix of the noise $\tilde{\mathbf{n}}$ is thus diagonal and, its inverse, the “*statistical weighting matrix*” \mathbf{W} , is also diagonal: $\mathbf{W} = \text{Cov}(\tilde{\mathbf{n}})^{-1} = \text{diag}(\mathbf{w})$ with \mathbf{w} the vector of diagonal elements of \mathbf{W} . In principle, the elements of \mathbf{w} are simply equal to the reciprocal of the variance of the corresponding measurement. At the cost of having an improper statistics, it is possible to rigorously account for bad or missing data by taking:

$$w_{k,\ell,m} = \begin{cases} 0 & \text{for bad pixels,} \\ 1/\text{Var}\{\tilde{n}_{k,\ell,m}\} & \text{else.} \end{cases} \quad (6)$$

This amounts to assuming that the variance of missing or bad pixels is infinite which is consistent with the fact that they provide no valuable information. Accounting for Poisson photon noise plus Gaussian detector noise and digitization yields:⁵

$$\text{Var}\{\tilde{y}_{k,\ell,m} \mid \mathbf{x}\} = \mathbb{E}\{\tilde{y}_{k,\ell,m} \mid \mathbf{x}\}/\gamma + (\sigma/\gamma)^2 = (1/\gamma) (\mathbf{H} \cdot \mathbf{x})_{k,\ell,m} + (\sigma/\gamma)^2, \quad (7)$$

where $\mathbb{E}\{\tilde{y}_{k,\ell,m} \mid \mathbf{x}\}$ is the expected value of the data given the model parameters \mathbf{x} and σ is the standard deviation of the detector noise (in e^- rms per pixel). For instance, to account for digitization errors:

$$\sigma = \sqrt{\sigma_{\text{ron}}^2 + \gamma^2/12}, \quad (8)$$

with σ_{ron} the standard deviation of the read-out noise (in e^- per pixel per frame).

2.3 Reduced fringe data

The real statistics of the raw data is quite complex as it combines the Poisson distribution of the signal, the distribution of the detector noise but also non-trivial effects such as digitization. Here, we assume that this complex statistics is fairly well approximated by a Gaussian distribution with a non-uniform variance given by Eq. (7). Then, maximizing the likelihood of the data amounts to minimizing the co-log-likelihood:

$$\mathcal{L}_{\text{raw}}(\mathbf{x}) = (1/2) \{ (\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x})^\top \cdot \mathbf{W} \cdot (\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x}) - \log|\mathbf{W}|_+ \} \quad (9)$$

where $|\mathbf{W}|_+$ is the product of the strictly positive eigenvalues of the statistical weighting matrix \mathbf{W} (described in the previous section). When \mathbf{W} is non-singular, it is positive definite, and $|\mathbf{W}|_+$ is just the determinant of \mathbf{W} . The term $\log|\mathbf{W}|_+$ comes from the normalization factor of a, possibly improper, Gaussian distribution.⁶

In an attempt to reduce the amount of data which has to be processed and to unify the outputs of the different types of instruments, we introduce $\tilde{\mathbf{x}}$, a particular value of variables \mathbf{x} which minimizes the first term in the right hand side of Eq. (9):

$$\tilde{\mathbf{x}} \in \arg \min_{\mathbf{x}} (\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x})^\top \cdot \mathbf{W} \cdot (\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x}) \iff (\mathbf{H}^\top \cdot \mathbf{W} \cdot \mathbf{H}) \cdot \tilde{\mathbf{x}} = \mathbf{H}^\top \cdot \mathbf{W} \cdot \tilde{\mathbf{y}}. \quad (10)$$

A solution of the so-called “normal equations” (after the \Leftrightarrow sign above) always exists but may not be unique. Now, using the equality in Eq. (10), it is possible to rewrite the first term in the right hand side of Eq. (9):

$$(\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x})^\top \cdot \mathbf{W} \cdot (\tilde{\mathbf{y}} - \mathbf{H} \cdot \mathbf{x}) = (\mathbf{x} - \tilde{\mathbf{x}})^\top \cdot \mathbf{Q} \cdot (\mathbf{x} - \tilde{\mathbf{x}}) + c \quad \text{with} \quad \mathbf{Q} \stackrel{\text{def}}{=} \mathbf{H}^\top \cdot \mathbf{W} \cdot \mathbf{H}, \quad (11)$$

and with $c = \tilde{\mathbf{y}}^\top \cdot \mathbf{W} \cdot \tilde{\mathbf{y}} - \tilde{\mathbf{x}}^\top \cdot \mathbf{Q} \cdot \tilde{\mathbf{x}}$ which is a constant with respect to the sought parameters \mathbf{x} as $\tilde{\mathbf{x}}$ only depends on the original data $\tilde{\mathbf{y}}$. Besides, we note that $\log|\mathbf{W}|_+ = \log|\mathbf{Q}|_+ - c'$ where $c' = \log|\mathbf{H}^\top \cdot \mathbf{H}|_+$ is independent on the sought parameters \mathbf{x} . Finally, after discarding the constant terms c and c' , finding the maximum likelihood parameters amounts to minimizing the following data fidelity criterion:

$$\mathcal{L}_{\text{reduc}}(\mathbf{x}) = (1/2) \{ (\mathbf{x} - \tilde{\mathbf{x}})^\top \cdot \mathbf{Q} \cdot (\mathbf{x} - \tilde{\mathbf{x}}) - \log|\mathbf{Q}|_+ \}. \quad (12)$$

Note that \mathbf{Q} only depends on the variables \mathbf{x} if the weighting matrix \mathbf{W} does depend on \mathbf{x} (see previous section). If \mathbf{W} is known *a priori* (it does not depend on the sought parameters), the term $\log|\mathbf{Q}|_+$ is a constant which can be dropped from Eq. (12) and the reduced data $\tilde{\mathbf{x}}$ are also the most likely values of the linear parameters \mathbf{x} .

As previously pointed, a solution of the normal equations always exists but is not necessarily unique unless the matrix \mathbf{Q} is invertible. However, the matrix \mathbf{Q} does not need to be invertible to define a particular solution to the normal equations. For instance, introducing \mathbf{Q}^\dagger , the pseudo inverse of \mathbf{Q} (which does always exist), yields the following particular solution of the normal equations (10):

$$\tilde{\mathbf{x}} = \mathbf{R} \cdot \tilde{\mathbf{y}} \quad \text{with} \quad \mathbf{R} = \mathbf{Q}^\dagger \cdot \mathbf{H}^\top \cdot \mathbf{W} = (\mathbf{H}^\top \cdot \mathbf{W} \cdot \mathbf{H})^\dagger \cdot \mathbf{H}^\top \cdot \mathbf{W}. \quad (13)$$

In order to compute $\tilde{\mathbf{x}}$ given the raw data $\tilde{\mathbf{y}}$, the pseudo-inverse \mathbf{Q}^\dagger does not need to be explicitly formed (*e.g.*, using the singular value decomposition (SVD) of \mathbf{Q}). There are faster means. For instance, for uncorrelated spectral channels, \mathbf{Q} is block diagonal and independent smaller problems have to be solved for the different channels. Starting with a null initial solution, the linear conjugate gradient method⁷ will produce the minimum norm solution to the problem (10) in a few iterations.⁸

It is important to realize that fitting the “reduced” interferometric data $\tilde{\mathbf{x}}$ by minimizing $\mathcal{L}_{\text{reduc}}$ in Eq. (12) is exactly equivalent to fitting the raw interferometric data $\tilde{\mathbf{y}}$ by minimizing \mathcal{L}_{raw} in Eq. (9): there are no loss of valuable information. In other words, the reduced variables $\tilde{\mathbf{x}}$ have “sufficient statistics”⁹ for our needs. There is a close relationship between the matrix \mathbf{R} and the “pixel to visibility matrix”^{2,3} (P2VM). It can be shown that $\tilde{\mathbf{x}} = \mathbf{R} \cdot \tilde{\mathbf{y}}$ defined in Eq. (13) is the “best linear unbiased estimator”¹⁰ (BLUE) of the raw complex visibilities and therefore has the highest possible SNR.

2.4 Statistics of the reduced fringe data

In order to derive the statistics of the reduced data, we consider a particular spectral channel and frame. To simplify the equations, we therefore discard the spectral index ℓ and the frame index m to only keep the pixel index k . For the simple interferogram given in Eq. (2), the direct model of k -th data pixel writes:

$$\tilde{y}_k = \frac{N}{K\gamma} [1 + \rho \cos(\varphi_k + \psi)] + \tilde{n}_k. \quad (14)$$

where N is the expected value of the total number of detected photons, K is the number of pixels which sample the fringe, γ is the detector gain in e^- per digital level, φ_k is the phase shift at pixel k , ρ is the actual contrast of the fringes (possibly accounting for some instrumental loss) and ψ is their phase. Using matrix notation:

$$\tilde{\mathbf{y}} = \mathbf{H} \cdot \mathbf{x} + \tilde{\mathbf{n}} \quad \text{with} \quad \mathbf{x} = N \begin{pmatrix} 1 \\ \rho \cos \psi \\ \rho \sin \psi \end{pmatrix} \quad \text{and} \quad \mathbf{H} = \frac{1}{K\gamma} \begin{pmatrix} \vdots & \vdots & \vdots \\ 1 & \cos(\varphi_k) & -\sin(\varphi_k) \\ \vdots & \vdots & \vdots \end{pmatrix}.$$

Thanks to the properties of the pseudo inverse (denoted by the \dagger exponent), the covariance matrix of the estimated linear fringe parameters $\tilde{\mathbf{x}} = \mathbf{R} \cdot \tilde{\mathbf{y}}$ defined in Eq. (13) is simply:

$$\text{Cov}(\tilde{\mathbf{x}} | \mathbf{x}) = \mathbf{R} \cdot \text{Cov}(\tilde{\mathbf{y}} | \mathbf{x}) \cdot \mathbf{R}^\top = (\mathbf{H}^\top \cdot \mathbf{W} \cdot \mathbf{H})^\dagger,$$

where $\mathbf{W} = \text{Cov}(\tilde{\mathbf{y}} | \mathbf{x})^{-1} = \text{diag}(\mathbf{w})$ is a diagonal weighting matrix (the pixels are statistically independent, cf. Section 2.2). From Eqs. (6)–(7), the weights are:

$$w_k = \frac{\gamma^2}{\gamma \mathbb{E}\{\tilde{y}_k | \mathbf{x}\} + \sigma^2} = \frac{K \gamma^2}{N + K \sigma^2} [1 + \varepsilon \cos(\varphi_k + \psi)]^{-1} \quad \text{with} \quad \varepsilon = \frac{N \rho}{N + K \sigma^2} \in [0, 1). \quad (15)$$

To derive a useful expression for $\text{Cov}(\tilde{\mathbf{x}} | \mathbf{x})$, it is more convenient* to consider the following invertible linear change of variables:

$$\tilde{\mathbf{u}} = \mathbf{Z}_\psi^\top \cdot \tilde{\mathbf{x}} \quad \iff \quad \tilde{\mathbf{x}} = \mathbf{Z}_\psi \cdot \tilde{\mathbf{u}} \quad \text{with} \quad \mathbf{Z}_\psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{pmatrix}.$$

Note that \mathbf{Z}_ψ is a unitary matrix such that $\mathbf{Z}_\psi^{-1} = \mathbf{Z}_{-\psi} = \mathbf{Z}_\psi^\top$. The covariance of the new variables writes $\text{Cov}(\tilde{\mathbf{u}} | \mathbf{x}) = \mathbf{Z}_\psi^\top \cdot \text{Cov}(\tilde{\mathbf{x}} | \mathbf{x}) \cdot \mathbf{Z}_\psi$ and, thanks to the properties of \mathbf{Z}_ψ , taking the pseudo inverse yields:

$$\text{Cov}(\tilde{\mathbf{u}} | \mathbf{x})^\dagger = \mathbf{Z}_\psi^\top \cdot \text{Cov}(\tilde{\mathbf{x}} | \mathbf{x})^\dagger \cdot \mathbf{Z}_\psi = (\mathbf{H} \cdot \mathbf{Z}_\psi)^\top \cdot \mathbf{W} \cdot (\mathbf{H} \cdot \mathbf{Z}_\psi) = \frac{1}{N + K \sigma^2} \begin{pmatrix} q_1 & q_2 & q_3 \\ q_2 & q_4 & q_5 \\ q_3 & q_5 & q_6 \end{pmatrix}$$

where:

$$\begin{aligned} q_1 &= \frac{1}{K} \sum_{k=1}^K \frac{1}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx \frac{1}{\sqrt{1 - \varepsilon^2}}, & q_2 &= \frac{1}{K} \sum_{k=1}^K \frac{\cos(\varphi_k + \psi)}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx -\frac{1 - \sqrt{1 - \varepsilon^2}}{\varepsilon \sqrt{1 - \varepsilon^2}}, \\ q_3 &= \frac{1}{K} \sum_{k=1}^K \frac{-\sin(\varphi_k + \psi)}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx 0, & q_4 &= \frac{1}{K} \sum_{k=1}^K \frac{\cos^2(\varphi_k + \psi)}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx \frac{1 - \sqrt{1 - \varepsilon^2}}{\varepsilon^2 \sqrt{1 - \varepsilon^2}}, \\ q_5 &= \frac{1}{K} \sum_{k=1}^K \frac{-\cos(\varphi_k + \psi) \sin(\varphi_k + \psi)}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx 0, & q_6 &= \frac{1}{K} \sum_{k=1}^K \frac{\sin^2(\varphi_k + \psi)}{1 + \varepsilon \cos(\varphi_k + \psi)} \approx \frac{1}{1 + \sqrt{1 - \varepsilon^2}}, \end{aligned}$$

where all the terms having the form $(1/K) \sum_{k=1}^K f(\varphi_k + \psi)$ have been approximated by the mean value of $f(\varphi_k + \psi)$ with respect to the sample offset φ_k . This is similar to the Riemann approximation of integrals by

*this trick makes all phase arguments of the trigonometric functions into the form: $\varphi_k + \psi$

discrete sums and assumes that the fringes are correctly and evenly sampled. Under these assumptions, replacing the q_i 's by their approximation in the expression of $\text{Cov}(\tilde{\mathbf{u}} | \mathbf{x})^\dagger$ and taking the inverse yields:

$$\text{Cov}(\tilde{\mathbf{u}} | \mathbf{x}) \approx (N + K \sigma^2) \begin{pmatrix} 1 & \varepsilon & 0 \\ \varepsilon & 1 + \sqrt{1 - \varepsilon^2} & 0 \\ 0 & 0 & 1 + \sqrt{1 - \varepsilon^2} \end{pmatrix}$$

and:

$$\text{Cov}(\tilde{\mathbf{x}} | \mathbf{x}) = \mathbf{Z}_\psi \cdot \text{Cov}(\tilde{\mathbf{u}} | \mathbf{x}) \cdot \mathbf{Z}_\psi^\top \approx (N + K \sigma^2) \times \begin{pmatrix} 1 & \varepsilon \cos \psi & \varepsilon \sin \psi \\ \varepsilon \cos \psi & 1 + \sqrt{1 - \varepsilon^2} & 0 \\ \varepsilon \sin \psi & 0 & 1 + \sqrt{1 - \varepsilon^2} \end{pmatrix}. \quad (16)$$

The real and imaginary parts of the measured raw complex visibilities (the second and third components of $\tilde{\mathbf{x}}$) are therefore approximately independent with the same variance and thus follow Goodman model.¹¹ To our knowledge, this is the first demonstration of this property for a realistic non-stationary and signal dependent noise model.

2.5 Least squares estimator

Another possibility is to use a least squares estimator for \mathbf{x} :

$$\tilde{\mathbf{x}}_{\text{LSQ}} \stackrel{\text{def}}{=} \arg \min_{\mathbf{x}} \|\mathbf{H} \cdot \mathbf{x} - \tilde{\mathbf{y}}\|_2 = \mathbf{R}_{\text{LSQ}} \cdot \tilde{\mathbf{y}} \quad \text{with} \quad \mathbf{R}_{\text{LSQ}} = (\mathbf{H}^\top \cdot \mathbf{H})^\dagger \cdot \mathbf{H}^\top. \quad (17)$$

Compared to the maximum likelihood estimator $\tilde{\mathbf{x}}$, this amounts to assuming that the noise is i.i.d. in the data $\tilde{\mathbf{y}}$. Using the same approximations as previously (*cf.* Section 2.4), we can compute the covariance matrix of the least squares estimator:

$$\text{Cov}\{\tilde{\mathbf{x}}_{\text{LSQ}} | \mathbf{x}\} = \mathbf{R}_{\text{LSQ}} \cdot \text{Cov}\{\tilde{\mathbf{y}} | \mathbf{x}\} \cdot \mathbf{R}_{\text{LSQ}}^\top \approx (N + K \sigma^2) \times \begin{pmatrix} 1 & \varepsilon \cos \psi & \varepsilon \sin \psi \\ \varepsilon \cos \psi & 2 & 0 \\ \varepsilon \sin \psi & 0 & 2 \end{pmatrix}. \quad (18)$$

Compared to the covariance matrix of the maximum likelihood estimator $\tilde{\mathbf{x}}$ given in Eq. (16), we can see that:

$$\text{Cov}\{\tilde{\mathbf{x}}_{\text{LSQ}} | \mathbf{x}\} \succeq \text{Cov}\{\tilde{\mathbf{x}} | \mathbf{x}\}.$$

In words: the covariance of the least-squares estimator is larger than that of the maximum likelihood one.

2.6 Simulation of noisy reduced fringe data

Given the detector and fringes parameters, simulating noisy reduced fringe data $\tilde{\mathbf{x}}$ can be done as follows:

$$\tilde{x}_1 = N + \varsigma_1 \xi_1 + \varsigma_2 (\xi_2 \cos \psi + \xi_3 \sin \psi) \quad (19a)$$

$$\tilde{x}_2 = N \rho \cos \psi + \varsigma_3 \xi_2 \quad (19b)$$

$$\tilde{x}_3 = N \rho \sin \psi + \varsigma_3 \xi_3 \quad (19c)$$

where $\xi_j \sim \mathcal{N}(0, 1)$, for $j \in \{1, 2, 3\}$, are 3 independant pseudo-random numbers drawn following a normal distribution and with:

$$\varsigma_1 = \left((N + K \sigma^2) \sqrt{1 - \varepsilon^2} \right)^{1/2}, \quad (20a)$$

$$\varsigma_2 = \left((N + K \sigma^2) \frac{\varepsilon^2}{1 + \sqrt{1 - \varepsilon^2}} \right)^{1/2}, \quad (20b)$$

$$\varsigma_3 = \left((N + K \sigma^2) (1 + \sqrt{1 - \varepsilon^2}) \right)^{1/2}. \quad (20c)$$

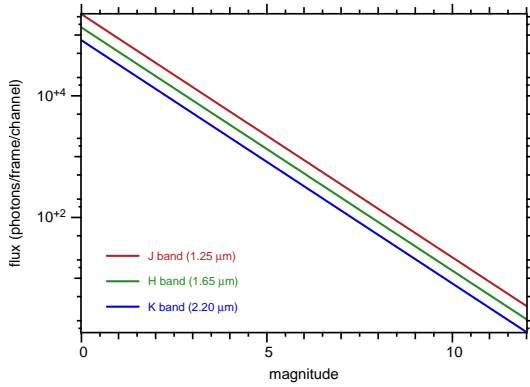


Figure 1. Detected flux per frame per channel as a function of the magnitude. Parameters: spectral resolution $R = 1500$, total throughput 1%, VLTI with UT's (subtract 3.3 to the magnitudes for the AT's) and 25 ms exposure time.

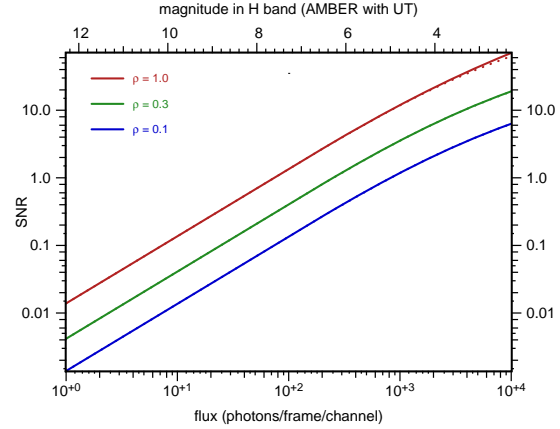


Figure 2. Fringe SNR for different flux levels and fringe contrasts. The solid curves are for the BLUES estimator, the dotted curves are for the least square estimator.

For all simulations presented in this paper, we considered different fringe contrasts ρ and number of photons N per frame per channel. The other parameters were set to correspond to AMBER instrument³ in H band ($\lambda = 1.65 \mu\text{m}$) at medium spectral resolution ($\lambda/\Delta\lambda = 1500$): $K = 32$ pixels per spectral channel, $\sigma = 9 e^-/\text{pixel}/\text{frame}$ and $L = 364$ spectral channels. Assuming a total throughput (optics + detector) of 1%, and a frame acquisition time of 25 ms. Figure 1 shows the expected number of photons N per frame per channel as a function of the magnitude.

2.7 Signal to noise ratio of the fringe signal

Considering the deterministic and stochastic terms in Eq. (19b) and Eq. (19c), the amplitude of the fringe signal in the reduced data is given by $N\rho$ while the standard deviation of the additive noise is equal to ς_3 . Thus, the signal noise ratio (SNR) of the fringe information is given by:

$$\text{SNR}_{\text{BLUE}} = \frac{N\rho}{\varsigma_3} = \frac{\varepsilon\sqrt{N+K\sigma^2}}{\sqrt{1+\sqrt{1-\varepsilon^2}}} = \frac{N\rho}{\sqrt{N+K\sigma^2 + \sqrt{(N+K\sigma^2)^2 - (N\rho)^2}}}. \quad (21)$$

Note that this expression only depends on three parameters: N the number of photons, $K\sigma^2$ the total variance of the detector noise and ρ the fringe contrast (a.k.a. the fringe visibility). Figure 2 shows the evolution of the SNR of a single frame from a single channel as a function of the number of photons N per frame per channel and for different fringe contrasts ρ . For comparison, the SNR of the least-squares estimator would be:

$$\text{SNR}_{\text{LSQ}} = \frac{N\rho}{\sqrt{2(N+K\sigma^2)}} \leq \text{SNR}_{\text{BLUE}}. \quad (22)$$

When $\varepsilon \rightarrow 1$ (that is photon noise regime and highly contrasted fringes), the BLUE estimator yields a $\sqrt{2}$ improvement with respect to the least squares one. When $\varepsilon \rightarrow 0$ (that is detector noise regime and fringes with low visibility), the two estimators have the same SNR. Except for magnitudes brighter than 4, AMBER is in this latter regime.

2.8 Maximum likelihood parameters from the reduced data

Putting all together and taking into account the statistical independence of the reduced data from different spectral channels and frames, maximizing the likelihood of the reduced data amount to minimize:

$$\mathcal{L}_{\text{reduc}} = \frac{1}{2} \sum_{\ell,m} \left\{ (\tilde{\mathbf{x}}_{\ell,m} - \mathbf{x}_{\ell,m})^\top \cdot \mathbf{C}_{\ell,m}^{-1} \cdot (\tilde{\mathbf{x}}_{\ell,m} - \mathbf{x}_{\ell,m}) + \log|\mathbf{C}_{\ell,m}|_+ \right\} \quad (23)$$

where $\tilde{\mathbf{x}}_{\ell,m}$ are the reduced data for ℓ -th spectral channel and m -th frame, $\mathbf{C}_{\ell,m} = \text{Cov}(\tilde{\mathbf{x}}_{\ell,m} | \mathbf{x}_{\ell,m})$ is the 3×3 covariance matrix given in Eq. (16) and:

$$\mathbf{x}_{\ell,m} = \begin{pmatrix} N_{\ell,m} \\ \mu_{\ell,m} \cos(\phi_{\ell} + \kappa_{\ell} p_m) \\ \mu_{\ell,m} \sin(\phi_{\ell} + \kappa_{\ell} p_m) \end{pmatrix} \quad \text{with} \quad 0 \leq \mu_{\ell,m} = N_{\ell,m} \rho_{\ell,m} \leq N_{\ell,m}, \quad (24)$$

their model. Here, the model parameters are: $N_{\ell,m}$ the expected number of photons, $\rho_{\ell,m} e^{i\phi_{\ell}}$ the complex visibility of the fringes not accounting for p_m the atmospheric piston. The phase ϕ_{ℓ} does not depend on the time (and hence on the frame index m) but the amplitude $\rho_{\ell,m}$ is subject of temporal variations[†] and hence depend on the frame index m .

3. COPHASING THE FRAMES

Fitting all the non-linear parameters on the reduced data is a difficult problem which involves global minimization of the multi-variate criterion $\mathcal{L}_{\text{reduc}}$ in Eq. (23). The most difficult part is certainly to estimate the differential piston between the frames, this process is called “*cophasing*” or “*fringe tracking*”. If this problem is properly solved, one can coherently add the fringe information from the different frames and thus obtain interferometric observables much more interesting (higher SNR and fewer missing information).

3.1 Maximum Likelihood Cophasing

Here only the “*fringe part*” of the reduced data is needed, so we discard the “*photometric*” part and introduce the measured “*raw visibility*”:

$$\tilde{\mathbf{z}}_{\ell,m} = [(\tilde{\mathbf{x}}_{\ell,m})_2, (\tilde{\mathbf{x}}_{\ell,m})_3]^{\top} \in \mathbb{R}^2 \quad \iff \quad \tilde{z}_{\ell,m} = (\tilde{\mathbf{x}}_{\ell,m})_2 + i (\tilde{\mathbf{x}}_{\ell,m})_3 \in \mathbb{C}. \quad (25)$$

As shown in Section 2.6, the covariance matrix of $\tilde{\mathbf{z}}_{\ell,m}$ is proportional to the identity. We denote the variance of the real and imaginary parts of $\tilde{z}_{\ell,m}$ as $\sigma_{\ell,m}$ and their statistical weight as $w_{\ell,m} = \sigma_{\ell,m}^{-2}$. In the case of the BLUE estimator:

$$\sigma_{\ell,m}^2 = N_{\ell,m} + K \sigma^2 - \sqrt{(N_{\ell,m} + K \sigma^2)^2 - \mu_{\ell,m}^2}. \quad (26)$$

To slightly simplify the cophasing problem, we will assume that the statistical weights are known and ignore their dependency with the sought parameters. This assumption can be relaxed by iteratively applying a reweighting technique.

In the limit of small errors, the variance of the phase of the complex data \tilde{z} is given by:¹²

$$\text{Var}\{\arg(\tilde{z})\} = \frac{\text{Re}(z)^2 \text{Var}\{\text{Im}(\tilde{z})\} - 2 \text{Re}(z) \text{Im}(z) \text{Cov}\{\text{Re}(\tilde{z}), \text{Im}(\tilde{z})\} + \text{Im}(z)^2 \text{Var}\{\text{Re}(\tilde{z})\}}{|z|^4} \quad (27)$$

with $z \stackrel{\text{def}}{=} \mathbb{E}\{\tilde{z}\}$. Applying this expression to the raw visibility $\tilde{z}_{\ell,m}$ yields:

$$\text{Var}\{\arg(\tilde{z}_{\ell,m})\} = \frac{\sigma_{\ell,m}^2}{|\mathbb{E}\{\tilde{z}_{\ell,m}\}|^2} = \text{SNR}^{-2}, \quad (28)$$

where SNR is the signal to noise ratio introduced in Eqs. (21) and (22).

In order to account for the degeneracies of the problem, we express the expected values of the fringe data as:

$$z_{\ell,m}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \mathbb{E}(\tilde{\mathbf{z}}_{\ell,m} | \boldsymbol{\theta}) = \mathbf{G}_{\ell,m} \cdot \mathbf{z}_{\ell} \quad \text{with: } \mathbf{G}_{\ell,m} = \alpha_{\ell,m} \begin{pmatrix} \cos(\kappa_{\ell} s_m) & -\sin(\kappa_{\ell} s_m) \\ \sin(\kappa_{\ell} s_m) & \cos(\kappa_{\ell} s_m) \end{pmatrix} \quad (29a)$$

$$z_{\ell,m}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \mathbb{E}(\tilde{\mathbf{z}}_{\ell,m} | \boldsymbol{\theta}) = g_{\ell,m} z_{\ell} \quad \text{with: } g_{\ell,m} = \alpha_{\ell,m} e^{i\kappa_{\ell} s_m}, \quad (29b)$$

[†]The fringe contrast is scaled by a factor $2 \sqrt{N_{j_1,\ell,m} N_{j_2,\ell,m}} / N_{\ell,m}$ where $N_{j,\ell,m}$ is the expected number of photons from j -th telescope in ℓ -th channel and m -th frame and $N_{\ell,m} = N_{j_1,\ell,m} + N_{j_2,\ell,m}$.

where $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \mathbf{s}, \mathbf{Z}\}$ represents the sought parameters with $\boldsymbol{\alpha} = \{\alpha_{\ell,m} \geq 0 \mid \ell = 1, \dots, L; m \in \mathbb{M}\}$ non-negative factors, $\mathbf{s} = \{s_m = p_m - p_{\text{ref}} \mid m \in \mathbb{M}\}$ the differential pistons for m -th frame and relative to an arbitrary piston p_{ref} and $\mathbf{Z} = \{z_\ell \in \mathbb{C} \text{ or } z_\ell \in \mathbb{R}^2 \mid \ell = 1, \dots, L\}$ where $z_\ell \sim \tilde{z}_\ell$ is the uncalibrated complex visibility equals to that of the object times some chromatic attenuation and up to a phase shift due to the piston p_{ref} . For instance, $z_\ell = z_{\ell,m_0}$ the expected value of \tilde{z}_{ℓ,m_0} given the sought parameters at some reference frame m_0 , then $p_{\text{ref}} = p_{m_0}$ and $\alpha_{\ell,m} = |z_{\ell,m}|/|z_{\ell,m_0}|$. In order to derive more general equations, we consider the problem of cophasing a subset \mathbb{M} of the frames and we assume that some measurement \tilde{z}_ℓ of z_ℓ of the uncalibrated visibility is available with w_ℓ its corresponding statistical weight (possibly $w_\ell \equiv 0$ if no such data exists). Under our assumptions and noting that:

$$[\tilde{z}_{\ell,m} - z_{\ell,m}(\boldsymbol{\theta})]^\top \cdot \text{Cov}(\tilde{z}_{\ell,m} \mid \boldsymbol{\theta})^{-1} \cdot [\tilde{z}_{\ell,m} - z_{\ell,m}(\boldsymbol{\theta})] = w_{k,\ell} \|\tilde{z}_{\ell,m} - z_{\ell,m}(\boldsymbol{\theta})\|_2^2 = w_{k,\ell} |\tilde{z}_{\ell,m} - z_{\ell,m}(\boldsymbol{\theta})|^2,$$

the cophasing is achieved in a maximum likelihood sense by minimizing:

$$\mathcal{L}_{\text{sync}}(\boldsymbol{\alpha}, \mathbf{s}, \mathbf{Z}) = \frac{1}{2} \sum_{\ell} \left\{ w_\ell |\tilde{z}_\ell - z_\ell|^2 + \sum_{m \in \mathbb{M}} w_{\ell,m} |\tilde{z}_{\ell,m} - g_{\ell,m} z_\ell|^2 \right\}, \quad (30)$$

where any additive constants which do not depend on the parameters have been discarded and where the complex notation has been used for it is the most compact.

To simply reduce the number of unknown, there are two possibilities: we can first optimize with respect to the z_ℓ or with respect to the $\alpha_{\ell,m}$ as the criterion depends quadratically on these variables. The former choice, which we known describe, yields a simpler algorithm. Given the complex gains $g_{\ell,m}$, the maximum likelihood estimator of the uncalibrated visibility z_ℓ is:

$$z_\ell^+(\boldsymbol{\alpha}, \mathbf{s}) = \frac{w_\ell \tilde{z}_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} g_{\ell,m}^* \tilde{z}_{\ell,m}}{w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} |g_{\ell,m}|^2} = \frac{w_\ell \tilde{z}_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} \mathbf{G}_{\ell,m}^\top \tilde{z}_{\ell,m}}{w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} \alpha_{\ell,m}^2}. \quad (31)$$

Which can be seen as a weighted average of the measured reduced raw visibilities corrected for their phase and amplitude. This is the “*best linear unbiased estimator*” (BLUE) of the uncalibrated visibility but it is conditioned to the knowledge of the gains, that is of $\boldsymbol{\alpha}$ and \mathbf{s} . It can be shown that the statistical weights of z_ℓ^+ are:

$$w_\ell^+(\boldsymbol{\alpha}) \stackrel{\text{def}}{=} \text{Var}(z_\ell^+(\boldsymbol{\alpha}, \mathbf{s}) \mid \boldsymbol{\alpha}, \mathbf{s})^{-1} = w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} |g_{\ell,m}|^2 = w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} \alpha_{\ell,m}^2, \quad (32)$$

which only depends on $\boldsymbol{\alpha}$.

Substituting into $\mathcal{L}_{\text{sync}}$ the uncalibrated visibilities z_ℓ by z_ℓ^+ yields a new criterion which only depends on $\boldsymbol{\alpha}$ and \mathbf{s} :

$$\mathcal{L}_{\text{sync}}^+(\boldsymbol{\alpha}, \mathbf{s}) \stackrel{\text{def}}{=} \mathcal{L}_{\text{sync}}(\boldsymbol{\alpha}, \mathbf{s}, \mathbf{Z}^+(\boldsymbol{\alpha}, \mathbf{s})) = \frac{1}{2} \sum_{\ell} \left\{ w_\ell |\tilde{z}_\ell|^2 + \sum_{m \in \mathbb{M}} w_{\ell,m} |\tilde{z}_{\ell,m}|^2 - \frac{|w_\ell \tilde{z}_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} g_{\ell,m}^* \tilde{z}_{\ell,m}|^2}{w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} |g_{\ell,m}|^2} \right\}.$$

Thus the maximum likelihood parameters $\boldsymbol{\alpha}$ and \mathbf{s} are:

$$\{\boldsymbol{\alpha}_{\text{ML}}, \mathbf{s}_{\text{ML}}\} \stackrel{\text{def}}{=} \arg \min_{\boldsymbol{\alpha} \geq 0, \mathbf{s}} \mathcal{L}_{\text{sync}}^+(\boldsymbol{\alpha}, \mathbf{s}) = \arg \max_{\boldsymbol{\alpha} \geq 0, \mathbf{s}} \left\{ \mathcal{E}_{\text{sync}}(\boldsymbol{\alpha}, \mathbf{s}) \stackrel{\text{def}}{=} \sum_{\ell} \frac{|w_\ell \tilde{z}_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} g_{\ell,m}^* \tilde{z}_{\ell,m}|^2}{w_\ell + \sum_{m \in \mathbb{M}} w_{\ell,m} |g_{\ell,m}|^2} \right\}. \quad (33)$$

All the difficulty now consists in maximizing the cophasing criterion $\mathcal{E}_{\text{sync}}$. We propose to split this difficult problem in easier to solve sub-problems. Starting by the cophasing of two frames which requires global optimization, we deduce a tractable estimate of all the delays. Finally, we fit all parameters by means of local optimization. Note that the global problem is similar to a data fusion problem for which frames have to be properly registered to add coherently.¹³ Like for the registration problem, the best cophasing parameters are the ones which maximize $\mathcal{E}_{\text{sync}}$ which is equivalent to maximizing the (squared) SNR of z_ℓ^+ summed over all spectral channels ℓ .

3.2 Cophasing of two frames

We consider here the problem of cophasing the reduced data of two frames, say m_0 and m_1 using all the spectral channels. Being able to solve this problem efficiently can be the building block of cophasing a complete sequence of frames. We first consider an approach based on the maximum likelihood and then introduce alternative methods where less variables need to be fitted and may thus yield better results.

3.2.1 Maximum likelihood cophasing of two frames

Using m_0 as the reference (thus $\tilde{z}_\ell = \tilde{z}_{\ell, m_0}$) and $\mathbb{M} = \{m_1\}$, the criterion (33) to be maximized writes:

$$\mathcal{C}_{\text{sync}}(\boldsymbol{\alpha}, s) = \sum_{\ell} \mathcal{C}_{\ell}^{\text{sync}}(\alpha_{\ell}, s) \quad \text{with: } \mathcal{C}_{\ell}^{\text{sync}}(\alpha, s) = \frac{|w_{\ell, m_0} \tilde{z}_{\ell, m_0} + w_{\ell, m_1} \alpha e^{-i \kappa_{\ell} s} \tilde{z}_{\ell, m_1}|^2}{w_{\ell, m_0} + w_{\ell, m_1} \alpha^2}, \quad (34)$$

where, to simplify the notation, $s \equiv s_{m_1}$ and $\alpha_{\ell} \equiv \alpha_{\ell, m_1}$. Exploiting the partial separability with respect to the wavelength, we maximize the criterion $\mathcal{C}_{\ell}^{\text{sync}}(\alpha, s)$ with respect to α subject to $\alpha \geq 0$. Taking the derivative:

$$\frac{\partial \mathcal{C}_{\ell}^{\text{sync}}(\alpha, s)}{\partial \alpha} = \frac{2 w_{\ell, m_0} w_{\ell, m_1} (c(s) + b \alpha - a(s) \alpha^2)}{(w_{\ell, m_0} + w_{\ell, m_1} \alpha^2)^2} \quad \text{with: } \begin{cases} a(s) = w_{\ell, m_1} q(s), \\ b = w_{\ell, m_1} |\tilde{z}_{\ell, m_1}|^2 - w_{\ell, m_0} |\tilde{z}_{\ell, m_0}|^2, \\ c(s) = w_{\ell, m_0} q(s), \\ q(s) = \text{Re}(\tilde{z}_{\ell, m_0} e^{i \kappa_{\ell} s} \tilde{z}_{\ell, m_1}^*). \end{cases}$$

As the denominator is strictly positive whatever α_{ℓ} , the stationary points of $\mathcal{C}_{\ell}^{\text{sync}}(\alpha, s)$ with respect to α are the roots of the numerator which is a quadratic polynomial. The discriminant of this quadratic is always non-negative, thus there are always two (possibly indistinct) roots. Maximizing $\mathcal{C}_{\ell}^{\text{sync}}(\alpha, s)$ with respect to α and subject to $\alpha \geq 0$ yields:

$$\alpha_{\ell}^+(s) = \arg \max_{\alpha \geq 0} \mathcal{C}_{\ell}^{\text{sync}}(\alpha, s) = \begin{cases} \frac{b + \sqrt{b^2 + 4 a(s) c(s)}}{2 a(s)} = \frac{2 c(s)}{\sqrt{b^2 + 4 a(s) c(s)} - b} & \text{if } q(s) > 0; \\ \max(w_{\ell, m_0} |\tilde{z}_{\ell, m_0}|^2, w_{\ell, m_1} |\tilde{z}_{\ell, m_1}|^2) & \text{otherwise.} \end{cases}$$

Remarks: The two expressions in the case $q(s) > 0$ are equivalent (the first one is numerically more stable for $b \geq 0$ while the second one is to be preferred if $b \leq 0$); in the case $q(s) \leq 0$ which correspond to $\tilde{z}_{\ell, m_0} e^{i \kappa_{\ell} s}$ and \tilde{z}_{ℓ, m_1} making an angle wider than 90° , the solution is either $\mathcal{C}_{\ell}^{\text{sync}}(\alpha = 0, s)$ or $\lim_{\alpha \rightarrow \infty} \mathcal{C}_{\ell}^{\text{sync}}(\alpha, s)$.

Putting the expression of $\alpha_{\ell}^+(s)$ in $\mathcal{C}_{\text{sync}}(\boldsymbol{\alpha}, s)$ finally yields a criterion:

$$\mathcal{C}_{\text{sync}}^+(s) = \sum_{\ell} \mathcal{C}_{\ell}^{\text{sync}}(\alpha_{\ell}^+(s), s) \quad (35)$$

which only depend on the single variable s and has to be maximized by means of 1D global optimization. Figure 3 shows a typical plot of $u(s) = -\mathcal{C}_{\text{sync}}^+(s)$. Compared to the full global optimization problem, such operation is tractable and far more easier to perform safely with a rather simple algorithm BRADI given below.

Global minimization algorithm BRADI (*bracket then dig*). Given the search bounds $s_{\min} < s_{\max}$ and a sampling step $\delta s > 0$, find the global minimum of the continuous function $u : \mathbb{R} \mapsto \mathbb{R}$ on $[s_{\min}, s_{\max}]$.

0. Initialize: $s_{\text{best}} = s_{\text{b}} = s_{\text{c}} = s_{\min}$ and $u_{\text{best}} = u_{\text{b}} = u_{\text{c}} = u(s_{\min})$.

1. Next sub-interval: $s_{\text{a}} \leftarrow s_{\text{b}}$ and $u_{\text{a}} \leftarrow u_{\text{b}}$;
 $s_{\text{b}} \leftarrow s_{\text{c}}$ and $u_{\text{b}} \leftarrow u_{\text{c}}$;
 $s_{\text{c}} \leftarrow \min(s_{\text{c}} + \delta s, s_{\max})$ and $u_{\text{c}} \leftarrow u(s_{\text{c}})$.

If $u_{\text{c}} < u_{\text{best}}$ then update the best solution so far: $s_{\text{best}} \leftarrow s_{\text{c}}$ and $u_{\text{best}} \leftarrow u_{\text{c}}$.

2. If $u_{\text{b}} \leq \min(u_{\text{a}}, u_{\text{c}})$, a local minimum has been *bracketed*, then use Brent's method¹⁴ to *dig* and find it:

$$s_{\text{m}} \approx \arg \min_{s \in (s_{\text{a}}, s_{\text{c}})} u(s),$$

and if $u_{\text{m}} = u(s_{\text{m}}) < u_{\text{best}}$ then update the best solution so far: $s_{\text{best}} \leftarrow s_{\text{m}}$ and $u_{\text{best}} \leftarrow u_{\text{m}}$.

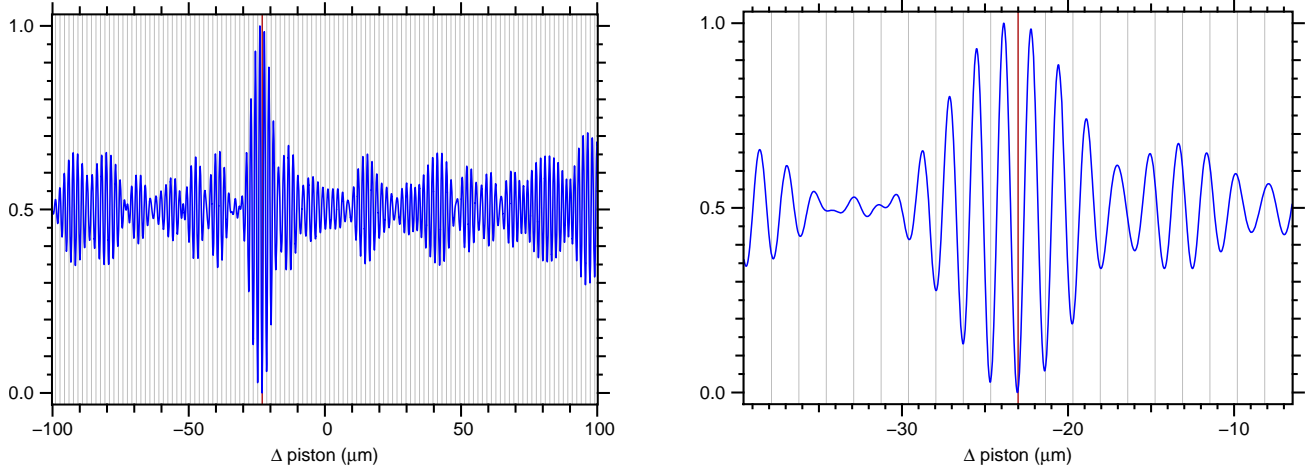


Figure 3. Criterion for cophasing two frames. The left panel displays the multi-modal criterion $u(s) = -\mathcal{C}_{\text{sync}}^+(s)$ for a differential piston $s \in [-0.1 \text{ mm}, +0.1 \text{ mm}]$. The right panel is a magnification around the global minimum. The red vertical line indicates the true differential piston, the vertical lines are spaced by the effective wavelength.

3. If $s_b < s_c$, then go to step 1; otherwise, exit with the solution s_{best} . ■

Remarks: (i) The BRADI algorithm is given in a rather standard form (that is to minimize a function). For our needs, we use $u(s) = -\mathcal{C}_{\text{sync}}^+(s)$. (ii) Our algorithm relies on the fact that not two local minima of $u(s)$ can be closer than $2\delta s$. In our case, since s is the differential piston, we expect close minima to be separated by at least the smallest wavelength so we take: $\delta s < (\min_{\ell} \lambda_{\ell})/2$. (iii) A variant of the algorithm is to use a given sampling of the search interval instead of the constant macro step size δs . (iv) The efficiency of the proposed algorithm is that it performs an exhaustive search (without any randomness) and requires at most $1 + \lceil (s_{\text{max}} - s_{\text{min}})/\delta s \rceil + N_{\text{locmin}} \times N_{\text{eval}}$ evaluations of the function (with N_{locmin} the number of local minima in the search interval and N_{eval} the number of function evaluations for Brent's method in a single interval). Typically in our case, $N_{\text{locmin}} \sim (s_{\text{max}} - s_{\text{min}})/(\min_{\ell} \lambda_{\ell})$ and $N_{\text{eval}} \lesssim 20$ although this latter depends on the requested precision. So the total cost scales as $\sim (2 + N_{\text{eval}})(s_{\text{max}} - s_{\text{min}})/(\min_{\ell} \lambda_{\ell})$ and, in any cases, terminates in a finite time and cannot fail (providing δs is adequate). In a future work, we foresee to compare our algorithm to other 1D global optimization algorithms notably GLOMIN by Brent¹⁴ which takes some random steps and STEP by Swarzberg *et al.*¹⁵ which performs an adaptive search but does not exploit the smoothness of the objective function.

3.2.2 Fitting scaled reduced data

If one form the variable:

$$\tilde{r}_{\ell} = \alpha_{\ell, m_0} \tilde{z}_{\ell, m_0} - \alpha_{\ell, m_1} \tilde{z}_{\ell, m_1} e^{-i\kappa_{\ell} s}$$

with $\alpha_{\ell, m}$ given (real or complex) factors and, as before, $s = p_{m_1} - p_{m_0}$, then:

$$\mathbb{E}\{\tilde{r}_{\ell} | \boldsymbol{\theta}\} = (\alpha_{\ell, m_0} |z_{\ell, m_0}| - \alpha_{\ell, m_1} |z_{\ell, m_1}|) e^{i(\phi_{\ell} + \kappa_{\ell} p_{m_0})},$$

and it is easy to show that the real and imaginary parts of \tilde{r}_{ℓ} are i.i.d. Gaussian variables with variance:

$$\text{Var}\{\text{Re}(\tilde{r}_{\ell}) | \boldsymbol{\theta}\} = \text{Var}\{\text{Im}(\tilde{r}_{\ell}) | \boldsymbol{\theta}\} = |\alpha_{\ell, m_0}|^2 \sigma_{\ell, m_0}^2 + |\alpha_{\ell, m_1}|^2 \sigma_{\ell, m_1}^2.$$

Thus, if $\alpha_{\ell, m_0} |z_{\ell, m_0}| \approx \alpha_{\ell, m_1} |z_{\ell, m_1}|$, then \tilde{r}_{ℓ} is approximately a centered complex variable whatever the phase of the object complex visibility and the piston difference s can be estimated by minimizing:

$$s_{\text{scl dif}}^+(\boldsymbol{\alpha}) = \arg \min_s \mathcal{E}_{\text{scl dif}}(s | \boldsymbol{\alpha}) \quad \text{with:} \quad \mathcal{E}_{\text{scl dif}}(s | \boldsymbol{\alpha}) = \sum_{\ell} \frac{|\alpha_{\ell, m_0} \tilde{z}_{\ell, m_0} - \alpha_{\ell, m_1} \tilde{z}_{\ell, m_1} e^{-i\kappa_{\ell} s}|^2}{|\alpha_{\ell, m_0}|^2 \sigma_{\ell, m_0}^2 + |\alpha_{\ell, m_1}|^2 \sigma_{\ell, m_1}^2}, \quad (36)$$

where, here, $\boldsymbol{\alpha} = \{\alpha_{\ell, m} | \ell = 1, \dots, L; m \in \{m_0, m_1\}\}$. Clearly the result only depends on the ratio $\alpha_{\ell, m_1}/\alpha_{\ell, m_0}$ ($\forall \ell$) but we keep our definition for symmetry. Whatever the choice for $\boldsymbol{\alpha}$, solving the problem (36) requires global

Table 1. Summary of the different methods for the cophasing of two frames.

Method	Description
Scaled differences:	
Method #0	$s_{\#0}^+ = \arg \max_s \mathcal{C}_{\text{scl dif}}^+(s \boldsymbol{\alpha})$ given by Eq. (36) with $\alpha_{\ell, m} \equiv 1$
Method #1	$s_{\#1}^+ = \arg \max_s \mathcal{C}_{\text{scl dif}}^+(s \boldsymbol{\alpha})$ given by Eq. (36) with $\alpha_{\ell, m_0} \equiv \tilde{z}_{\ell, m_1} $ and $\alpha_{\ell, m_1} \equiv \tilde{z}_{\ell, m_0} $
Cross product:	
Method #2	$s_{\#2}^+ = \arg \min_s \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha})$ with $\alpha_\ell = \tilde{z}_{\ell, m_0} \tilde{z}_{\ell, m_1} $
Method #3	$s_{\#3}^+ = \arg \min_s \{ \min_{\boldsymbol{\alpha}} \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) \}$
Method #4	$s_{\#4}^+ = \arg \min_s \{ \min_{\boldsymbol{\alpha} \geq 0} \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) \}$
Maximum likelihood:	
Method #5	$s_{\#5}^+ = \arg \max_s \mathcal{C}_{\text{sync}}^+(s)$ given in Eq. (35)

1D optimization which we carry out with the BRADI algorithm described above. In our tests, we considered two implementations of this approach: we neglect the variation of amplitude in the reduced data and take $\alpha_{\ell, m} \equiv 1$ (*Method #0*) or we take $\alpha_{\ell, m_0} \equiv |\tilde{z}_{\ell, m_1}|$ and $\alpha_{\ell, m_1} \equiv |\tilde{z}_{\ell, m_0}|$ (*Method #1*).

3.2.3 Cophasing with the cross-product

In order to get rid of most of the nuisance variables (the phase of the object complex visibility and, partially, the amplitudes of the reduced data), we also considered using the cross product:

$$\tilde{q}_\ell = \tilde{z}_{\ell, m_0}^* \tilde{z}_{\ell, m_1}.$$

Since, for $m_0 \neq m_1$, \tilde{z}_{ℓ, m_0}^* and \tilde{z}_{ℓ, m_1} are statistically independent and owing to their particular distributions, it is easy to compute the expected value and the terms of the covariance matrix of \tilde{q}_ℓ (conditioned to the knowledge of the parameters $\boldsymbol{\theta}$):

$$\begin{aligned} \mathbb{E}\{\tilde{q}_\ell | \boldsymbol{\theta}\} &= \mathbb{E}\{\tilde{z}_{\ell, m_0} | \boldsymbol{\theta}\}^* \mathbb{E}\{\tilde{z}_{\ell, m_1} | \boldsymbol{\theta}\} = \alpha_\ell e^{i \kappa_\ell s}, \\ \text{Var}\{\text{Re}(\tilde{q}_\ell) | \boldsymbol{\theta}\} &= |z_{\ell, m_0}|^2 \sigma_{\ell, m_1}^2 + |z_{\ell, m_1}|^2 \sigma_{\ell, m_0}^2 + 2 \sigma_{\ell, m_0}^2 \sigma_{\ell, m_1}^2, \\ \text{Var}\{\text{Im}(\tilde{q}_\ell) | \boldsymbol{\theta}\} &= \text{Var}\{\text{Re}(\tilde{q}_\ell) | \boldsymbol{\theta}\}, \\ \text{Cov}\{\text{Re}(\tilde{q}_\ell), \text{Im}(\tilde{q}_\ell) | \boldsymbol{\theta}\} &= 0, \end{aligned}$$

with $\alpha_\ell = |z_{\ell, m_0}| |z_{\ell, m_1}|$ and $s = p_{m_1} - p_{m_0}$. Thus the real and imaginary parts of \tilde{q}_ℓ are uncorrelated and have the same variance, even though they are certainly not Gaussian. Nevertheless under a Gaussian approximation, we can fit the cross product \tilde{q}_ℓ simultaneously at all wavelengths (*i.e.* $\forall \ell$) to get an estimator of the differential piston s . Depending on how we estimate $\boldsymbol{\alpha} = \{\alpha_\ell | \ell = 1, \dots, L\}$, we have several possibilities:

$$s_{\#2}^+ = \arg \min_s \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) \quad \text{with} \quad \alpha_\ell = |\tilde{z}_{\ell, m_0}| |\tilde{z}_{\ell, m_1}|, \quad (37a)$$

$$s_{\#3}^+ = \arg \min_s \left\{ \min_{\boldsymbol{\alpha}} \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) \right\}, \quad (37b)$$

$$s_{\#4}^+ = \arg \min_s \left\{ \min_{\boldsymbol{\alpha} \geq 0} \mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) \right\}, \quad (37c)$$

where:

$$\mathcal{C}_{\text{cross}}(s, \boldsymbol{\alpha}) = \sum_{\ell} w_{\ell}^{\text{cross}} \left| \tilde{z}_{\ell, m_0}^* \tilde{z}_{\ell, m_1} - \underbrace{\alpha_\ell e^{i \kappa_\ell s}}_{\mathbb{E}\{\tilde{q}_\ell | \boldsymbol{\theta}\}} \right|^2, \quad (37d)$$

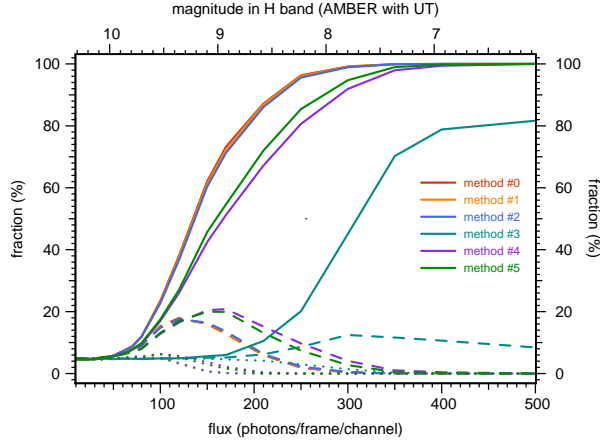


Figure 4. Fraction of frames correctly aligned as a function of the flux for different cophasing methods (see Table 1). With ε the error for the differential piston: $|\varepsilon| < \lambda/2$ (solid lines), $\lambda/2 \leq |\varepsilon| < \lambda$ (dashed lines) and $\lambda \leq |\varepsilon| < 3\lambda/2$ (dotted lines).

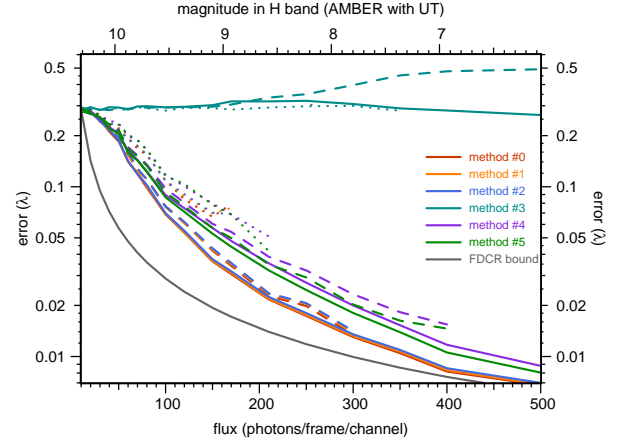


Figure 5. Precision of the cophasing as a function of the flux for different cophasing methods and ranges for the absolute error $|\varepsilon|$ (see Fig. 4). The error is measured with respect to nearest multiple of the effective wavelength λ . The best theoretically achievable accuracy is indicated by the gray curve.

with the statistical weights approximated from the variance of \tilde{z}_ℓ by replacing unknown terms, namely $|z_{\ell,m_0}|^2$ and $|z_{\ell,m_1}|^2$, by their empirical value[‡]:

$$w_\ell^{\text{cross}} = \frac{w_{\ell,m_0} w_{\ell,m_1}}{w_{\ell,m_0} |z_{\ell,m_0}|^2 + w_{\ell,m_1} |z_{\ell,m_1}|^2 + 2} \approx \frac{w_{\ell,m_0} w_{\ell,m_1}}{w_{\ell,m_0} |\tilde{z}_{\ell,m_0}|^2 + w_{\ell,m_1} |\tilde{z}_{\ell,m_1}|^2 + 2}. \quad (37e)$$

The optimal value of α given s is trivial to compute (the problem being quadratic and separable with respect to the spectral channel):

$$\begin{aligned} \arg \min_{\alpha_\ell} \mathcal{C}_{\text{cross}}(s, \alpha) &= \text{Re}(\tilde{z}_{\ell,m_0}^* \tilde{z}_{\ell,m_1} e^{-i\kappa_\ell s}), \\ \arg \min_{\alpha_\ell \geq 0} \mathcal{C}_{\text{cross}}(s, \alpha) &= \max\{0, \text{Re}(\tilde{z}_{\ell,m_0}^* \tilde{z}_{\ell,m_1} e^{-i\kappa_\ell s})\}. \end{aligned}$$

Though this is not the case for s , again we use BRADI for the related 1D global optimization.

3.2.4 Theoretical limit

The best achievable accuracy of an unbiased estimator is provided by the Fréchet-Darmois-Cramér-Rao (FDCR) bound. Assuming all other parameters are known, the variance of any unbiased estimator \hat{s} of the differential piston s is bounded by:

$$\text{Var}\{\hat{s}\} \geq \mathbb{E} \left\{ \frac{\partial^2 \mathcal{L}(\tilde{\mathbf{z}} | \boldsymbol{\theta})}{\partial s^2} \middle| \boldsymbol{\theta} \right\}^{-1} = \left\{ \sum_\ell \kappa_\ell^2 \frac{\text{SNR}_{\ell,m_0}^2 \text{SNR}_{\ell,m_1}^2}{\text{SNR}_{\ell,m_0}^2 + \text{SNR}_{\ell,m_1}^2} \right\}^{-1} \quad \text{with} \quad \text{SNR}_{\ell,m} = |z_{\ell,m}| / \sigma_{\ell,m}, \quad (38)$$

with $\mathcal{L}(\tilde{\mathbf{z}} | \boldsymbol{\theta}) \stackrel{\text{def}}{=} \text{const} - \log \text{Pr}\{\tilde{\mathbf{z}} | \boldsymbol{\theta}\}$ the co-log-likelihood of the data $\tilde{\mathbf{z}}$ given the parameters $\boldsymbol{\theta}$ and $\text{SNR}_{\ell,m}$ the signal to noise ratio of the raw visibility in ℓ -th channel during m -th frame as given by Eq. (21) or by Eq. (22) depending whether maximum likelihood or least squares raw visibilities are used.

3.2.5 Comparison of pairwise cophasing methods

The proposed 6 cophasing methods for a pair of frames are summarized in Table 1. We compared these methods for an intermediate fringe contrast ($\rho = 0.3$) and under various fluxes. Figure 4 shows the fraction of pairs of

[‡]however this slightly overestimate the denominator of w_ℓ^{cross} , notably at low SNR's, because of a positive bias term and thus underestimate the weights

frames which are correctly cophased, or within one or two wavelengths. Note that if the ratio $\Delta\lambda/\lambda$ (with $\Delta\lambda$ the total spectral bandwidth) is small (usually $\Delta\lambda/\lambda \sim 0.1$), then phasing the frames up to a $k\lambda$ offset, with k a small integer in magnitude, can be very effective. Figure 5 shows the achieved accuracy. As expected, as the number of photons decreases, the fraction of misregistered frames augments and the error worsens. Methods #0, #1 and #2 behave the best, then come the maximum likelihood method #5 and method #4 (fitting the cross-product with non-negative amplitude). Method #3 is grossly inefficient which is not unexpected as it does not enforces the non-negativity of the amplitude parameter α_ℓ and thus yields a lot of spurious minima, roughly at multiple of $\lambda/2$ with respect to the true value. The effectiveness of positivity constraints has also been demonstrated for detection methods.¹⁶

In our simulations, method #0 uses the exact scaling factors (*i.e.* $\alpha = 1$) so its performances should be considered as the ultimate ones, at least for the scaled difference criterion in Eq. (36). Methods #1 and #2 which nearly achieve the same best performances use empirical estimators for the scale factors and for the amplitude α . Methods #3, #4 and #5 fit the value of α so as to optimize the corresponding criterion and behave notably worse (in particular method #4) than the best methods. In fact looking closely at Eq. (37d) and Eq. (36) when α is replaced by its empirical value, it appears that these criteria can both be rewritten as weighted sums of squared phasors differences:

$$\mathcal{E}_{\text{phasor}}(s) = \sum_{\ell} w_{\ell, m_0, m_1}^{\text{phasor}} |\tilde{u}_{\ell, m_1} - \tilde{u}_{\ell, m_0} e^{i\kappa_{\ell} s}|^2 \quad \text{with} \quad \tilde{u}_{\ell, m} \stackrel{\text{def}}{=} \tilde{z}_{\ell, m} / |\tilde{z}_{\ell, m}|, \quad (39a)$$

where only the expression of the weights change between method #1 and method #2:

$$w_{\ell, m_0, m_1}^{\text{phasor}} = \frac{w_{\ell, m_0}^{\text{phasor}} w_{\ell, m_1}^{\text{phasor}}}{w_{\ell, m_0}^{\text{phasor}} + w_{\ell, m_1}^{\text{phasor}} + \beta} \quad \text{with} \quad w_{\ell, m}^{\text{phasor}} \stackrel{\text{def}}{=} w_{\ell, m} |\tilde{z}_{\ell, m}|^2, \quad (39b)$$

where $\beta = 0$ for the scaled differences (method #1) and $\beta = 2$ for the cross product (method #2). A future direction of investigation could be to derive improved weighting for the criterion in Eq. (39a). The results of our tests however indicate that this is not a critical issue. Comparing the performances of the proposed methods to optimal statistical bounds such as the Fréchet-Darmonis-Cramér-Rao (FCDR) bound (the gray curve in Fig. 5) shows that the ultimate performances are only achieved for the highest fluxes considered ($\gtrsim 500$ photons/channel/frame). Note that for the lowest considered fluxes the measured accuracy is biased by the modulo λ wrapping so it cannot be compared to the FCDR limit and it tends to $\lambda/\sqrt{12} \simeq 0.28\lambda$, *i.e.* the standard deviation of a uniform distribution on $[-\lambda/2, +\lambda/2]$. The expression in Eq. (39b) motivates the use of $w'_{\ell, m} = w_{\ell, m} |\tilde{z}_{\ell, m}|^2$ as the statistical weight of an empirical phasor.

4. CONCLUSION AND PERSPECTIVES

The best methods for cophasing two given frames amount to fit the raw phasors (not the raw complex visibilities). From this preliminary study, phase tracking seems possible with less than 100 photons/channel/frame. The FCDR bound is however not achieved except for high fluxes (≥ 400 photons/channel/frame).

Further studies include: (i) Design a global strategy for cophasing all the frames of a sequence. (ii) Deal with the fluctuations of the fringe contrast (due to speckle noise). (iii) Apply our method to real data.

ACKNOWLEDGMENTS

This work has been supported by MiTiV (ANR-09-EMER-008), POLCA (ANR-10-BLAN-0511) and OPTICON projects.

REFERENCES

- [1] Jaffe, W. J., “Coherent fringe tracking and visibility estimation for midi,” in [*Proc. SPIE 5491, New Frontiers in Stellar Interferometry*], 715 (2004). 1

- [2] Tatulli, E. and LeBouquin, J.-B., “Comparison of fourier and model-based estimators in single-mode multi-axial interferometry,” *Monthly Notices of the Royal Astronomical Society* **368**(3), 1159–1168 (2006). 1, 3
- [3] Tatulli, E., Millour, F., Chelli, A., Duvert, G., Acke, B., Hernandez Utrera, O., Hofmann, K.-H., Kraus, S., Malbet, F., Mège, P., Petrov, R. G., Vannier, M., Zins, G., Antonelli, P., Beckmann, U., Bresson, Y., Dugué, M., Gennari, S., Glück, L., Kern, P., Lagarde, S., Le Coarer, E., Lisi, F., Perraut, K., Puget, P., Rantakyö, F., Robbe-Dubois, S., Roussel, A., Weigelt, G., Accardo, M., Agabi, K., Altariba, E., Arezki, B., Aristidi, E., Baffa, C., Behrend, J., Blöcker, T., Bonhomme, S., Busoni, S., Cassaing, F., Clause, J.-M., Colin, J., Connot, C., Delboulbé, A., Domiciano de Souza, A., Driebe, T., Feautrier, P., Ferruzzi, D., Forveille, T., Fossat, E., Foy, R., Fraix-Burnet, D., Gallardo, A., Giani, E., Gil, C., Glentzlin, A., Heiden, M., Heininger, M., Kamm, D., Kiekebusch, M., Le Contel, D., Le Contel, J.-M., Lesourd, T., Lopez, B., Lopez, M., Magnard, Y., Marconi, A., Mars, G., Martinot-Lagarde, G., Mathias, P., Monin, J.-L., Mouillet, D., Mourard, D., Nussbaum, E., Ohnaka, K., Pacheco, J., Perrier, C., Rabbia, Y., Rebattu, S., Reynaud, F., Richichi, A., Robini, A., Sacchettini, M., Schertl, D., Schöller, M., Solscheid, W., Spang, A., Stee, P., Stefanini, P., Tallon, M., Tallon-Bosc, I., Tasso, D., Testi, L., Vakili, F., von der Lühse, O., Valtier, J.-C., and Ventura, N., “Interferometric data reduction with AMBER/VLTI. principle, estimators, and illustration,” *Astron. Astrophys.* **464**, 29–42 (2007). 1, 3, 6
- [4] Gordon, J. A. and Buscher, D. F., “Detection noise bias and variance in the power spectrum and bispectrum in optical interferometry,” *Astron. Astrophys.* **541**, A46 (May 2012). 1
- [5] Foi, A., Trimeche, M., Katkovnik, V., and Egiazarian, K., “Practical poissonian-gaussian noise modeling and fitting for single-image raw-data,” *IEEE Transactions on Image Processing* **17**, 1737–1754 (2008). 2
- [6] Trouvé, P., Champagnat, F., Le Besnerais, G., and Idier, J., “Single image local blur identification,” in [*18th IEEE International Conference on Image Processing*], 621–624, Institute of Electrical and Electronics Engineers (Sep 2011). 3
- [7] Hestenes, M. R. and Stiefel, E., “Methods of conjugate gradients for solving linear systems,” *Journal of Research of the National Bureau of Standards* **49**(6), 409–436 (1952). 3
- [8] Béchet, C., Tallon, M., and Thiébaud, É., “Comparison of minimum-norm maximum likelihood and maximum a posteriori wavefront reconstructions for large adaptive optics systems,” *J. Opt. Soc. America A* **26**(3), 497–508 (2009). 3
- [9] Kay, S. M., [*Fundamentals of Statistical Signal Processing: Detection Theory*], vol. 2 of *Signal Processing Series*, Prentice-Hall (1998). 3
- [10] Kay, S. M., [*Fundamentals of Statistical Signal Processing: Estimation Theory*], vol. 1 of *Signal Processing Series*, Prentice-Hall (1998). 3
- [11] Goodman, J. W., [*Statistical Optics*], John Wiley & Sons (1985). 5
- [12] Thiébaud, É., *Imagerie astrophysique à la limite de diffraction des grands télescopes. Application à l’observation des objets froids.*, PhD thesis, Université de Paris 7 (1994). 7
- [13] Gratadour, D., Mugnier, L., and Rouan, D., “Sub-pixel image registration with a maximum likelihood estimator. application to the first adaptive optics observations of Arp 220 in the L’ band,” *Astron. Astrophys.* **443**, 357–365 (2005). 8
- [14] Brent, R., [*Algorithms for Minimization without Derivatives*], Englewood Cliffs, NJ: Prentice-Hall (1973). 9, 10
- [15] Swarzberg, S., Seront, G., and Bersini, H., “S.T.E.P.: the easiest way to optimize a function,” *IEEE World Congress on Computational Intelligence, Proceedings of the First IEEE Conference on Evolutionary Computation* **1**, 519–524 (1994). 10
- [16] Thiébaud, É. and Mugnier, L., “Maximum a posteriori planet detection and characterization with a nulling interferometer,” in [*IAU Colloq. 200: Direct Imaging of Exoplanets: Science & Techniques*], Aime, C. and Vakili, F., eds., 547–552, Cambridge University Press (2006). 13